Mapping the Architecture of Protein Complexes in *Arabidopsis* Using Cross-Linking Mass Spectrometry

Cao Son Trinh[1,2]*, Ruben Shrestha[1]*, William C Conner[1], Andres V Reyes[1,2], Sumudu S Karunadasa[1,2], Grace Liu[1], Ken Hu[1], Shou-Ling Xu[1,2]

[1]Department of Plant Biology, Carnegie Institution for Science, Stanford, California, USA; [2]Carnegie Mass Spectrometry Facility, Carnegie Institution for Science, Stanford, California, USA

*These authors contributed equally.

Corresponding author: Shou-Ling Xu, slxu@stanford.edu

Classification: Research Reports

Keywords: cross-linking mass spectrometry (XL-MS), structural analysis, protein-protein interaction network mapping, gene annotation.

Abstract

Capturing molecular machines in action is essential to understanding the architecture of protein complexes, cellular regulation, and gene function. Here, we present a robust cross-linking mass spectrometry (XL-MS) platform that uses the PhoX cross-linker to map proteome-wide interactions in *Arabidopsis thaliana* under semi-native conditions. Using whole-cell lysates, chloroplasts, and nuclei, we identified 47,119 unique cross-links, including 3,527 inter-protein cross-links. We then constructed a high-confidence protein-protein interaction (PPI) network containing 1,229 proteins and 1,446 PPIs. In silico validation using STRING confirmed 637 heteromeric interactions with a combined confidence score of at least 0.4, including 532 interactions with a score of at least 0.7. The remaining interactions are novel. Our dataset provides direct experimental evidence of physical interactions and defines residue-level binding interfaces. XL-MS resolved the spatial topology of structurally challenging protein complexes, such as the chloroplast 70S ribosome stalks, and transient intermediates involved in photosystem repair. XL-MS also revealed both conserved and distinct interaction patterns across photosystem complexes, ribosomes, and chromatin-associated proteins. Notably, the data suggest that the ribosome undergoes ubiquitination and is regulated by the 26S proteasome. This finding is further supported by *in vivo* diGly proteomic analysis. Among the novel interactions uncovered is a metabolite enzyme that interacts with histones, which could potentially contribute to histone modification. This work establishes XL-MS as a powerful tool for advancing the functional annotation of genomes and the understanding of the architecture of protein complexes and cellular regulation in plants, providing a valuable resource for the plant research community.

35 **Significance Statement**

36 Understanding how proteins interact within living cells is essential to deciphering cellular
37 architecture and function. However, capturing native protein-protein interactions (PPIs) on a
38 global scale has proven technically challenging. Here, we present a proteome-wide cross-linking
39 mass spectrometry (XL-MS) platform that can systematically map direct PPIs in plant cells without
40 requiring transgenic manipulation. This approach reveals thousands of interactions spanning major
41 subcellular compartments and resolves the architecture of critical protein assemblies, such as
42 photosystems, ribosomes, and chromatin-associated connectivity. We mapped both known and
43 novel interactions. This comprehensive dataset advances protein annotation and provides a
44 valuable resource for investigating the dynamic organization of the plant proteome.

45

Introduction

Understanding how molecular machines operate within cells - and capturing them in action - is essential to unraveling the principles of life and gene function. A powerful tool for this purpose is cross-linking mass spectrometry (XL-MS), an emerging high-throughput approach that enables the capture of protein complexes in their native context across diverse organisms (1, 2). Unlike affinity purification mass spectrometry (AP-MS) and proximity labeling mass spectrometry (PL-MS) (3-5), XL-MS does not require antibodies or genetically tagged transgenes. Instead, it relies on chemical cross-linkers that covalently link spatially proximal residues within or between proteins in solution, followed by protease digestion and mass spectrometric identification of the resulting cross-linked peptides. In addition to generating a "parts list" of interacting proteins, XL-MS provides direct evidence of physical interactions and precisely maps contact sites at the amino acid level, providing unparalleled structural insight into macromolecular assemblies.

Despite its promise, large-scale proteome-wide XL-MS - such as in whole-cell lysates, isolated organelles, or *in vivo* systems - remains challenging due to the inherent low abundance and heterogeneity of cross-linked peptides and the complexity of data analysis (1, 6). In plants, additional barriers further complicate its application, including rigid cell walls and large central vacuoles that dilute cross-linkers, as well as secondary metabolites that can chemically interfere with cross-linking chemistry (7, 8). Recent pioneering studies in Arabidopsis, algae, and soybean have demonstrated the potential of XL-MS in plant systems (7, 9, 10), but its overall implementation remains limited.

To address these challenges, we developed a cross-linking strategy optimized for proteome-wide mapping in complex Arabidopsis samples using the enrichable cross-linker PhoX. PhoX is a phosphonic acid-based reagent that, like BS3 and DSSO, contains NHS esters and

69 primarily targets lysine residues (11). Its phosphonic acid handle allows efficient enrichment by
70 immobilized metal affinity chromatography (IMAC). With a maximum range of 35Å - consisting
71 of a 20 Å Cα-Cα distance from its rigid structure and an additional 15 Å to account for protein
72 flexibility - PhoX is well suited for high confidence, large-scale structural proteomics (12). To
73 reduce sample complexity and improve cross-linking efficiency, we used subcellular
74 fractionations. To further improve the detection of the low abundance cross-linked peptides, we
75 performed extensive high-pH reversed-phase HPLC fractionation after IMAC enrichment.
76 Together, these measures allowed for comprehensive proteome-wide XL-MS mapping.

77 We identified a total of 47,119 unique cross-linked peptide pairs from cross-linked cell
78 lysates, chloroplasts, and nuclei, including 43,592 intra-protein (self-link) and 3,527 inter-protein
79 cross-links. These results enabled us to construct a PPI network comprising 1,229 proteins and
80 1,446 interactions, of which 1,390 are heteromeric and 56 are homomultimeric. STRING analysis,
81 which catalogs only heteromeric PPIs, validated 637 interactions with a combined confidence
82 score of at least 0.4. The rest represent previously uncharacterized interactions. Our dataset
83 provides experimental evidence of direct physical binding and defines protein binding interfaces
84 with amino acid resolution for both known and novel interactions.

85 Beyond interaction mapping, our XL-MS data provide structural and functional insights
86 into key cellular processes. By resolving flexible or transient components within the 70S
87 chloroplast ribosome and photosystem complexes, our results complement existing cryo-EM
88 structures. XL-MS also enabled a comprehensive mapping of the photosystem, 80S ribosome, and
89 histone interactomes within heterogeneous protein populations, revealing structural biology and
90 capturing both known and novel interactions. We identified a metabolic O-acyltransferase that
91 interacts with histones, suggesting a potential new mechanism of chromatin regulation via post-

translational modification. Together, these results improve protein annotation, shed light on protein complex architectures, and provide as a valuable resource for the broader research community - laying the foundation for future discoveries.

Results

Efficient cross-linking workflow for complex protein samples identifies numerous cross-links in Arabidopsis

To gain structural insights and visualize molecular machines in their native cellular context within Arabidopsis cells, we performed XL-MS (see Methods) on fractionated seedling samples, including cell lysates and chloroplasts from light-grown seedlings and nuclei from dark-grown seedlings. Using the PhoX cross-linker (Fig. 1A), we established a robust pipeline optimized for proteome-wide detection of low-abundance cross-linked peptides in complex Arabidopsis samples.

Optimization of the PhoX concentration showed a shift to higher molecular weights at 2-5 mM (red arrows, Fig. S1A-D), indicating successful cross-linking, with no additional improvement at higher concentrations. Based on these results, we used 2 and 5 mM PhoX for subsequent experiments (see Methods). After cross-linking, we precipitated the proteins with phenol, digested them with trypsin, and enriched the cross-linked peptides using $GaCl_3$-based IMAC (13). We then fractionated the enriched peptides by high-pH reversed-phase chromatography, analyzed them by LC-MS/MS, and identified cross-links using pLINK2 (14). High-resolution, high-accuracy data for both MS1 (precursor ions) and MS2 (fragment ions) were acquired on an Orbitrap Eclipse mass spectrometer to ensure reliable cross-link identification.

To optimize data acquisition, we compared the standard XL workflow with real-time spectral library search (RTLS) (15), both targeting highly charged precursors (+3 and above).

115 RTLS uses a survey MS2 scan to reduce mono-link detection. In cell lysates, RTLS reduced mono-
116 and loop-link identifications by ~50% and ~31%, respectively (Fig. S2), but also significantly
117 reduced intra-link (43%) and inter-link (40%) identifications. Because of this trade-off, we used
118 the standard XL workflow to analyze cross-linked chloroplast and nuclear samples.

119 In total, we identified 47,119 unique cross-linked peptide pairs (intra- and inter-links)
120 representing 5,759 proteins. Specifically, we detected 25,795, 17,090, and 19,612 cross-links from
121 the nucleus, chloroplast, and cell lysate, respectively (Fig. 1B and. S3, Table S1). Subcellular
122 localization analysis using SUBA5 revealed that the cross-linked proteins were distributed in
123 various cellular compartments, including the nucleus, plastid, cytosol, mitochondria and other
124 organelles (Fig. 1C, Table S2).

125 To assess the quality of cross-linked peptide identification, we examined the identified
126 cross-links (Table S1) and found consistent peptide pairs mapping to the same sites or regions. For
127 example, in ribosomal subunits, we detected site-specific cross-links with variable miscleavages,
128 such as those between uL15x-K92 and eL13z-K166/K172 (Fig. S4A). Conserved region cross-
129 links were also observed, such as VHA-E3-K49 with both VHA-B1-K437 and VHA-B2-K438
130 (Fig. S4B), and TOP1A-K672 with ERD14-K38 and COR47-K46 (Fig. S4C). These independent
131 and consistent observations underscore the robustness and reliability of our results.

132 **Validation of intra- and inter-protein cross-links by structural mapping**

133 To evaluate the validity of the cross-links, we first used ComMap (16) to map intra-links to high-
134 resolution protein structures available in the Protein Data Bank (PDB) and predicted structures by
135 AlphaFold. Specifically, 721 Arabidopsis proteins had high-resolution structures in the PDB, and
136 3,975 were available through AlphaFold (Table S1). A total of 21,981 cross-links were
137 successfully mapped, with 94% of the PDB-mapped cross-links and 89% of the AlphaFold-
138 mapped cross-links satisfying the 35 Å distance constraint (Fig. 2A-B).

139 To evaluate both intra- and inter-protein cross-links, we mapped 174 cross-links onto the
140 L_8S_8 Rubisco assembly modeled from the Arabidopsis L_2S_2 crystal structure (PDB:5IU0) (17). Of
141 these, 169 cross-links were mapped to the large (479 aa) and small (181 aa) subunits, while five
142 could not be mapped because the PDB structure lacks the N-terminal region (residues 1–11) (Fig.
143 2C). The majority (126, 74.6%) of mapped cross-links fell within the expected Cα-Cα distance of
144 ≤35 Å; the remaining 43 (25.4%) exceeded this range, suggesting regions of structural flexibility
145 (Fig. 2D–E). Similar mapping results were obtained using the hornwort Rubisco cryo-EM structure
146 (Fig. S5A-B), further supporting the validity of our cross-linking data.

147 To assess whether cross-linked lysine residues are positioned within predicted interaction
148 interfaces, we examined several PPIs using available structural data. The cross-link between
149 ACINUS-K633 and SR45-K154 maps near the RSB domain of ACINUS (aa 599-631) and within
150 the RRM domain of SR45 in the ASAP core complex (PDB:4A8X) (18-20) (Fig. 2F, Fig. S6A).
151 Similarly, the cross-link between SR45-K154 and PININ-K274 aligns with the same RRM domain
152 and the predicted RSB domain of PININ (Fig. 2F, Fig. S6B). For PRK-K217 and CP12.1-K65, the
153 cross-link localizes to the interaction interface in the GAPDH-CP12-PRK complex (PDB:6KEZ)
154 (21) with a Cα-Cα distance of 13.3 Å (Fig. 2G-H). These results confirm that the cross-linked
155 residues are located at or within a few amino acids of the interaction interfaces, consistent with
156 previous observations (22).

157 **XL-MS systematically identifies direct protein-protein interactions and reveals known and**
158 **novel complexes**

159 We imported 3,527 inter-protein cross-links into Cytoscape (23), revealing 1,446 PPIs
160 among 1,229 proteins, including 1,390 heteromeric and 56 homomultimeric interactions (Fig. 3A,
161 Table S1). Proteins were classified using the PANTHER protein class tool, and nodes were

162 colored accordingly (24). Many interacting proteins shared the same classification, supporting the
163 reliability of our PPI map. Notably, 457 of the 1,210 proteins involved in heteromeric interactions
164 lacked PANTHER annotations, contributing to 819 interactions - 369 between unclassified
165 proteins and 450 between classified and unclassified proteins.

166 We also incorporated STRING confidence scores (25) into the network and color-coded
167 the edges accordingly (Fig. 3A). These scores combine evidence from eight sources (Fig.3B) (25).
168 For heteromeric PPIs, 38.2% of PPIs had high STRING confidence: 388 scored ≥ 0.9 and 144
169 scored 0.7-0.9. The remainder included 105 medium (0.4-0.7), 144 low (<0.4), and 609 with no
170 STRING score (Fig.3B). The PANTHER and STRING results indicate that both known and novel
171 PPIs were captured, highlighting the power of systematic XL-MS to reveal previously
172 uncharacterized protein complexes and pathways.

173 Even among PPIs with good STRING combined scores, many lack experimental support
174 for direct or indirect interactions in Arabidopsis or homologs in other species. Among the
175 heteromeric PPIs identified, only 28.9% (n = 402) have experimental support (score ≥ 0.4), while
176 71.1% have low (<0.4) or no scores (Fig. 3B). Our XL-MS data provide evidence for their direct
177 interactions and contact sites. For example, ORP3B (AT3G09300, oxysterol-binding protein) and
178 PLDRP1 (AT5G39570, PLD-regulated protein 1) have a combined score of 0.602 but a low
179 experimental score (0.055); we identified a cross-link between ORP3B-K7 and PLDRP1-K193
180 (Fig. 3C), suggesting a common role in lipid trafficking and signaling. Similarly, NRPA2
181 (AT1G29940, RNA polymerase predicted to localize to both the nucleus and chloroplasts) and
182 uL1c (AT3G63490, chloroplast 50S ribosomal protein) have a combined score of 0.588 but no
183 experimental support; a cross-link between NRPA2-K394 and uL1c-K127 suggests a direct
184 interaction and potential coupling of chloroplast transcription and translation (Fig. 3D).

185 **XL-MS resolves the topology of chloroplast 70S ribosome stalk regions inaccessible by Cryo-**
186 **EM**

187 We mapped the topologies of several protein complexes (Fig. S7A-G), including the
188 flexible stalk regions of the chloroplast 70S ribosome (Fig. 4A-C), which have remained difficult
189 to resolve due to the dynamic nature of proteins such as uL1c, uL10c, uL11c, and bL12c (26, 27).
190 In total, we identified 424 cross-links across 34 ribosomal subunits of the 70S ribosome, including
191 352 intraprotein, 64 heteromeric, and 8 homomultimeric cross-links. Notably, many intra- and
192 interprotein cross-links involved these flexible stalk proteins. When cross-links were mapped to
193 spinach 70S ribosome structures (PDB: 5MMM and 5X8P; Fig. S8A-B), most cross-links
194 involving uL1c, uL10c, uL11c, and bL12c could not be precisely assigned to specific locations
195 within the stalk regions because these proteins are either missing or only partially resolved in the
196 cryo-EM structures.

197 Arabidopsis bL12c has a plant-specific disordered N-terminus, followed by conserved
198 structured N-terminal (NTD) and C-terminal (CTD) domains connected by a flexible hinge (Fig.
199 4B). Essential for translation, bL12c uses hinge-mediated mobility to recruit specific translation
200 factors(28). XL-MS identified interactions between bL12c (the two highly homologous proteins
201 AT3G27850/bL12cx and AT3G27830/bL12cz) and several ribosomal core proteins, as well as the
202 elongation factor CPEF-G, mainly through its structured CTD (Fig. 4C). Although bL12c dimers
203 are typically mediated by the NTD (29), XL-MS revealed unexpected homodimer cross-links
204 within the CTD - the same region involved in translation factor binding (Fig. 4C). This suggests
205 that CTD dimerization may serve as an interface for the recruitment of interaction partners. Taken
206 together, these results highlight both conserved and plant-specific features of bL12c and
207 demonstrate the power of XL-MS in resolving dynamic aspects of ribosomal architecture.

208 **XL-MS reveals dynamic connectivity and unique features in photosystem complexes**

209 Photosystems I and II (PSI and PSII) and their associated light-harvesting complexes
210 (LHCI and LHCII) are essential for light energy harvesting and carbon fixation, but they require
211 continuous repair, especially under high and fluctuating light conditions (30-32). Although
212 progress has been made in understanding certain steps of the PSII repair (33, 34), many aspects,
213 including the detailed mechanism, identification of repair intermediates, and specific contributing
214 factors, remain unresolved.

215 Using XL-MS, we identified extensive interactions involving the PSII repair factor Psb27
216 (AT1G03600) (Fig. 5A). Unlike its cyanobacterial counterpart, which functions in both assembly
217 and repair, Arabidopsis Psb27 appears to be specialized for repair. In contrast, the less homologous
218 Psb27-H1(AT1G05385) is involved in assembly. Notably, Arabidopsis Psb27 lacks a lipid anchor
219 and is structurally more flexible than the cyanobacterial version (35) (Fig. S9A). A partial structure
220 from Chlamydomonas (PDB: 6KAC) captured residues 91-114 of Psb27 (36). Our XL-MS data
221 revealed that Psb27 interacts with both the PSII reaction center and the oxygen-evolving complex
222 (OEC), involving K93 (as in 6KAC) and a second site at K168 (Fig. 5A, Fig. S9B, S10A-H). These
223 distance constraints (Fig. S9B) suggest that Psb27 engages PSII in a more dynamic and flexible
224 manner in plants.

225 XL-MS also revealed extensive interactions between PSI, PSII, LHCI, LHCII, ATP
226 synthase, ATP-dependent proteases (FtsH, SPPA) and transient proteins involved in state
227 transitions, complex assembly and repair (Fig. 5B). Notably, chlorophyll biosynthesis and
228 catabolic proteins (e.g., ACCELERATED CELL DEATH 1 (ACD1), magnesium protoporphyrin
229 IX O-methyltransferase (CHLM), FLUORESCENT IN BLUE LIGHT (FLU)) were found to
230 directly associate with LHCII, suggesting coordinated chlorophyll delivery and reactive oxygen

231 species (ROS) control during LHCII assembly. In addition, many assembly and repair factors were
232 found to associate with LHCII, indicating their readiness for rapid engagement during dynamic
233 photosystem remodeling.

234 **XL-MS of ribosome complexes reveals conserved ribosome-associated factors**

235 We identified extensive interactions involving the 80S ribosome, a complex of 60S and
236 40S subunits that synthesizes nuclear-encoded proteins. Ribosomes assemble in the nucleolus,
237 localize to the cytoplasm, and associate with the endoplasmic reticulum (ER). As a dynamic
238 macromolecular machine, the ribosome coordinates translation with numerous proteins and RNAs.
239 Errors in translation that produce misfolded proteins, or defects in the ribosome itself, can trigger
240 ribosome-associated quality control pathways (37). Using XL-MS, we mapped detailed structural
241 information within this highly heterogeneous complex, capturing not only intra- and inter-subunit
242 contacts (Fig. S11A-D), but also interactions with proteins involved in ribosome biogenesis,
243 translation, folding, quality control, stress response, and other regulatory pathways (Fig. 6A).
244 These interactions were resolved at the amino acid level, providing detailed insight into the
245 architecture of the functional interface (Fig. 6B, Table S4).

246 While many interactions showed good STRING scores (≥0.4), our analysis also revealed
247 numerous novel interactions (green edges) between specific protein pairs. Although homologs of
248 these proteins have been studied in other organisms for their ribosome-related functions, their
249 specific roles in plants remain largely uncharacterized. Our XL-MS data suggest conserved
250 functions: for example, the UTP7 homologs AT3G10530 and NOG1-1 may contribute to ribosome
251 biogenesis (38, 39), whereas AT5G50840 (α-taxilin-like) and ILITHYIA(ILA) may be involved
252 in protein folding or translation (40, 41). MBF1A, a transcriptional coactivator, may be involved
253 in stress responses at stalled ribosomes, similar to its role in yeast (42) . Other candidates include

254 a methyltransferase potentially involved in rRNA processing and polysome assembly (43, 44) , a
255 RIO kinase associated with ribosome biogenesis (45, 46), and PARP1, which may regulate
256 translation via ADP-ribosylation (47) (Fig. 6C). By mapping these interactions to specific
257 ribosomal proteins, our XL-MS dataset provides a valuable resource for future functional studies
258 of ribosome-associated factors in plants.

259 We identified several interactors related to ubiquitination, including two E3 ligases (EBF1
260 and AT4G02760) and one deubiquitinase UBP14. In addition, we detected direct interactions
261 between ribosomes and 26S proteasome components, such as RPN8A and PAD1 (Fig.6A). Given
262 the need for active ribosome turnover to maintain quality control and these XL interaction data,
263 we hypothesized that ribosome proteins undergo *in vivo* ubiquitination as a regulatory mechanism.
264 To test this hypothesis, we conducted a diGly ubiquitin proteomic analysis (48), which revealed
265 the *in vivo* ubiquitination of 37 large and 27 small ribosomal subunits (Fig. 6A, Table S4). These
266 findings suggest that the ribosomes are ubiquitinated *in vivo* and may be targeted for degradation
267 by the 26S proteosome.

268 **Histone interactome reveals conserved and novel chromatin-associated interactions**

269 Our XL-MS analysis revealed extensive interactions between histones and both known and
270 novel interactors, providing new insights into the networks involved in gene expression and
271 chromatin regulation (Fig. 7A). We identified several nuclear complexes associated with
272 chromatin, including nucleosome assembly proteins (NAPs), chromosome structure maintenance
273 proteins (SMCs), nuclear RNA polymerase components (NRPs), chromatin regulators (PWWP
274 DOMAIN PROTEIN 3(PDP3), MULTICOPY SUPPRESSOR OF IRA1 4(MSI4)), transcriptional
275 activator and acetylation reader (AT2G36410 and ACETYLATED INTERACTING PROTEIN
276 1(ACIP1)), and subunits of the PAF complex (VERNALIZATION INDEPENDENCE 2 and 4
277 (VIP2, VIP4)) (Fig. 7A).

278 Although many proteins are known to influence chromatin architecture and gene regulation
279 in plants, their direct interactions with histones and the nature of their contact sites remain largely
280 uncharacterized. Using XL-MS, we identified a diverse array of proteins that interact with histone,
281 including chaperones such as SUVH1/3-interacting DNAJ domain-containing protein 2(SDJ2) and
282 NUCLEOLIN 1(NUC-L1)), chromatin remodeling factors such as EARLY BOLTING IN SHORT
283 DAYS(EBS), SPLAYED(SYD) and SPIKE1(SPK1)), transcription factors such as SQUAMOSA
284 PROMOTER BINDING PROTEIN-LIKE 3(SPL3), co-regulators such as SIN3-LIKE 6(SNL6),
285 DNA repair proteins such as MUTS HOMOLOG 2 (MSH2), and structural proteins such as HIGH-
286 MOBILITY GROUP BOX 7 (HMGB7). Notably, we found that MOS7, a nucleoporin
287 homologous to the human NUP88, cross-linked with several histones. This provides direct
288 evidence of nuclear envelope-histone interactions that may contribute to chromatin organization
289 in plants (49). In addition, we identified histone interactions with several enzymatic regulators,
290 including the kinase MLK4, the poly(ADP-ribose) polymerase PARP1 (both of which were cross-
291 linked to H1), and the ubiquitin-like protein NEDD8 (which was cross-linked to H2B and H3).
292 Surprisingly, XL-MS identified four distinct cross-links between histone H2B and the O-
293 acyltransferase WSD4 (Fig. 7A and S12A-E). Members of the WSD family, including WSD4, are
294 traditionally associated with lipid metabolism and function as wax ester synthase/acyl-CoA-
295 dependent O-acetyltransferases (50, 51). However, WSD4 is preferentially expressed in roots,
296 which is unusual because wax deposition is uncommon in these tissues (51). This suggests that
297 WSD4 may have an alternative function unrelated to wax biosynthesis. Notably, TAIR predicts
298 that WSD4 is localized in nucleus. To test whether WSD4 co-localizes with histones, we co-
299 expressed mCherry-WSD4 and GFP-HTB11 in *Nicotiana benthamiana*, and observed overlapping
300 nuclear signals, as indicated by the arrows in Fig. 7B. Together, the XL-MS and co-localization

301 data raise the intriguing possibility that WSD4 may catalyze histone acylation, a hypothesis that
302 warrants further investigation (Fig.7C).

303 Taken together, these findings reveal previously uncharacterized histone-associated
304 proteins and enzymatic activities in plants, providing new insights into chromatin biology and
305 laying the groundwork for future studies of plant gene regulation.

306 **Discussion**

307 Systematic mapping of PPIs has greatly advanced our understanding of cellular function,
308 but plant PPI networks remain largely unexplored (5, 52). While comprehensive maps exist for
309 humans and other model organisms (53-55), similar efforts are just beginning in plants (56-58).
310 This gap represents a critical opportunity, as plants are central to addressing global challenges such
311 as food security, climate change, and sustainability. Expanding proteome-wide PPI mapping and
312 structural biology studies in plants will accelerate our mechanistic understanding of cellular
313 processes and drive advances in agriculture and biotechnology.

314 We developed a systems biology approach to capture molecular machines in action by
315 mapping PPIs and their interfaces using XL-MS. To address the inherent challenges of XL-MS in
316 complex biological samples, we established an optimized pipeline that integrates the PhoX
317 crosslinker, subcellular fractionation, and high-pH reversed-phase HPLC offline fractionation.
318 This workflow enabled the identification of over 47,000 crosslinks. The high quality of the data
319 was confirmed by mapping the intralinks onto experimentally determined structures in the PDB
320 and predicted models from AlphaFold.

321 Our XL-MS analysis revealed the topological organization of multiple protein complexes,
322 including the 70S ribosome, 26S proteasome, photosystems I and II, ATP synthases, and
323 chromatin-associated factors. We demonstrate that XL-MS is highly effective in resolving spatial

relationships within dynamic and heterogenous assemblies. Notably, we mapped flexible regions of the 70S ribosome stalk and established connectivity for components that are missing or only partially resolved in the existing cryo-EM structures. This includes flexible elements such as BL12c, for which we discovered unexpected C-terminal domain (CTD) dimerization within the translation factor binding region – suggesting that CTD dimerization may serve as an interface for recruiting interaction partners. Additionally, we mapped interactions involving transient repair proteins in PSII. Together, these results highlight the unique strength of XL-MS in complementing cryo-EM by providing structural insights into conformationally dynamic regions, underscoring the value of our approach in advancing in situ structural biology.

Our XL-MS analysis identified 1,446 high-confidence PPIs among 1,229 proteins, establishing a comprehensive interaction map in *Arabidopsis*. Functional categorization using PANTHER and network analysis via STRING analysis revealed that these interactions span a wide range of biological processes. Notably, 637 heteromeric PPIs had a combined STRING confidence score $\geq$ 0.4, including 532 with score $\geq$ 0.7, underscoring the robustness of our interaction mapping. In addition, we identified 56 homomultimeric PPIs, which are not represented in STRING, as the database catalogs only heteromeric interactions. Recent studies suggest that dimerization is essential for the function of many proteins (59, 60).

Further validation of our dataset is provided by genetic evidence: mutations in several newly identified interacting pairs result in shared or opposing phenotypic outcomes. For example, the chromatin remodeling protein EBS promotes early flowering, whereas its interactor SYD, a component of the chromatin structure remodeling complex, delays flowering when mutated(61, 62). In another example, loss-of-function mutations in both FLK (a flowering locus KH domain protein) and its interactor KHZ1 (a CCCH zinc finger and KH domain protein) result in late

347 flowering (63, 64). These phenotype correlations support the functional relevance of the identified
348 interactions and highlight the potential of XL-MS-based interaction maps to accelerate gene
349 discovery.

350 In addition, the interaction between ribosomal proteins and E3 ligases, deubiquitinases,
351 and 26S proteasome components, such as RPN8A and PAD1, suggests that ribosomes may be
352 regulated by ubiquitin-mediated pathways. Our ubiquitin diGly proteomics analysis supports this
353 hypothesis by identifying the *in vivo* ubiquitination of 37 large and 27 small ribosomal subunits.
354 These identifications were of high confidence, with many di-Gly peptides detected in both light
355 (^{14}N) and heavy (^{15}N) forms (65, 66). Together with the XL-MS data, these findings imply the
356 existence of a potential ribosome quality control mechanism similar to those pathways described
357 in yeast (67). The exact role of E3 ligases and deubiquitinases remains to be confirmed
358 functionally. EBF1 has been shown to regulate ethylene signaling by targeting EIN3/EIL1 for
359 degradation (47); however, its ribosome-related functions have yet to be explored. UBP14 is
360 essential for development in Arabidopsis but dispensable in yeast (48).

361 XL-MS uncovered a wealth of novel PPIs, with over half not previously annotated in the
362 STRING database. For example, we identified both well-established and unreported PPIs
363 involving ribosomal proteins and histones (see Figs. 6A and 7A), expanding our understanding of
364 these conserved molecular machines across species (68, 69). We also uncovered a broad array of
365 plant-specific PPIs, particularly within chloroplasts (Fig. 5B), reflecting unique features of the
366 plant cellular machinery.

367 Among the histone interactors, we identified several enzymatic regulators, including
368 MLK4 and PARP1 (cross-linked to H1), and NEDD8 (cross-linked to H2B and H3). Although
369 PARP1 and NEDD8 are known histone modifiers in other systems, their roles in plants remain

undefined (70-73). Previous studies have shown that MLK4 phosphorylates H2A and regulate flowering (74); thus, it may have broader functions in histone regulation.

We identified WSD4, an O-acyltransferase, as a potential novel histone regulator. Four independent cross-links were identified between WSD4 and H2B, and their interaction was further supported by nuclear co-localization. Although O-acyltransferase are usually linked with lipid metabolism(50), their function in protein modification remains largely unknown. Given the known examples of protein acylation by metabolic enzymes and the regulatory role of lysine acylation(75-78), we propose that WSD4 may catalyze histone acylation. This would reveal a previously unrecognized chromatin regulatory mechanism.

Despite decades of study, many *Arabidopsis* genes remain uncharacterized (79). Our work demonstrates that XL-MS efficiently maps PPIs under semi-native conditions, accelerating gene discovery and functional annotation. In addition, the extensive intra-protein cross-links generated in this study can aid in developing structural prediction tools such as AlphaFold. This dataset is a valuable resource for Arabidopsis functional genomics, and this approach can be applied to other species, particularly crops. Further developing this technique will advance comparative proteomics and deepen our understanding of fundamental biological processes.

Acknowledgement

This work was supported by the National Institutes of Health (NIGMS R01GM135706 and S10OD030441 to S-L.X.) and by the Carnegie endowment fund for the Carnegie Mass Spectrometry Facility. We thank Prof. Laura Helen Gunn for providing the structure of Arabidopsis Rubisco L8S8, and Drs. Mike Trnka and Robert Chalkley (UCSF) and Dr. Rosa Viner (Thermo Fisher Scientific) for technical assistance. We also thank the Hong Qiao lab (UT Austin) for the nuclear isolation protocol and the Adrien Burlacot lab for the chloroplast isolation protocol.

We thank Dr. Ajeet Chaudhary and Prof. Arthur Grossman for their critical reading of the manuscript.

Author contributions

S.L.X conceived the project and supervised the study. C.S.T. and R.S. designed and performed most of the XL-MS experiments. R.S. developed the initial XL workflow. W.C.C performed chloroplast XL-MS. S.K. performed ubiquitination enrichment. C.S.T and S.L.X analyzed all the MS data and prepared figures, with assistance from A.V.R, G.L. and K.H. C.S.T and S.L.X interpreted the data and wrote the manuscript with feedback from all authors.

Data Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset accession numbers PXD066234 and PXD066291. All other data is available from the corresponding author on reasonable request.

Supplemental Table

Table S1: Summary of cross-linked peptide, including intra- and inter- protein cross-links, PPI annotations based on STRING and mapping of intra-cross-links to known structures from PDB and predicted structure from AlphaFold.

Table S2: Summary of cross-linked protein information and SUBA annotation.

Table S3: Analysis of protein classification using PANTHER.

Table S4: Detected cross-links for ribosomal and histones PPIs, and a list of ribosome-related ubiquitinated peptides.

Table S5: Names, accession numbers and description used in the figures and a glossary.

Material and methods

Plant growth condition

Arabidopsis thaliana wild-type (Col-0) seeds were sterilized and sown on ½ Murashige and Skoog (MS) agar square plates supplemented with 0.6% (w/v) phytoblend and lined with nylon fabric. After stratification at 4°C for 3 days, the plates were transferred to growth conditions.

Light-grown seedlings: plants were vertically grown in a growth chamber at 22°C under a 16-hour light / 8-hour dark cycle for 10-12 days, then flash frozen in liquid nitrogen and stored at -80°C for crosslinking experiments on cell lysates and isolated chloroplasts. Dark-grown seedlings: plants were grown horizontally at 22°C in complete darkness for 3 days in a growth chamber, then flash frozen in liquid nitrogen and stored at -80°C for nucleus isolation. Dark growth was used to reduce Rubisco content.

Frozen Arabidopsis tissue was cryo-ground to a fine powder. Different amounts of tissue were used for cross-linking experiments: 6 g from 10 d seedlings for cell lysates, 195 g from 12d seedlings for chloroplasts, and 32 g (replicate 1 with PhoX) or 14 g (replicate 2 with tbu-PhoX) from 3d dark-grown seedlings for nuclei.

Fraction of plant seedling tissues for cell lysate, chloroplast and nucleus

*Cell lysates***:** Homogenized 10 d old light growth tissues were extracted using MOPS buffer [100 mM MOPS pH 7.6, 150 mM NaCl, 1% Nonidet P-40, 1 mM PMSF, 1x Roche protease inhibitor] at a tissue:buffer ratio of 1:1 (g/mL, w/v). The samples were centrifuged at 20,000 g for 15 min at 4ºC, and the supernatant was then filtered twice with Miracloth and used for the cross-linking reaction.

*Chloroplast isolation***:** Frozen, homogenized tissue powder from 12 d old seedlings was mixed with a grinding buffer similar to (80), using HEPES instead of Tricine for crosslinker compatibility. Briefly, 5 g of tissue powder was ground in a mortar and pestle with 30ml of grinding buffer [20mM HEPES-KOH pH 8.0, 330mM sorbitol, 5 mM EDTA, 5 mM EGTA, 5 mM

440 $MgCl_2$, and 10 mM $NaHCO_3$] at 4°C, and under green illumination only. The resuspended
441 homogenate was filtered twice through two layers of Miracloth. 30 ml of the filtrate was carefully
442 transferred onto 10 ml of 40% Percoll [20 mM HEPES-KOH pH 8.0, 330 mM sorbitol, 40% v/v
443 Percoll, 2.5 mM EDTA, 5 mM $MgCl_2$, and 10 mM $NaHCO_3$] in a 50ml centrifuge tube, and
444 centrifuged at 2,600g for 10 min at 4°C. The bottom pellet, enriched in intact chloroplasts, was
445 collected and resuspended in 30 ml of wash buffer [20 mM HEPES-KOH pH 8.0, 330 mM sorbitol,
446 2.5 mM EDTA, 2 mM $MgCl_2$, and 10 mM $NaHCO_3$], and centrifuged at 2,600g for 3 min at 4°C.
447 Each pellet derived from 5 g of tissue was washed twice with 4.5 ml of 1x PBS supplemented with
448 $MgCl_2$ [137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, and 2.5 mM $MgCl_2$]
449 and then resuspended in 1mL 1x PBS supplemented with $MgCl_2$ for the cross-linking reaction.
450 This process was repeated across multiple batches, with a total of 195 g of tissue used for cross-
451 linking.

452 *Nuclear isolation*: The method for nuclear isolation was adapted from (81) with slight
453 modification. Powdered dark-growth Arabidopsis tissue of approximately 2 g was completely
454 homogenized with lysis buffer [20 mM HEPES (pH 7.4), 20 mM KCl, 2 mM EDTA, 2.5 mM
455 $MgCl_2$, 25% glycerol, 250 mM sucrose, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride
456 (PMSF), 1× protease inhibitor cocktail, and 0.2% formaldehyde] at tissue:buffer ratio of 1:5 (g/mL,
457 w/v). The homogenate was then filtered through a 40 μm cell strainer and centrifuged at 4°C,
458 1,500g for 10 min. The pellet was retained as the nuclear fraction. The nuclear pellet was then
459 washed 5 times with 5 mL of nuclear washing buffer [20 mM HEPES (pH 7.4), 2.5 mM $MgCl_2$,
460 25% glycerol, 0.2% Triton X-100, 1 mM PMSF, and 1× protease inhibitor cocktail] and
461 centrifuged at 4°C, 1,500g for 10 min to collect the pellets after each wash. The washed nuclear
462 pellet was then resuspended in 0.5 mL of nuclear resuspension buffer 1 [20 mM HEPES (pH 7.4),

463 10 mM $MgCl_2$, 250 mM sucrose, 0.2% Triton X-100, 1 mM PMSF, and 1× protease inhibitor
464 cocktail] and carefully laid on top of nuclear resuspension buffer 2 [20 mM HEPES (pH 7.4), 10
465 mM $MgCl_2$, 1.7 M sucrose, 0.2% Triton X-100, 1 mM PMSF, and 1× protease inhibitor cocktail],
466 then centrifuged at 4°C,16,000g for 45 min. The supernatant was carefully removed, and the
467 nuclear pellet was resuspended in 0.5 mL 1 X PBS (pH 7.5) buffer for crosslinking reaction. The
468 intactness and purity of the nuclear pellet was assessed by DAPI staining under a Leica SP8
469 confocal microscope. Multiple samples were prepared in parallel, and final isolated nuclei were
470 pooled before crosslinking.

471 **Cross-linking reaction**

472 PhoX or tBU-PhoX cross-linker was freshly prepared at a stock concentration of 100 mM in
473 DMSO for each cross-linking experiment (supplementary Fig. S3). A working concentration of 2-
474 5 mM cross-linker was added to buffer solution containing whole cell lysate, isolated chloroplasts
475 or isolated nuclei and incubated for 2 h at 4°C (dark condition for chloroplast cross-linking) with
476 end-to-end rotation. The cross-linking reaction was quenched by adding Tris pH 8.0 to a final
477 concentration of 20 mM and subjected to protein extraction.

478 **Protein extraction, digestion, IMAC enrichment and RPLC fractionation**

479 Protein extraction from the cross-linked samples, digestion to peptides, immobilized metal affinity
480 chromatography (IMAC) enrichment, and high pH reverse HPLC fractionation followed the same
481 procedure as described in (13). Multiple batches of PhoX-cross-linked peptides from isolated intact
482 chloroplasts were combined before IMAC enrichment. More information can be found in
483 supplementary Fig. S3. Multiple batches of samples were combined prior to high pH reversed-
484 phase HPLC offline fractionation (RPLC) on Vanquish Flex HPLC, and fractions were collected
485 for further LC-MS/MS analysis.

486 To generate selected proteomic libraries, total peptides from cell lysate and chloroplast samples
487 were analyzed by LC-MS/MS. Chloroplast peptides were further fractionated into eight fractions
488 using the Pierce High pH Reverse-Phase Peptide Fractionation Kit.

489 **diGly proteomics analysis to identify ubiquitinated proteins *in vivo***

490 Metabolically labeled Arabidopsis plants with ^{14}N and ^{15}N isotope were harvested, followed by
491 protein extraction and trypsin digestion as previously described (13, 65). Ubiquitinated peptides
492 were enriched using the PTMScan® Ubiquitin Remnant Motif (K-ε-GG) Kit (Cell Signaling
493 Technology) according to the manufacturer's instructions. The enriched peptides were then
494 analyzed by LC-MS/MS.

495 **LC-MS/MS analysis**

496 Cross-linked peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-
497 MS/MS) using an Easy-nLC 1200 ultra-performance liquid chromatography (UPLC) system
498 (Thermo Fisher) coupled to an Orbitrap Eclipse Tribrid quadrupole Orbitrap mass spectrometer
499 (Thermo Fisher). Peptides were first loaded onto a C18 trap column (Acclaim PepMap 100 C18)
500 and then separated on an analytical C18 column (Aurora Series, 25 cm × 75 µm ID, Ion Opticks).
501 The flow rate was set to 300 nL/min and the separation was performed over a 120-minute gradient.
502 Peptides were eluted using a gradient from 3% to 28% solvent B over 106 minutes, followed by
503 an increase from 28% to 44% over 15 minutes and a final wash at 90% solvent B for 14 minutes.
504 Solvent A consisted of 0.1% formic acid and solvent B consisted of 80% (v/v) acetonitrile with
505 0.1% (v/v) formic acid.

506 Data was acquired on an Orbitrap Eclipse mass spectrometer. Full MS scans were acquired
507 over a mass-to-charge (m/z) range of 375–1600 at a resolution of 120,000, with an automatic gain
508 control (AGC) target of 200,000, a maximum injection time of 50 ms, normalized AGC target set

509 to 50%, and RF lens set to 30%. The most intense multiply charged precursors (charge states 2–8)
510 were selected for fragmentation with a cycle time of 3 seconds. MS/MS scans were acquired at a
511 resolution of 15,000 with an AGC target of 5×10^4, maximum injection time of 22 ms, isolation
512 window of 1.4 m/z, and normalized AGC target set to 100%. Fragmentation was performed using
513 higher-energy collisional dissociation (HCD) with a normalized collision energy (NCE) of 27.
514 Dynamic exclusion was enabled for 30 seconds. For cross-linked peptide detection, charge states
515 3 to 8 were used.

516 For Real-time library search (RTLS) acquisition (15), a low-resolution MS2 survey scan
517 performed in the ion trap over an m/z range of 120–500 was included. The resulting MS2 peaks
518 were matched against a custom spectral library containing signature peak patterns characteristic of
519 mono-links and cross-links (signature m/z values: 201.1231, 215.1387, 312.0632, and 377.1269,
520 each with characteristic relative intensities for mono- or cross-linked species). Only precursor ions
521 whose survey MS2 spectra matched the cross-link pattern using the "Similarity Search" function
522 were selected for high-resolution MS2 acquisition.

523 **Data search using pLINK2 and Protein Prospector**

524 Liner peptide searches were searched using Protein Prospector against Arabidopsis TAIR database
525 as described in (65). Precursor and fragment mass tolerance were set up as 10 and 20 ppm.

526 Crosslinked peptide searches were performed using pLink 2.3.11 with the following
527 parameters: crosslinker set to PhoX (209.972 Da); peptide length range: 6-60 amino acids;
528 enzyme: trypsin with up to 3 missed cleavages. Fixed modification: Cys carbamidomethylation;
529 variable modifications: protein N-terminal acetylation and Met oxidation. Precursor and fragment
530 mass tolerances were set to 5 ppm and 20 ppm, respectively.

531 MS2 spectra were searched against the TAIR 10 protein database (35,386 entries), as well
532 as selective database for cell lysate (7,835 entries) and chloroplast (9,897 entries), as described in
533 Supplementary Fig.S3. Search results were filtered at 5% FDR at the peptide spectrum matches
534 (PSMs) level. FDRs for monolinks, loop-links, intra- and inter-protein cross-links were calculated
535 separately using pLINK2. To compare two acquisition methods (XL-Standard [Standard] and
536 Real-time library search [RTLS]), data were acquired on the whole cell lysate XL samples, search
537 against TAIR10 database, and directly compared.

538 **Data compilation and reporting**

539 Intra- and inter-protein cross-links were compiled from multiple experiments (Supplementary Fig.
540 S3) and reported separately (Supplementary Table S1). The total number of cross-linked peptide
541 pairs was determined by counting unique lysine–lysine (K–K) residue pairs along with their
542 corresponding backbone sequences. Peptides with identical backbone sequences and the same K–
543 K pair were collapsed into a single count, regardless of charge state or peptide backbone
544 modifications (e.g., oxidation).

545 **Subcellular localization analysis of proteins using SUBA5**

546 The subcellular localization of cross-linked proteins was analyzed using SUBA5
547 (https://suba.live/). Accession IDs of proteins identified in both intra- and inter-protein cross-links
548 were submitted as input. Only the localization consensus data were used for classification. Proteins
549 assigned to multiple compartments were counted once in each corresponding localization category
550 when calculating total counts.

551 **PANTHER protein class analysis and mapping intralinks to PDB and AlphaFold structures**
552 **using ComMap**

553 Arabidopsis gene identifiers (ATG accession numbers) were used for protein class analysis
554 using PANTHER™ Protein Class tool (24), which contains 210 classification terms (version 19.0,
555 released 2024-06-20).

556 To map the results of intralinks against experimentally determined structures from the
557 Protein Data Bank (PDB) and predicted structures from AlphaFold protein structure database, the
558 protein accession IDs were first converted to UniProt IDs. The intralink information was then
559 mapped to available structures using ComMap, a tool that enables large-scale structure-based
560 mapping by integrating XL-MS data with existing structural models and performing distance
561 calculations (16).

562 **Crosslinking data visualization**

563 *Cytoscape visualization*: The entire cross-link-based protein-protein interaction (PPI)
564 network was visualized in Cytoscape (version 3.10, https://cytoscape.org/) using inter-protein
565 cross-link data, including heteromeric and homomultimeric cross-links. Nodes were colored
566 according to protein class information from the PANTHER classification system, and edges were
567 colored according to the STRING interaction score (https://string-db.org/).

568 *xiVIEW:* PPIs, including both inter- and intra-protein cross-links, were visualized using
569 xiVIEW (https://xiview.org/) (82). PPIs were categorized via the built-in STRING data loader
570 (score cutoff 0.4). Cross-linked peptides supporting the same lysine pair were consolidated into a
571 unique K-K linkage, collapsing redundancy (e.g. due to miscleavage events) and reporting each
572 linkage only once.

573 *Mapping of crosslinks to high-resolution structures of proteins/complexes*. Intra- and inter-
574 crosslinks were superimposed on PDB or AlphaFold structures using the built-in function of
575 xiVIEW, which also calculated the corresponding crosslink distances within the high-resolution

structures. Alignments of protein orthologs were performed and visualized using PyMOL (https://www.pymol.org/).

Co-localization assay

Co-localization was assessed in *Nicotiana Benthamian* leaves using transient co-expression. WSD4 was cloned into a modified pMDC43 Gateway binary vector containing an N-terminal mCherry tag to generate 35S::mCherry-WSD4, and HTB11 was introduced to pGWB6 gateway vector to generate 35S::sGFP-HTB11. Agrobacterium-mediated infiltration was performed, and fluorescence signals were visualized three days post-infiltration using a Leica SP8 confocal microscope.

Reference

1. Yu C & Huang L (2018) Cross-Linking Mass Spectrometry: An Emerging Technology for Interactomics and Structural Biology. *Anal Chem* 90(1):144-165.
2. Piersimoni L, Kastritis PL, Arlt C, & Sinz A (2022) Cross-Linking Mass Spectrometry for Investigating Protein Conformations and Protein-Protein Interactions horizontal line A Method for All Seasons. *Chem Rev* 122(8):7500-7531.
3. Michaelis AC*, et al.* (2023) The social and structural architecture of the yeast protein interactome. *Nature* 624(7990):192-200.
4. Xu SL, Shrestha R, Karunadasa SS, & Xie PQ (2023) Proximity Labeling in Plants. *Annu Rev Plant Biol* 74:285-312.
5. Greenblatt JF, Alberts BM, & Krogan NJ (2024) Discovery and significance of protein-protein interactions in health and disease. *Cell* 187(23):6501-6517.
6. Lee K & O'Reilly FJ (2023) Cross-linking mass spectrometry for mapping protein complex topologies in situ. *Essays Biochem* 67(2):215-228.
7. Zhu X*, et al.* (2016) In planta chemical cross-linking and mass spectrometry analysis of protein structure and interaction in Arabidopsis. *Proteomics* 16(13):1915-1927.
8. Jayachandran B, Parvin TN, Alam MM, Chanda K, & Mm B (2022) Insights on Chemical Crosslinking Strategies for Proteins. *Molecules* 27(23).
9. Smythers AL, Iannetta AA, & Hicks LM (2021) Crosslinking mass spectrometry unveils novel interactions and structural distinctions in the model green alga Chlamydomonas reinhardtii. *Mol Omics* 17(6):917-928.
10. Dai S*, et al.* (2023) Capturing the hierarchically assorted modules of protein-protein interactions in the organized nucleome. *Mol Plant* 16(5):930-961.
11. Steigenberger B, Pieters RJ, Heck AJR, & Scheltema RA (2019) PhoX: An IMAC-Enrichable Cross-Linking Reagent. *ACS Cent Sci* 5(9):1514-1522.
12. Jiang PL*, et al.* (2022) A Membrane-Permeable and Immobilized Metal Affinity Chromatography (IMAC) Enrichable Cross-Linking Reagent to Advance In Vivo Cross-Linking Mass Spectrometry. *Angew Chem Int Ed Engl* 61(12):e202113937.

13. Xu S-L, *et al.* (2017) Proteomic analysis reveals O-GlcNAc modification on proteins with key regulatory functions in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* 114(8):E1536-E1543.
14. Chen ZL, *et al.* (2019) A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat Commun* 10(1):3404.
15. Ruwolt M, *et al.* (2023) Real-Time Library Search Increases Cross-Link Identification Depth across All Levels of Sample Complexity. *Anal Chem* 95(12):5248-5255.
16. Zhang W, *et al.* (2023) ComMap: a software to perform large-scale structure-based mapping for cross-linking mass spectrometry. *Bioinformatics* 39(2).
17. Valegard K, *et al.* (2018) Structural and functional analyses of Rubisco from arctic diatom species reveal unusual posttranslational modifications. *J Biol Chem* 293(34):13033-13043.
18. Murachelli AG, Ebert J, Basquin C, Le Hir H, & Conti E (2012) The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex. *Nat Struct Mol Biol* 19(4):378-386.
19. Bi Y, *et al.* (2021) Arabidopsis ACINUS is O-glycosylated and regulates transcription and alternative splicing of regulators of reproductive transitions. *Nat Commun* 12(1):945.
20. Shrestha R, *et al.* (2025) Next-Generation Mapping of the ACINUS-Mediated Alternative Splicing Machinery and Its Regulation by O-glycosylation in Arabidopsis. . *BioRxiv*.
21. Yu A, *et al.* (2020) Photosynthetic Phosphoribulokinase Structures: Enzymatic Mechanisms and the Redox Regulation of the Calvin-Benson-Bassham Cycle. *Plant Cell* 32(5):1556-1573.
22. Wheat A, *et al.* (2021) Protein interaction landscapes revealed by advanced in vivo cross-linking-mass spectrometry. *Proc Natl Acad Sci U S A* 118(32).
23. Shannon P, *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498-2504.
24. Mi H, *et al.* (2021) PANTHER version 16: a revised family classification, tree-based classification tool, enhancer regions and extensive API. *Nucleic Acids Res* 49(D1):D394-D403.
25. Szklarczyk D, *et al.* (2021) The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 49(D1):D605-D612.
26. Ahmed T, Yin Z, & Bhushan S (2016) Cryo-EM structure of the large subunit of the spinach chloroplast ribosome. *Sci Rep* 6:35793.
27. Bieri P, Leibundgut M, Saurer M, Boehringer D, & Ban N (2017) The complete structure of the chloroplast 70S ribosome in complex with translation factor pY. *EMBO J* 36(4):475-486.
28. Younkin AD, Gregory ST, & O'Connor M (2020) Alterations in the ribosomal protein bL12 of E. coli affecting the initiation, elongation and termination of protein synthesis. *Biochimie* 175:173-180.
29. Liljas A & Sanyal S (2018) The enigmatic ribosomal stalk. *Q Rev Biophys* 51:e12.
30. Komenda J, Sobotka R, & Nixon PJ (2024) The biogenesis and maintenance of PSII: Recent advances and current challenges. *Plant Cell* 36(10):3997-4013.
31. Chen H, *et al.* (2006) A Psb27 homologue in Arabidopsis thaliana is required for efficient repair of photodamaged photosystem II. *Plant Mol Biol* 61(4-5):567-575.
32. Wei L, *et al.* (2010) LPA19, a Psb27 homolog in Arabidopsis thaliana, facilitates D1 protein precursor processing during PSII biogenesis. *J Biol Chem* 285(28):21391-21398.
33. Li A, *et al.* (2024) Structural basis for an early stage of the photosystem II repair cycle in Chlamydomonas reinhardtii. *Nat Commun* 15(1):5211.
34. Weisz DA, *et al.* (2019) A novel chlorophyll protein complex in the repair cycle of photosystem II. *Proc Natl Acad Sci U S A* 116(43):21907-21913.

35. Xingxing C*, et al.* (2018) Crystal structure of Psb27 from Arabidopsis thaliana determined at a resolution of 1.85 A. *Photosynth Res* 136(2):139-146.
36. Sheng X*, et al.* (2019) Structural insight into light harvesting for photosystem II in green algae. *Nat Plants* 5(12):1320-1330.
37. Martinez-Seidel F, Beine-Golovchuk O, Hsieh YC, & Kopka J (2020) Systematic Review of Plant Ribosome Heterogeneity and Specialization. *Front Plant Sci* 11:948.
38. Chaker-Margot M (2018) Assembly of the small ribosomal subunit in yeast: mechanism and regulation. *RNA* 24(7):881-891.
39. Lee S*, et al.* (2017) The small GTPase, nucleolar GTP-binding protein 1 (NOG1), has a novel role in plant innate immunity. *Sci Rep* 7(1):9260.
40. Yoshida K*, et al.* (2005) Interaction of the taxilin family with the nascent polypeptide-associated complex that is involved in the transcriptional and translational processes. *Genes Cells* 10(5):465-476.
41. Faus I*, et al.* (2018) Arabidopsis ILITHYIA protein is necessary for proper chloroplast biogenesis and root development independent of eIF2alpha phosphorylation. *J Plant Physiol* 224-225:173-182.
42. Kim KQ*, et al.* (2024) Multiprotein bridging factor 1 is required for robust activation of the integrated stress response on collided ribosomes. *Mol Cell* 84(23):4594-4611 e4599.
43. Tafforeau L*, et al.* (2013) The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. *Mol Cell* 51(4):539-551.
44. Chen J*, et al.* (2024) The Ptch/SPOUT1 methyltransferase deposits an m(3)U modification on 28S rRNA for normal ribosomal function in flies and humans. *Sci Adv* 10(50):eadr1743.
45. Ghandadi M, Dobi A, & Malhotra SV (2024) A role for RIO kinases in the crosshair of cancer research and therapy. *Biochim Biophys Acta Rev Cancer* 1879(3):189100.
46. Messling JE*, et al.* (2022) Targeting RIOK2 ATPase activity leads to decreased protein synthesis and cell death in acute myeloid leukemia. *Blood* 139(2):245-255.
47. Kim DS, Challa S, Jones A, & Kraus WL (2020) PARPs and ADP-ribosylation in RNA biology: from RNA expression and processing to protein translation and proteostasis. *Genes Dev* 34(5-6):302-320.
48. Udeshi ND, Mertins P, Svinkina T, & Carr SA (2013) Large-scale identification of ubiquitination sites by mass spectrometry. *Nat Protoc* 8(10):1950-1960.
49. Van de Vosse DW, Wan Y, Wozniak RW, & Aitchison JD (2011) Role of the nuclear envelope in genome organization and gene expression. *Wiley Interdiscip Rev Syst Biol Med* 3(2):147-166.
50. Li F*, et al.* (2008) Identification of the wax ester synthase/acyl-coenzyme A: diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in Arabidopsis. *Plant Physiol* 148(1):97-107.
51. Patwari P*, et al.* (2019) Surface wax esters contribute to drought tolerance in Arabidopsis. *Plant J* 98(4):727-744.
52. Vazquez A, Flammini A, Maritan A, & Vespignani A (2003) Global protein function prediction from protein-protein interaction networks. *Nat Biotechnol* 21(6):697-700.
53. Giurgiu M*, et al.* (2019) CORUM: the comprehensive resource of mammalian protein complexes-2019. *Nucleic Acids Res* 47(D1):D559-D563.
54. Luck K*, et al.* (2020) A reference map of the human binary protein interactome. *Nature* 580(7803):402-408.
55. Burke DF*, et al.* (2023) Towards a structurally resolved human protein interaction network. *Nat Struct Mol Biol* 30(2):216-225.
56. Brandao MM, Dantas LL, & Silva-Filho MC (2009) AtPIN: Arabidopsis thaliana protein interaction network. *BMC Bioinformatics* 10:454.

57. Herrera-Ubaldo H, *et al.* (2023) The protein-protein interaction landscape of transcription factors during gynoecium development in Arabidopsis. *Mol Plant* 16(1):260-278.
58. Yang X, *et al.* (2020) PlaPPISite: a comprehensive resource for plant protein-protein interaction sites. *BMC Plant Biol* 20(1):61.
59. Balch S, *et al.* (2025) Cryo-EM structure of human telomerase dimer reveals H/ACA RNP-mediated dimerization. *Science* 389(6756):eadr5817.
60. Dang DT (2022) Molecular Approaches to Protein Dimerization: Opportunities for Supramolecular Chemistry. *Front Chem* 10:829312.
61. Yang Z, *et al.* (2018) EBS is a bivalent histone reader that regulates floral phase transition in Arabidopsis. *Nat Genet* 50(9):1247-1253.
62. Shu J, *et al.* (2021) Genome-wide occupancy of Arabidopsis SWI/SNF chromatin remodeler SPLAYED provides insights into its interplay with its close homolog BRAHMA and Polycomb proteins. *Plant J* 106(1):200-213.
63. Lim MH, *et al.* (2004) A new Arabidopsis gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. *Plant Cell* 16(3):731-740.
64. Yan Z, Shi H, Liu Y, Jing M, & Han Y (2020) KHZ1 and KHZ2, novel members of the autonomous pathway, repress the splicing efficiency of FLC pre-mRNA in Arabidopsis. *J Exp Bot* 71(4):1375-1386.
65. Shrestha R, *et al.* (2022) 15N Metabolic Labeling Quantification Workflow in Arabidopsis Using Protein Prospector. *Front. Plant Sci.* 13:832562.
66. Reyes AV, Shrestha R, Baker PR, Chalkley RJ, & Xu S-L (2022) Application of Parallel Reaction Monitoring in 15N Labeled Samples for Quantification. *Front. Plant Sci.* 13:832585.
67. Matsuo Y, *et al.* (2017) Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat Commun* 8(1):159.
68. Simsek D, *et al.* (2017) The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* 169(6):1051-1065 e1018.
69. Peng Y, Markov Y, Goncearenco A, Landsman D, & Panchenko AR (2021) Human Histone Interaction Networks: An Old Concept, New Trends. *J Mol Biol* 433(6):166684.
70. Zhang S, Yu Q, Li Z, Zhao Y, & Sun Y (2024) Protein neddylation and its role in health and diseases. *Signal Transduct Target Ther* 9(1):85.
71. Alemasova EE & Lavrik OI (2019) Poly(ADP-ribosyl)ation by PARP1: reaction mechanism and regulatory proteins. *Nucleic Acids Res* 47(8):3811-3827.
72. Messner S, *et al.* (2010) PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucleic Acids Res* 38(19):6350-6362.
73. Nosella ML, *et al.* (2024) Poly(ADP-ribosyl)ation enhances nucleosome dynamics and organizes DNA damage repair components within biomolecular condensates. *Mol Cell* 84(3):429-446 e417.
74. Su Y, *et al.* (2017) Phosphorylation of Histone H2A at Serine 95: A Plant-Specific Mark Involved in Flowering Time Regulation and H2A.Z Deposition. *Plant Cell* 29(9):2197-2213.
75. Sabari BR, Zhang D, Allis CD, & Zhao Y (2017) Metabolic regulation of gene expression through histone acylations. *Nat Rev Mol Cell Biol* 18(2):90-101.
76. Goudarzi A (2019) The recent insights into the function of ACAT1: A possible anti-cancer therapeutic target. *Life Sci* 232:116592.
77. Liu Y, *et al.* (2022) Mechanisms and inhibition of Porcupine-mediated Wnt acylation. *Nature* 607(7920):816-822.
78. Nile AH & Hannoush RN (2016) Fatty acylation of Wnt proteins. *Nat Chem Biol* 12(2):60-69.

79. Xue B & Rhee SY (2023) Status of genome function annotation in model organisms and crops. *Plant Direct* 7(7):e499.
80. Casazza AP, Tarantino D, & Soave C (2001) Preparation and functional characterization of thylakoids from Arabidopsis thaliana. *Photosynth Res* 68(2):175-180.
81. Shao Z*, et al.* (2024) Nuclear pyruvate dehydrogenase complex regulates histone acetylation and transcriptional regulation in the ethylene response. *Sci Adv* 10(30):eado2825.
82. Combe CW, Graham M, Kolbowski L, Fischer L, & Rappsilber J (2024) xiVIEW: Visualisation of Crosslinking Mass Spectrometry Data. *J Mol Biol* 436(17):168656.

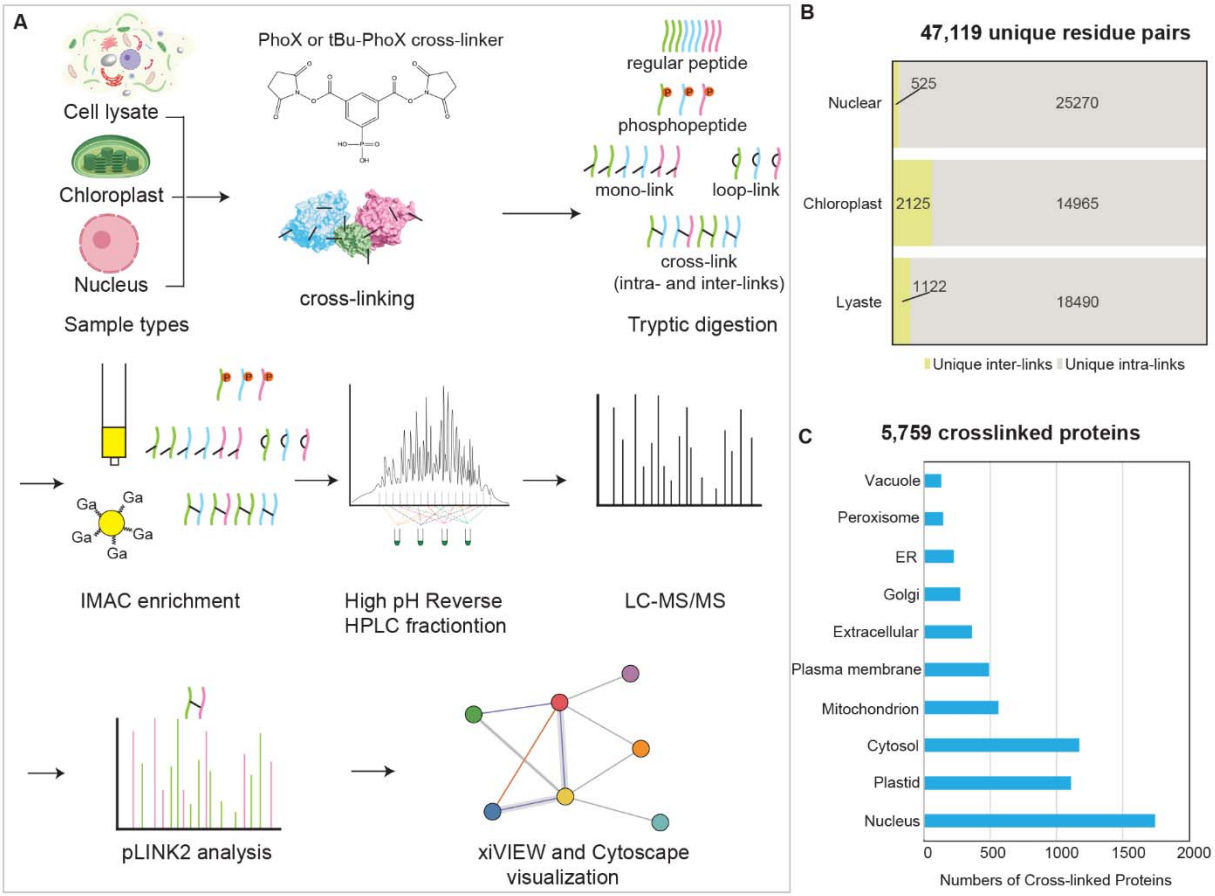


Figure 1. Efficient cross-linking workflow for complex protein samples identifies numerous cross-links in Arabidopsis.

(**A**) Schematic of the workflow using PhoX or tBu-PhoX to covalently cross-link spatially proximal lysine residues in Arabidopsis subcellular fractions (lysate, chloroplast, nucleus), followed by digestion, IMAC enrichment, high pH reversed-phase HPLC fractionation, and LC-MS/MS. Data were analyzed using pLINK2, Cytoscape, and xiVIEW.

(**B**) Summary of unique intra- and inter-protein cross-links from each subcellular fraction. Intra-links connect non-overlapping peptides within a protein; inter-links include heteromeric and homomultimeric cross-links.

(**C**) Subcellular localization of cross-linked proteins based on SUBA5, with redundancy of proteins assigned to multiple compartments.

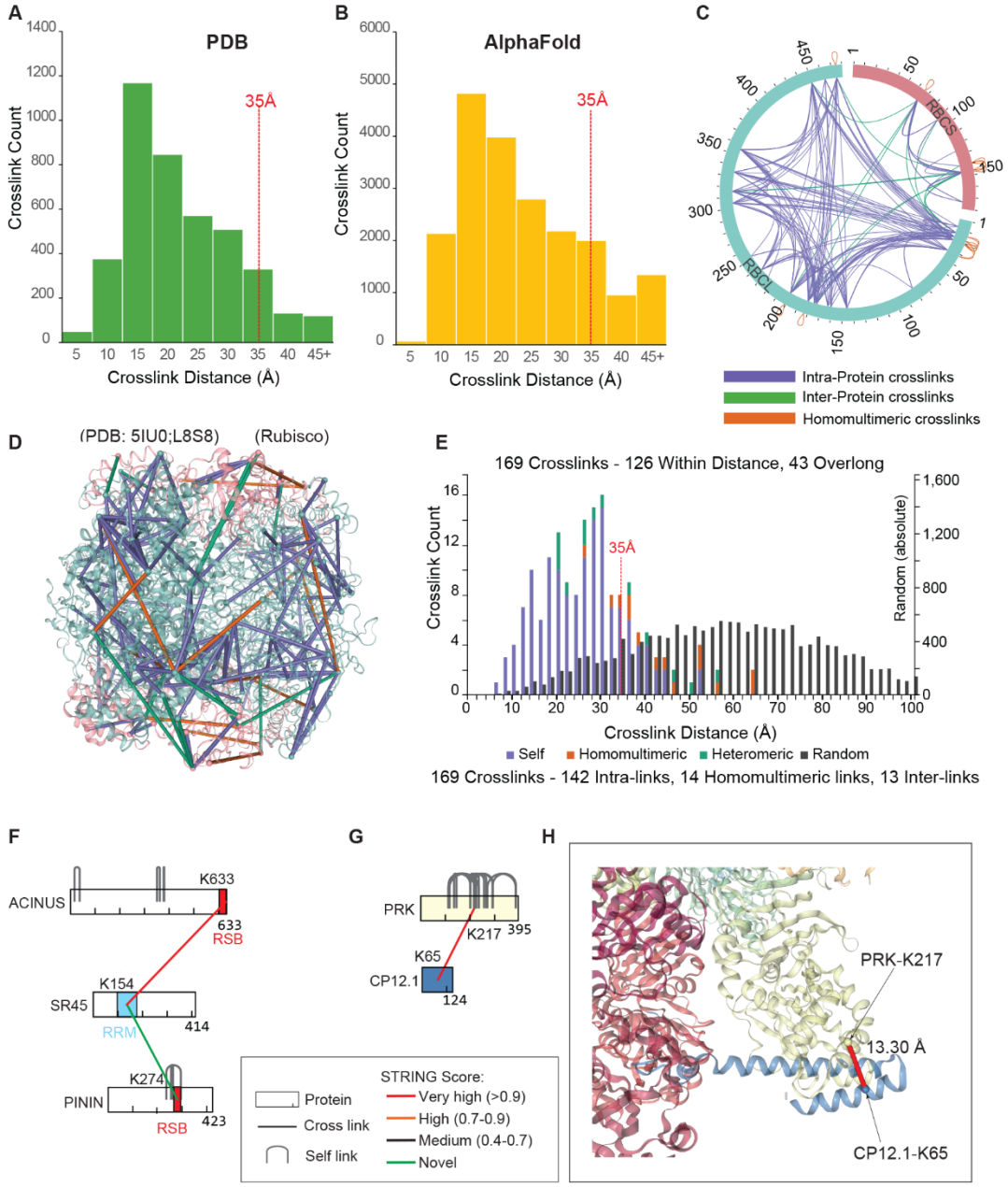


Figure 2. Validation of intra- and inter-protein cross-links by structural mapping.
(**A**-**B**) Histogram of Euclidean Cα-Cα distances of intra-protein cross-links mapped to Arabidopsis structures from PDB (A) or AlphaFold (B), showing 94% and 89% within the 35Å constraint, respectively.
(**C**-**E**) Cross-links mapped to the Arabidopsis Rubisco L_8S_8 complex modeled from the L_2S_2 structure (PDB:5IU0). RBCL is shown in cyan and RBCS is shown in pink. Self(intra), heteromeric, and homomultimeric links are shown in purple, green, and orange, respectively. The circular view (C) and structural mapping (D-E) show that 74.56% of the cross-links are within 35Å.

(**F**) Cross-links between SR45 and ACINUS or PININ occur at known interaction interfaces involving the RRM domain of SR45 and the RSB domain of ACINUS and PININ.
(**G-H**) Intra- and inter-protein cross-links between PRK and CP12.1 mapped onto the GAPDH-CP12-PRK complex (PDB:6KEZ), aligning with the known interaction interface.

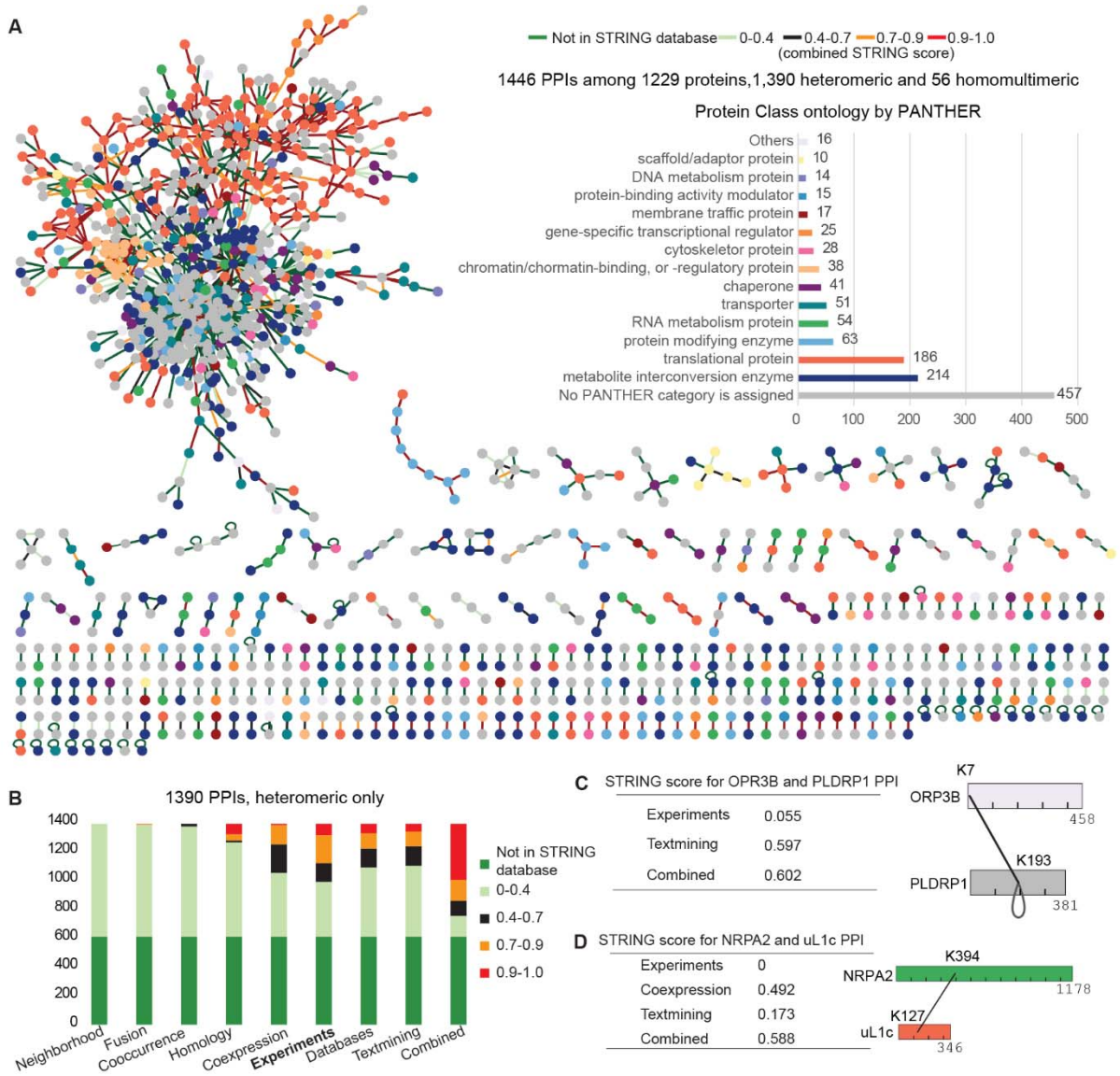


Figure 3. XL-MS identifies direct protein-protein interactions and reveals known and novel complexes.

(**A**) XL-MS identified an Arabidopsis interactome comprising 1,446 PPIs, including 1,390 heteromeric and 56 homomultimeric interactions among 1,229 proteins. Nodes represent proteins, colored by PANTHER protein class; edges represent interactions, colored by the combined STRING confidence score. Of the total PPIs, 637 PPIs have a combined confidence score of at least 0.4, including 532 with a score of at least 0.7.

(**B**) Summary of the STRING score, showing the contribution from eight evidence sources. XL-MS provides experimental support for 71% of the heteromeric PPIs that either have low confidence scores (<0.4) or lack prior experimental evidence.

(**C-D**) Examples of specific PPIs detected by XL-MS that complement low confidence or missing experimental STRING scores: the ORP3B-PLDRP1 interaction (cross-linked at ORP3B-K7/PLDRP1-K193) and the NRPA2-uL1c interaction (NRPA2-K394/uL1c-K127).

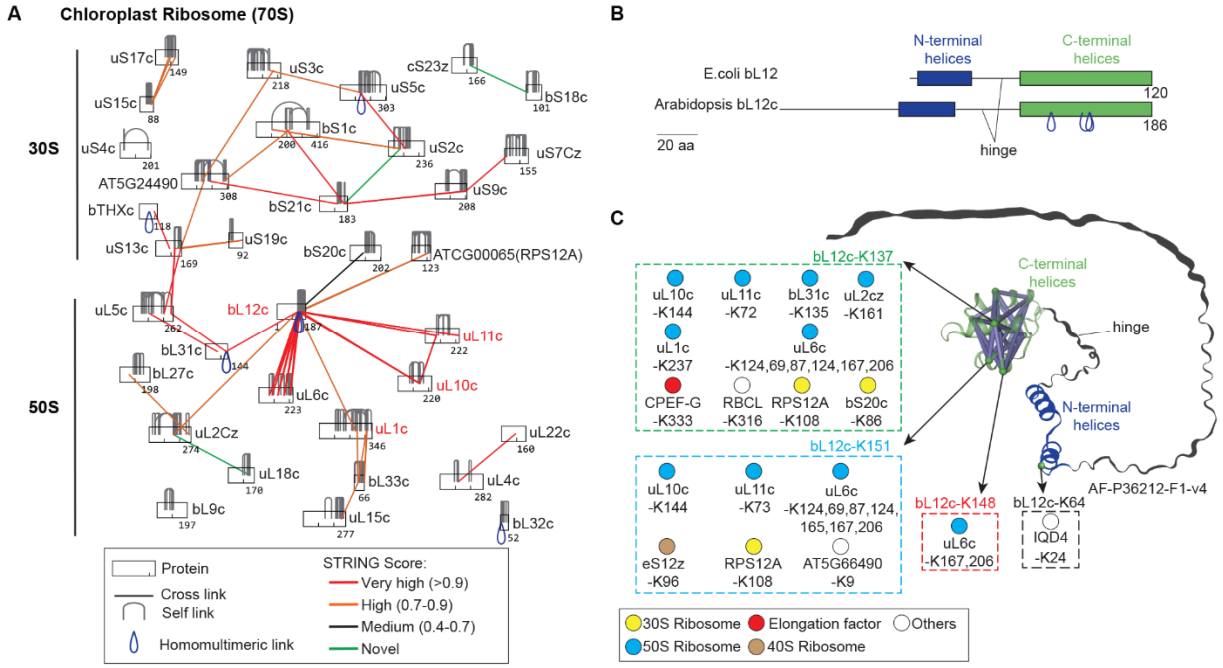


Figure 4. XL-MS resolves the topology of chloroplast 70S ribosome stalk regions inaccessible by cryo-EM.

(**A**) XL-MS identifies flexible stalk proteins (bL12c, uL1c, uL10c, uL11c), with cross-links color-coded by STRING confidence.

(**B**) Domain comparison between Arabidopsis bL12c and *E. coli* bL12. Homodimer interfaces are found at the C-terminus of bL12c.

(**C**) Cross-links mapped onto the predicted bL12c structure (AF-P36212-F1-v4) showing lysines K64, K137, K148, and K151 interacting with ribosomal subunits and the elongation factor CPEF-G. Nodes are color-coded according to annotation category.

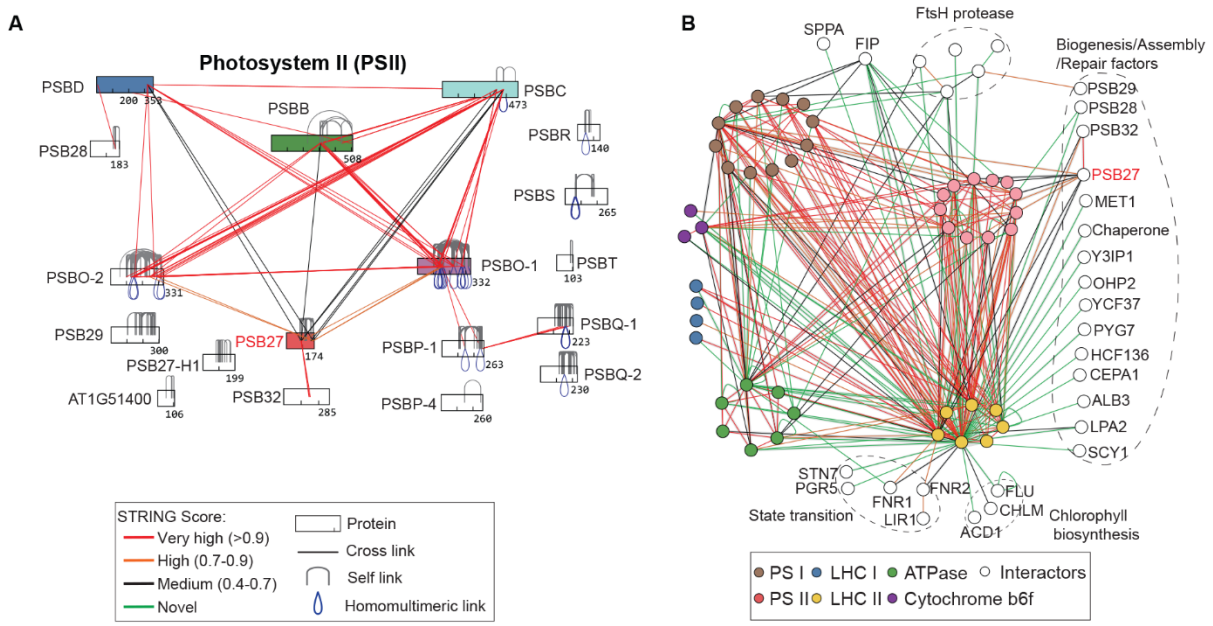
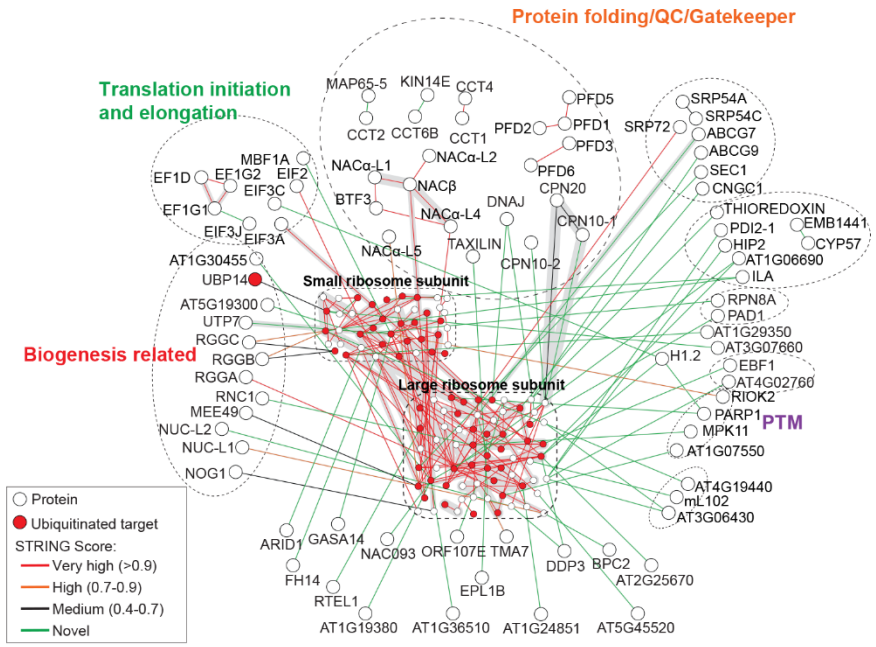


Figure 5. XL-MS reveals dynamic connectivity and unique features in photosystem complexes.

(**A**) PSII topology showing cross-links color-coded by STRING confidence. Notable PSB27 cross-links (K93, K168) highlight interactions with PSII core and OEC components.

(**B**) XL-MS uncovers extensive interactions among PSI, PSII, LHCI, LHCII, ATPase, FtsH/SPPA proteases, and transient proteins involved in state transitions, assembly, repair, and chlorophyll metabolism.

A



B

Interactor	Lys(K)	Ribosome	Lys(K)
MEE49	22	uL29z	35
RNC1	85	uL4z	289
AT5G19300	3	eS1z	19
NOG1	532	eL31z	58
UBP14	154	eS31z	48
UTP7	453	eS30z	3
UTP7	451	eS30z	20
RGGA	255	eS31y	119
RGGA	266	uS19u	65
RGGB	245	eS31y	119
RGGB	256	uS19u	65
RGGC	42	eL8z	250
NUC-L1	79	uL23y	32
NUC-L2	228	eL8y	116
EIF2	429	eL24z	130
AT1G30455	169	eL24y	155
EIF3A	568	uL24z	41
EIF3A	568	uL24z	130
MBF1A	19	eS30z	53
NACα-like 5	92	uL29y	43
NACβ	18	eL31z	58
NACβ	18	eL31z	61
NACα-like 4	95	uL29y	43
TAXILIN	339	eL31y	38
SRP72	579	uL29y	77
ILA	793	eS8z	152
ILA	793	uS19u	65
AT1G06690	85	uL23z	32
AT1G06690	85	eL8z	227
THIOREDOXIN	229	eL38z	29
PDI 2-1	73	uL18z	279
HIP2	161	uL18z	240
PAD1	243	eL24y	155
RPN8A	40	eS1z	27
EBF1	365	uL15x	118
AT4G02760	7	uL18z	254
MPK11	38	uL4z	289
AT1G07550	77	eL27z	47
RIOK2	13	eS30z	3
PARP1	61	eL43y	7
AT1G29350	5	uS15z	43
AT3G07760	16	eS31z	6
H1.2	217	uL23z	24

C

Detected	Description/Function	Human homolog	Ribosome related function (based on human/yeast homolog)	Ribosome related function in plant
AT5G19300	methytransferase C9orf114	SPOUT1	Ribosome synthesis trans-Acting factor	uncharacterized
PARP1	ADP-ribosylation	PARPs (17)	Polysome formation	uncharacterized
RIOK2	Kinase/ATPase	RIOKs(3)	40S maturation	uncharacterized
MPK11	Kinase	MPKs	uncharacterized	uncharacterized
UBP14	De-ubiquitinase	UBPs	Maintain Ribosome assembly	uncharacterized
EBF11	F-box E3 ligase	FBXL20	uncharacterized	uncharacterized
AT4G02760	F box E3 ligase with LRR	FBXL2-like	uncharacterized	uncharacterized

Figure 6. XL-MS of ribosome complexes reveals conserved ribosome-associated factors. (**A**) XL-MS identified extensive interactions within ribosomal subunits as well as with proteins involved in ribosome biogenesis, translation, folding, quality control, stress response, and other regulatory pathways. Ribosomal proteins are overlaid with ubiquitination data; red nodes indicate ubiquitinated proteins.

(**B**) Table showing covalently cross-linked lysine residues between ribosomal proteins, reflecting spatial proximity. Protein labels are color-coded according to the functional categories shown in (A).

(**C**) Table of XL-MS-identified ribosome-associated interactors with potential post-transcriptional or post-translational regulatory functions, many of which remain uncharacterized in plants.

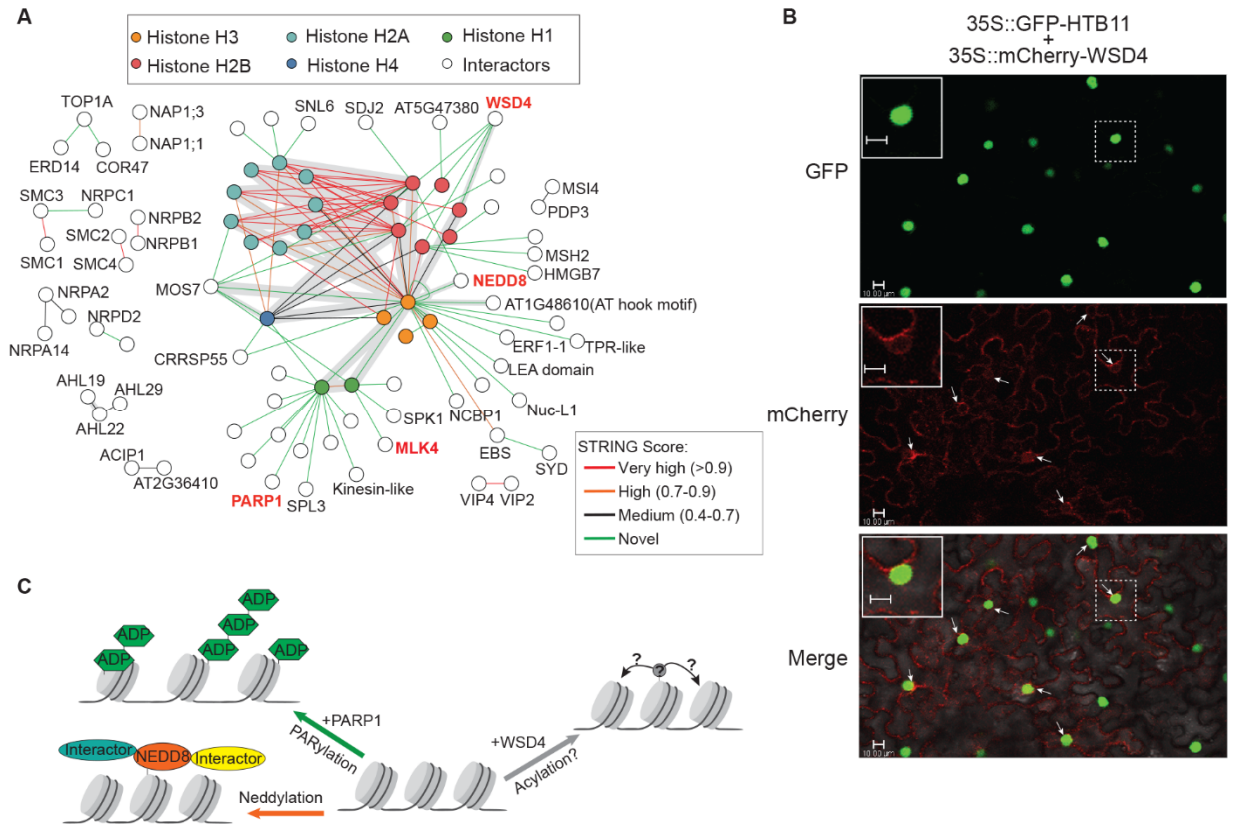
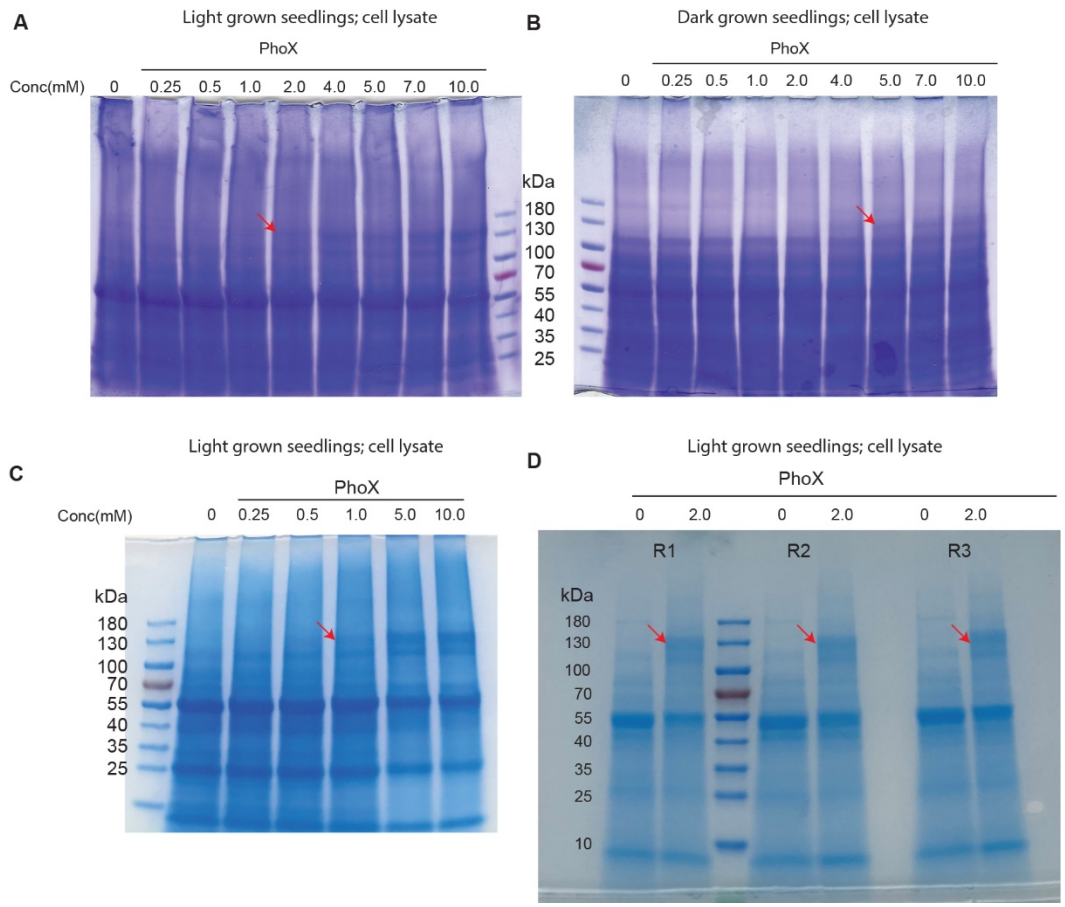


Figure 7. Histone interactome reveals conserved and novel chromatin-associated interactions.

(**A**) XL-MS reveals extensive interactions between histones and both known and previously uncharacterized proteins, including components of chromatin-associated complexes. Enzymatic regulators are highlighted in red.

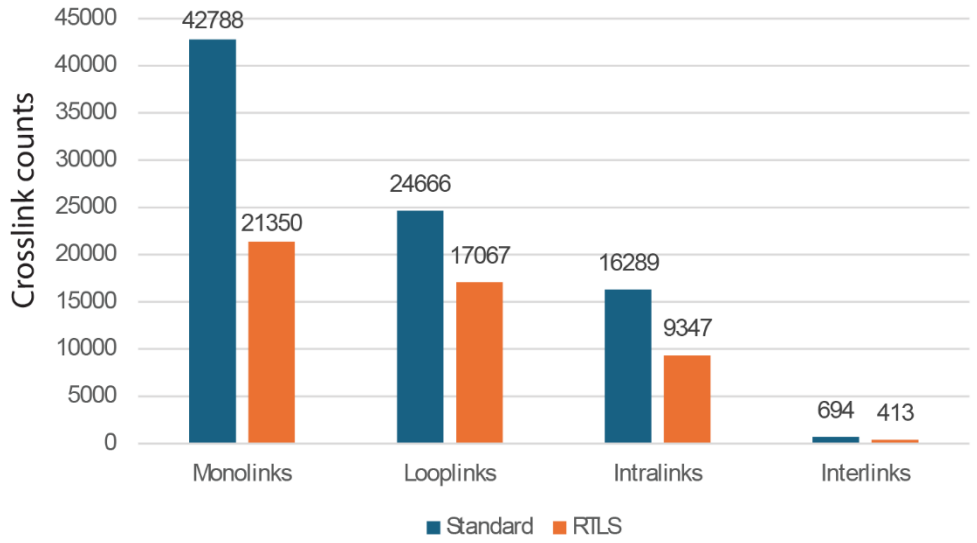
(**B**) Visualization of the co-localization of GFP-HTB11 and mCherry-WSD4 in *N.benthamiana.* Green: GFP-HTB11; red: mCherry-WSD4. Overlapping signals are indicated by arrows, with a zoom-in view highlighted in the inset. Scale bar=10μm.

(**C**) XL-MS identifies interaction between histones and PARP1, NEDD8 and WSD4, suggesting their potential roles in histone modifications and polynucleosome organization. Multiple cross-links between histone H2B proteins and WSD4, an O-acyltransferase, support a possible function for WSD4 in histone acylation.



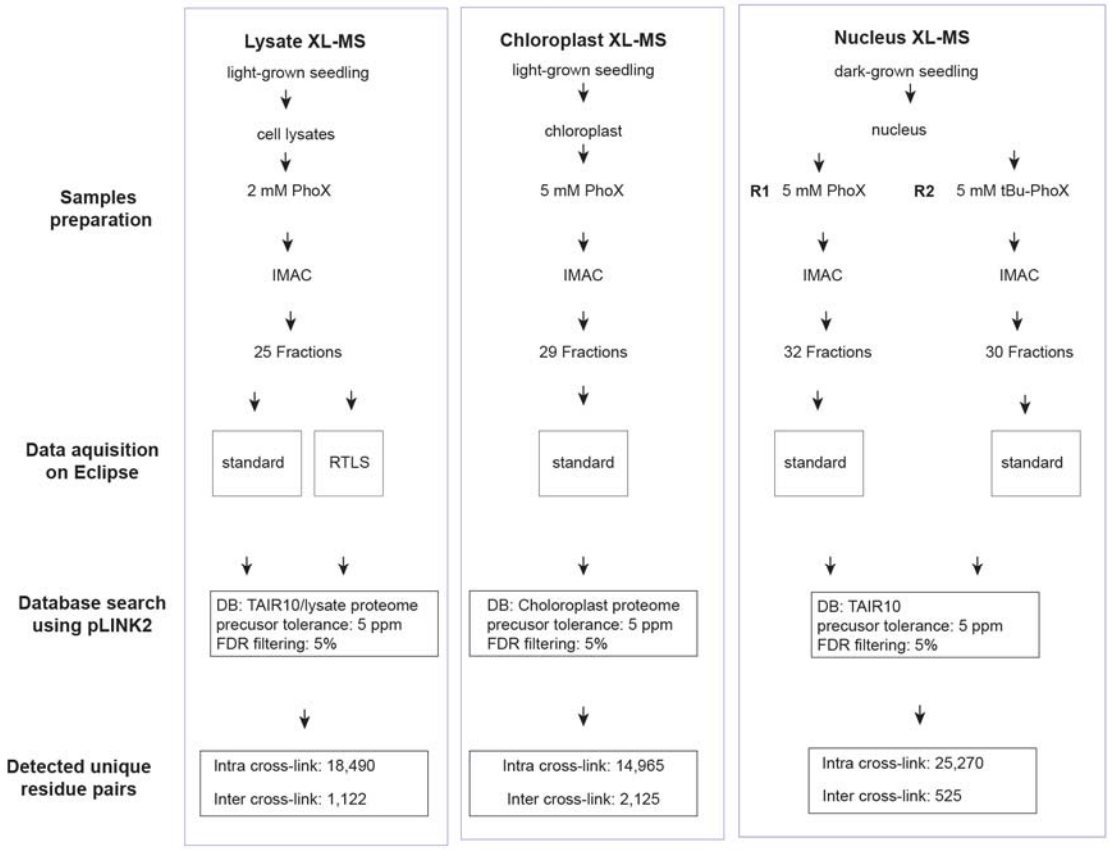
Supplementary Figure S1: Optimization of the PhoX cross-linker concentration for XL-MS.

(**A**-**D**) Different concentrations of PhoX cross-linker (0 to 10 mM) were tested on cell lysates from light (A, C, D) or dark grown seedlings (B). Increased protein cross-linking, indicated by the accumulation of higher molecular weight proteins (red arrow), was observed with increasing cross-linker concentration. However, no additional visible cross-linking was detected at higher concentrations (7-10 mM).

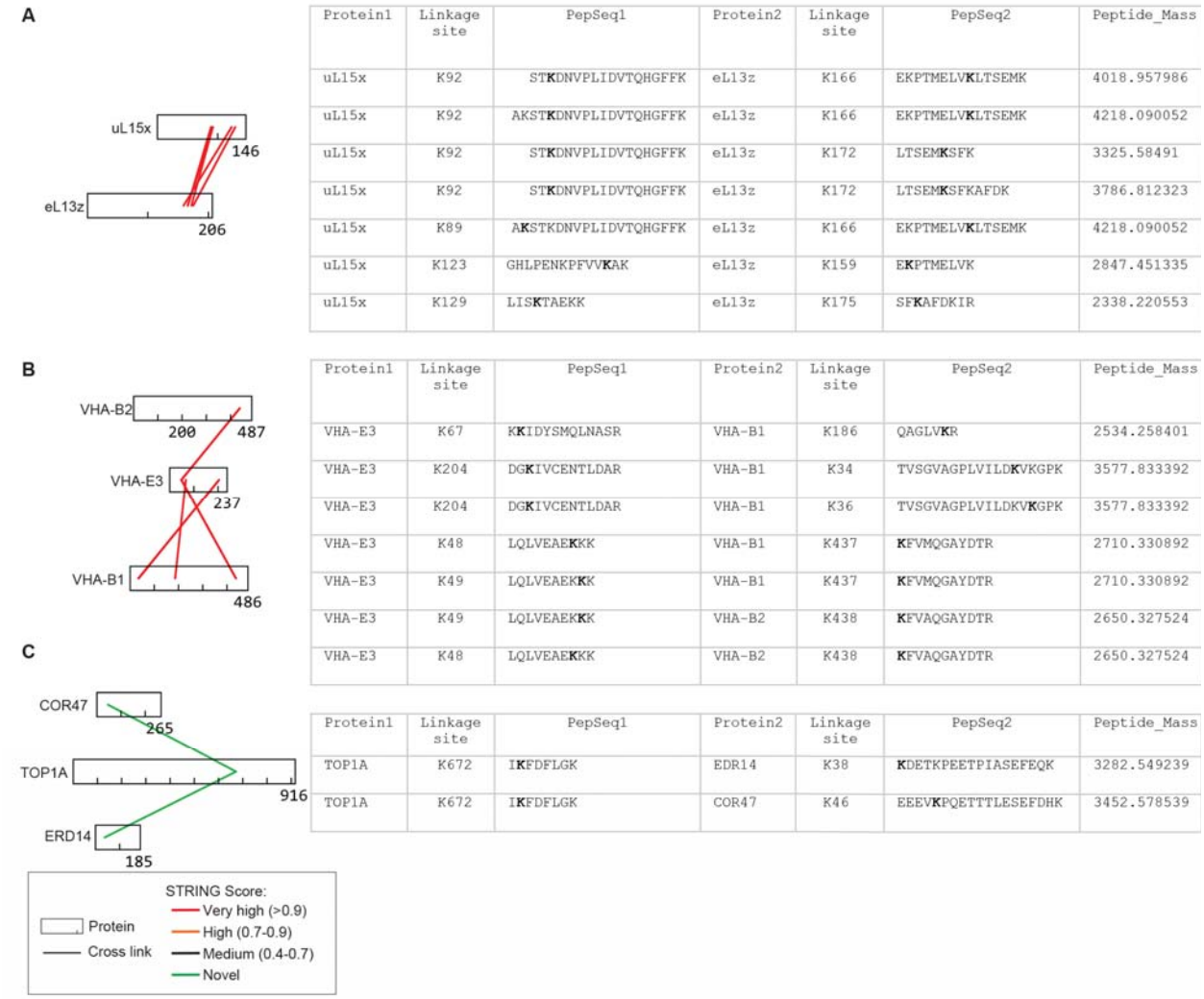


Supplementary Figure S2: **Standard XL acquisition outperforms Real-Time Library Search (RTLS) acquisition.**

Data were acquired on an Orbitrap Eclipse mass spectrometer using PhoX-enriched cross-linked peptides from Arabidopsis cell lysate samples, then searched against the entire Arabidopsis protein database. RTLS acquisition significantly reduced the identification of monolinks, where one NHS ester group is cross-linked to a peptide while the other is hydrolyzed, by approximately 50%, and loop links, which are cross-links within a single peptide, by about 31%. However, RTLS also substantially decreased the detection of inter-peptide cross-links, including both intraprotein (intralinks) and interprotein (interlinks) cross-links, by 43% and 40%, respectively.



Supplementary Fig. S3: Summary of XL-MS experiments.
Key details of sample preparation, data acquisition using the Orbitrap Eclipse mass spectrometer, and database searching using pLINK2 are provided. A summary of the unique cross-linked peptide residue pairs identified across experiments is also included.



Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
uL15x	K92	STKDNVPLIDVTQHGFFK	eL13z	K166	EKPTMELVKLTSEMK	4018.957986
uL15x	K92	AKSTKDNVPLIDVTQHGFFK	eL13z	K166	EKPTMELVKLTSEMK	4218.090052
uL15x	K92	STKDNVPLIDVTQHGFFK	eL13z	K172	LTSEMKSFK	3325.58491
uL15x	K92	STKDNVPLIDVTQHGFFK	eL13z	K172	LTSEMKSFKAFDK	3786.812323
uL15x	K89	AKSTKDNVPLIDVTQHGFFK	eL13z	K166	EKPTMELVKLTSEMK	4218.090052
uL15x	K123	GHLPENKPFVVKAK	eL13z	K159	EKPTMELVK	2847.451335
uL15x	K129	LISKTAEKK	eL13z	K175	SFKAFDKIR	2338.220553

Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
VHA-E3	K67	KKIDYSMQLNASR	VHA-B1	K186	QAGLVKR	2534.258401
VHA-E3	K204	DGKIVCENTLDAR	VHA-B1	K34	TVSGVAGPLVILDKVKGPK	3577.833392
VHA-E3	K204	DGKIVCENTLDAR	VHA-B1	K36	TVSGVAGPLVILDKVKGPK	3577.833392
VHA-E3	K48	LQLVEAEKKK	VHA-B1	K437	KFVMQGAYDTR	2710.330892
VHA-E3	K49	LQLVEAEKKK	VHA-B1	K437	KFVMQGAYDTR	2710.330892
VHA-E3	K49	LQLVEAEKKK	VHA-B2	K438	KFVAQGAYDTR	2650.327524
VHA-E3	K48	LQLVEAEKKK	VHA-B2	K438	KFVAQGAYDTR	2650.327524

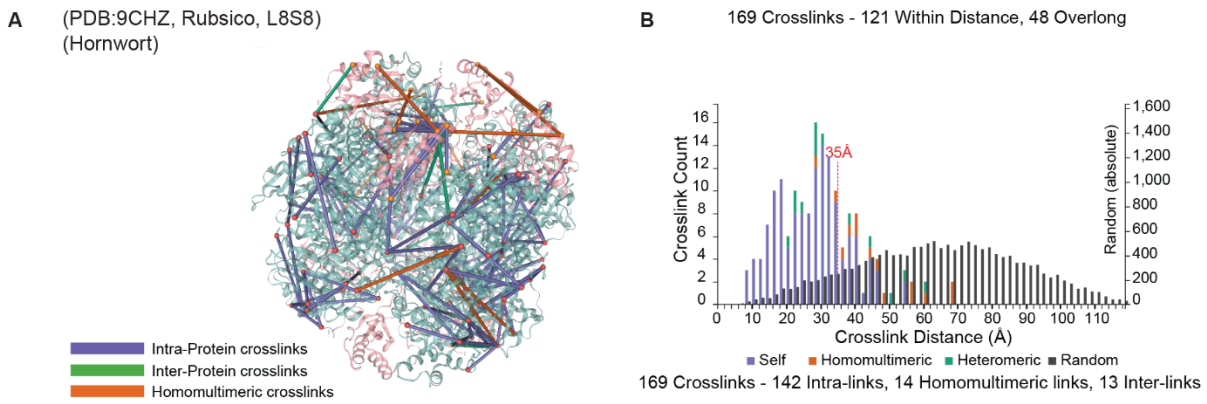
Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
TOP1A	K672	IKFDFLGK	EDR14	K38	KDETKPEETPIASEFEQK	3282.549239
TOP1A	K672	IKFDFLGK	COR47	K46	EEEVKPQETTTLESEFDHK	3452.578539

Supplementary Fig. S4. Representative cross-links supporting the high confidence of XL-MS analysis.

(**A**) Multiple cross-links were identified between the ribosomal proteins uL15x and eL13z, including peptide pairs with different degrees of miscleavage cross-linked at identical sites (e.g., uL15x-K92 to eL13z-K166/K172).

(**B**) Cross-links map conserved regions, with the ATPase subunit VHA-E3 (K49) cross-linked to both VHA-B1 (K437) and VHA-B2 (K438).

(**C**) Similar cross-linking patterns were observed between homologous interactions, as shown by DNA topoisomerase 1 alpha (TOP1A-K672) cross-linked to two dehydrin proteins, ERD14-K38 and COR47-K46.

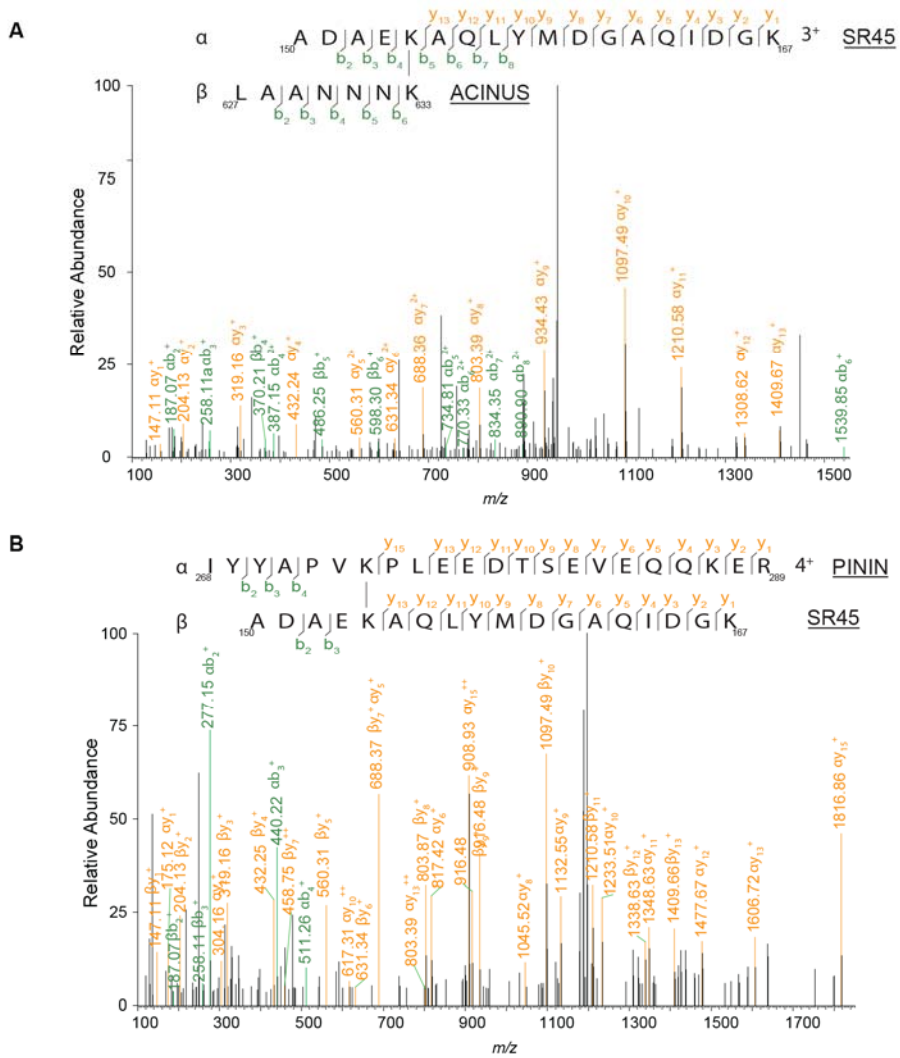


Supplementary Figure S5. Mapping of Rubisco cross-links on a high-resolution orthogonal structure of hornwort Rubisco.

Cross-links are color-coded by type: purple for self (intra)links, green for heteromeric (inter)links, and orange for homomultimeric links.

(**A**) Cross-links were mapped onto the high-resolution structure of L8S8 heterotetrameric Rubisco from hornwort (PDB: 9CHZ). RBCL subunits are shown in cyan and RBCS subunits in pink.

(**B**) Distance distribution of cross-links mapped onto Rubisco 9CHZ, showing that 71.6% of the identified cross-links fall within the expected spatial constraint of 35 Å.

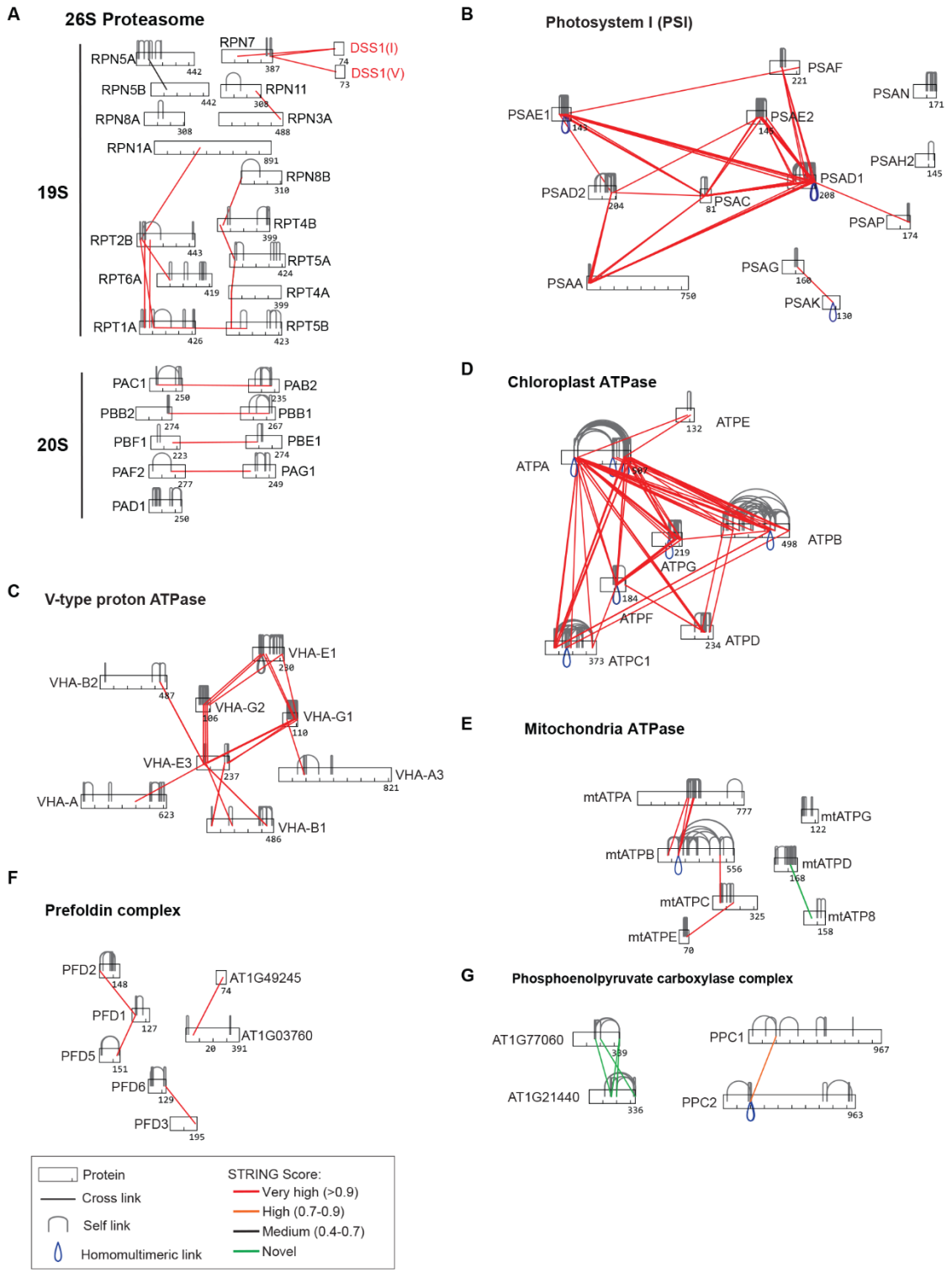


Supplementary Figure S6. High-confidence identifications of cross-linked peptides between ACINUS–SR45 and PININ–SR45.

b-ions are shown in green and y-ions in yellow. A series of fragment ions, including those from α- and β-peptides, as well as cross-linker-containing ions, were detected, supporting the reliability of the cross-link identifications.

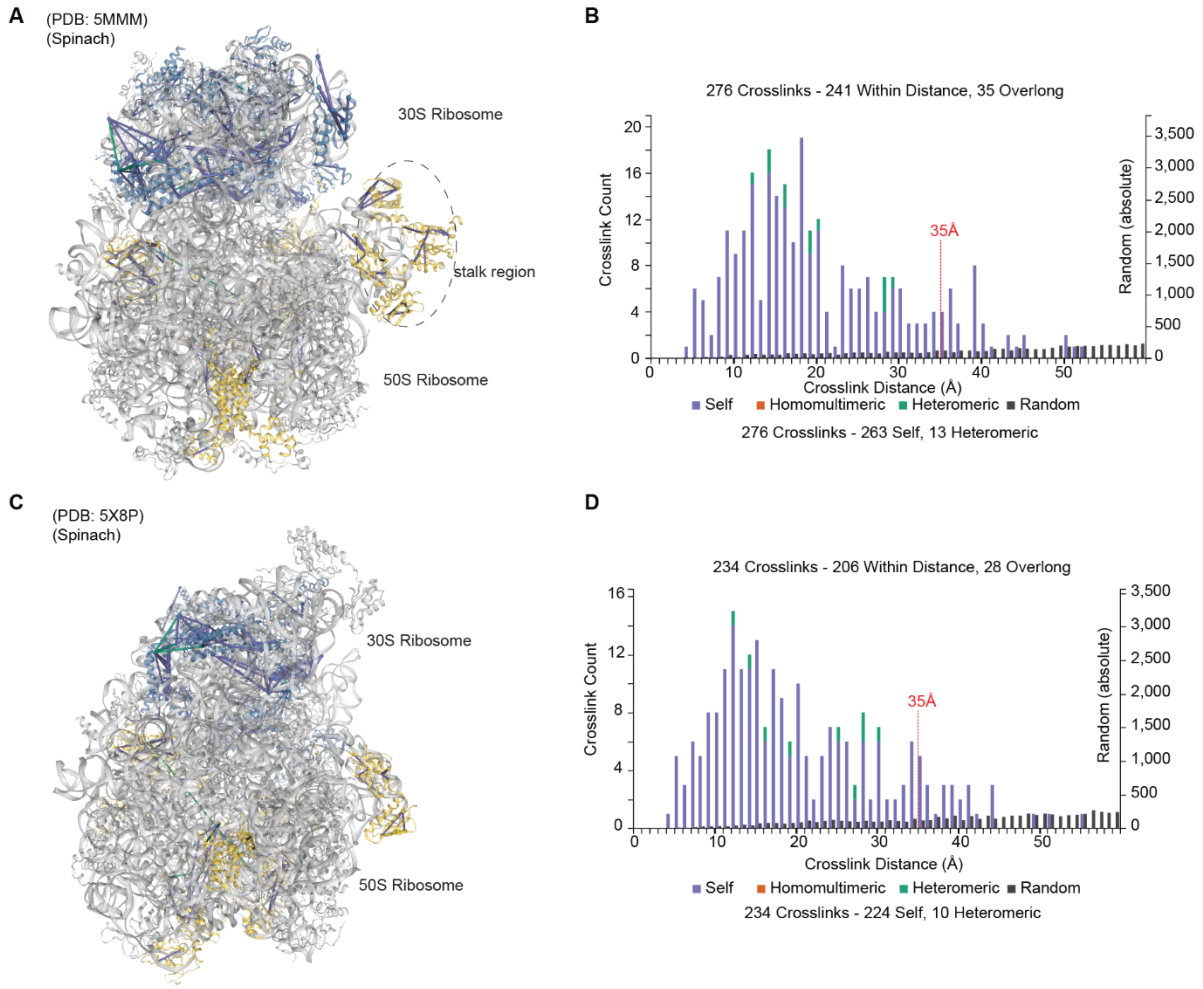
(**A**) Spectrum of the cross-linked peptide between ACINUS-K633 and SR45-K154.

(**B**) Spectrum of the cross-linked peptide between PININ-K274 and SR45-K154.

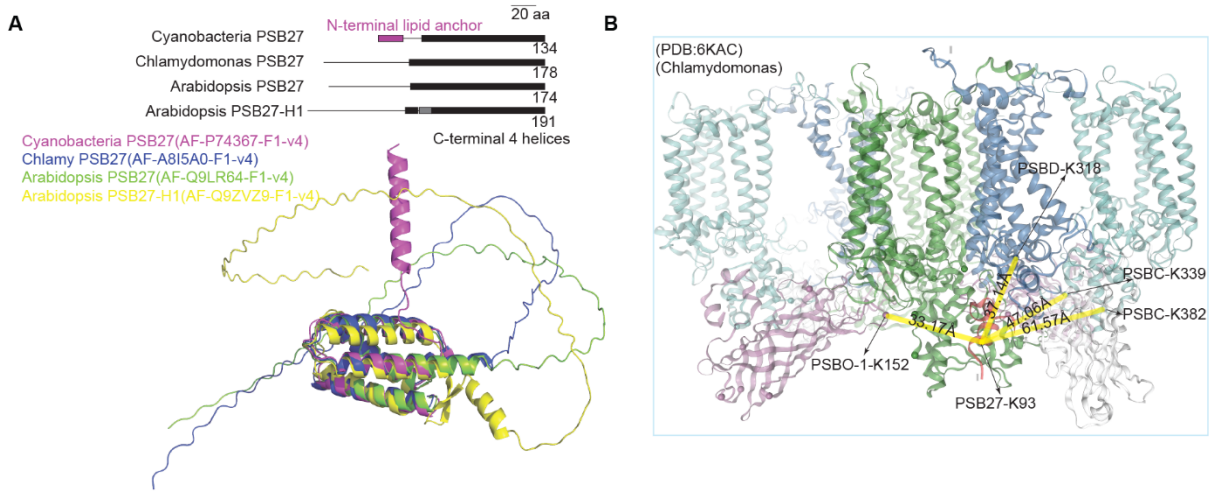


Supplementary Fig. S7. Topological mapping of multiple protein complexes by XL-MS. Proteins are shown as rectangles scaled to their respective lengths. Cross-links are color-coded based on STRING score confidence thresholds, highlighting both known and novel interactions.

(**A**) Mapping of 26S proteasome subunits, highlighting interaction interfaces between RPN7 and DSS1(I) and DSS1(V). DSS1 proteins are particularly challenging to map in cryo-EM due to their small size, intrinsic disorder, and dynamic interaction properties.
(**B**) Mapping of photosystem I (PSI) showing connectivity between different subunits.
(**C**-**E**) Mapping of different ATPase complexes, including chloroplast, V-type, and mitochondrial ATPases.
(**F**) Mapping of the prefoldin complex.
(**G**) Mapping of the phosphoenolpyruvate carboxylase (PEPC) complex.



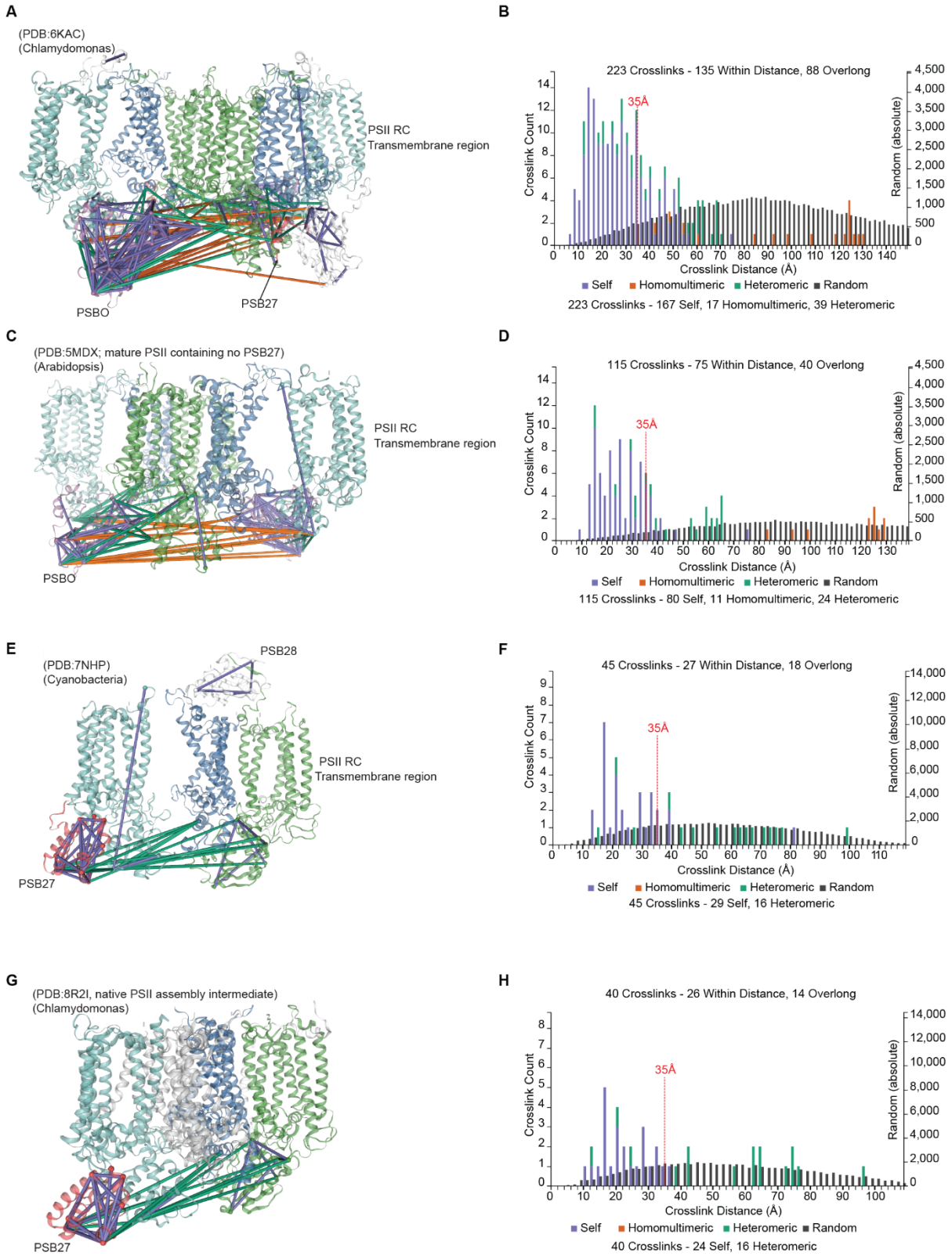
Supplementary Figure S8. Mapping of cross-links in the 70S ribosome complex onto high-resolution structures of spinach 70S ribosomes (PDB: 5MMM and 5X8P).
Cross-links are color-coded by type: purple for intralinks, green for heteromeric inter-protein cross-links, and orange for homomultimeric inter-protein cross-links.
(**A**) Structural mapping of 70S ribosome cross-links on the spinach cryo-EM structure (PDB: 5MMM), which lacks bL12 and other flexible proteins in the stalk region.
(**B**) Distance distribution of the cross-links mapped in (A). Of the 424 total cross-links detected, 276 were mapped (263 self-links and 13 heteromeric), with 241 falling within the 35 Å spatial constraint.
(**C**) Structural mapping of 70S ribosome cross-links onto an alternative spinach cryo-EM structure (PDB: 5X8P), which also lacks most of the flexible proteins in the stalk region.
(**D**) Distance distribution of the cross-links mapped in (C). Of the 424 cross-links detected, 234 were mapped (224 self-links and 10 heteromeric), with 206 falling within the 35 Å constraint.



Supplementary Figure S9. Mapping of PSB27 cross-links onto a high-resolution structure of 6KAC from Chlamydomonas.

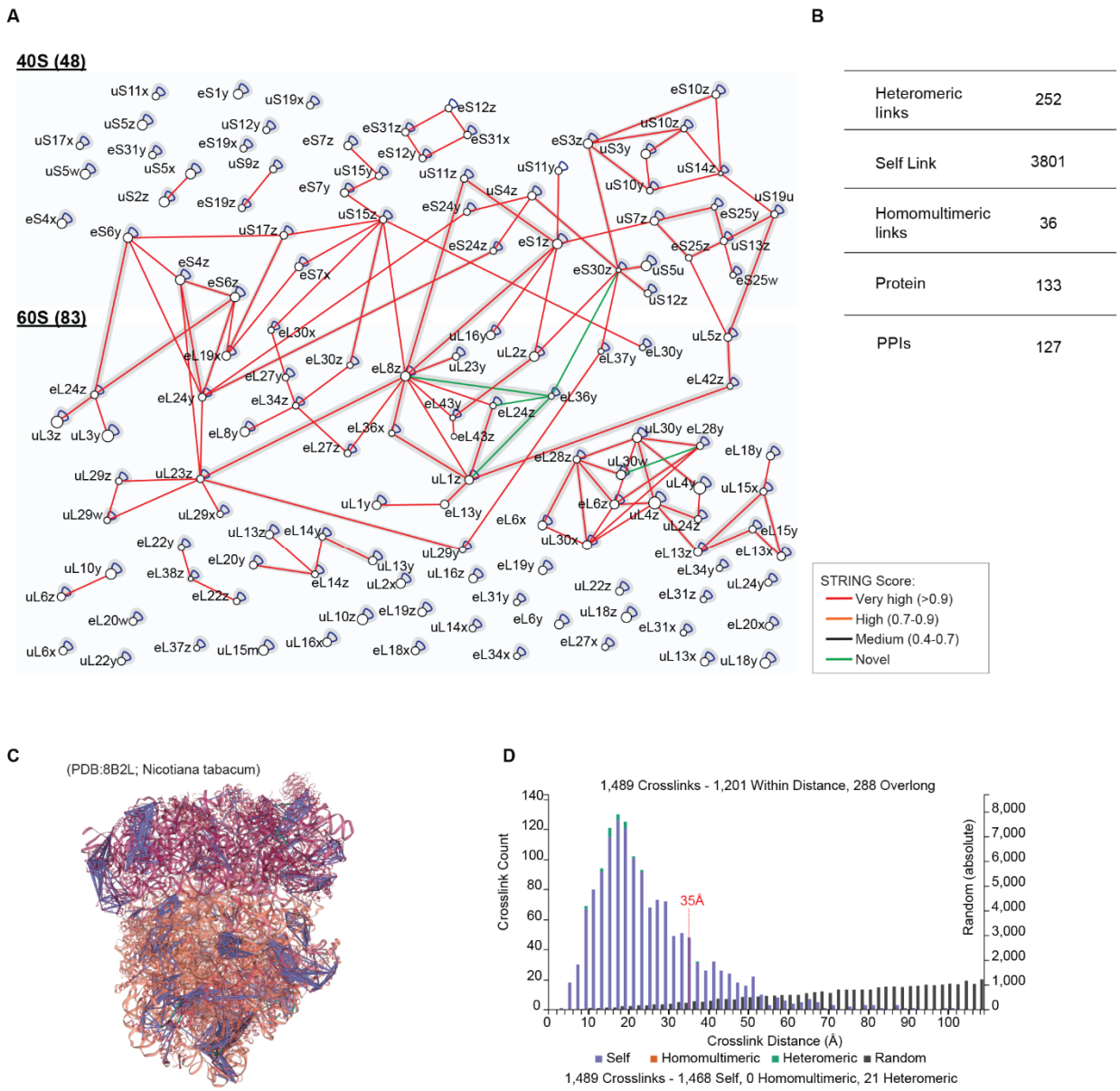
(**A**) Superimposed AlphaFold structures of PSB27 homologs reveal unique features of Arabidopsis PSB27. While it contains a conserved C-terminal four-helix bundle, Arabidopsis PSB27 lacks a lipid anchor and shows divergence in sequence and surface charge.

(**B**) Cross-links between Arabidopsis PSB27 and PSII core/oxygen-evolving complex (OEC) subunits mapped onto the Chlamydomonas PSII structure (PDB: 6KAC), which captures residues 91-114 of PSB27. The distance constraints suggest a more dynamic interaction between PSB27 and the PSII complex.



Supplementary Figure S10. Mapping of PSII complex cross-links onto several available high-resolution structures. A total of 581 cross-links were mapped, including 504 self-links

and 77 heteromeric links. Cross-links are colored according to the type of cross-link (purple for self (intra) link; green for heteromeric (inter) link; orange for homomultimeric link). The distribution of cross-link distances from the integrated cross-link model mapping is shown alongside the structural mapping.
(**A-B**) Structure mapping of PSII cross-links onto the Chlamydomonas PSII-LHCII supercomplex (PDB: 6KAC), which includes the PSII reaction center, PsbO, and a short sequence of Psb27.
(**C-D**) Structure mapping of PSII cross-links onto the Arabidopsis PSII-LHCII supercomplex, (PDB: 5MDX) which represents mature PSII, containing the PSII reaction center and PsbO but lacking Psb27.
(**E-F)** Cross-links mapped onto the cyanobacterial PSII assembly intermediate structure (PDB: 7NHP), which includes the PSII reaction center together with Psb27 and Psb28. Cyanobacteria have a single copy of the PSB27 gene. Fewer cross-links were mapped onto this assembly intermediate.
(**G**-**H**) Cross-links mapped onto the *Chlamydomonas* PSII assembly intermediate structure (PDB: 8R2I), representing a native assembly intermediate containing the PSII reaction center and Psb27. Similarly, fewer cross-links were mapped to this structure.



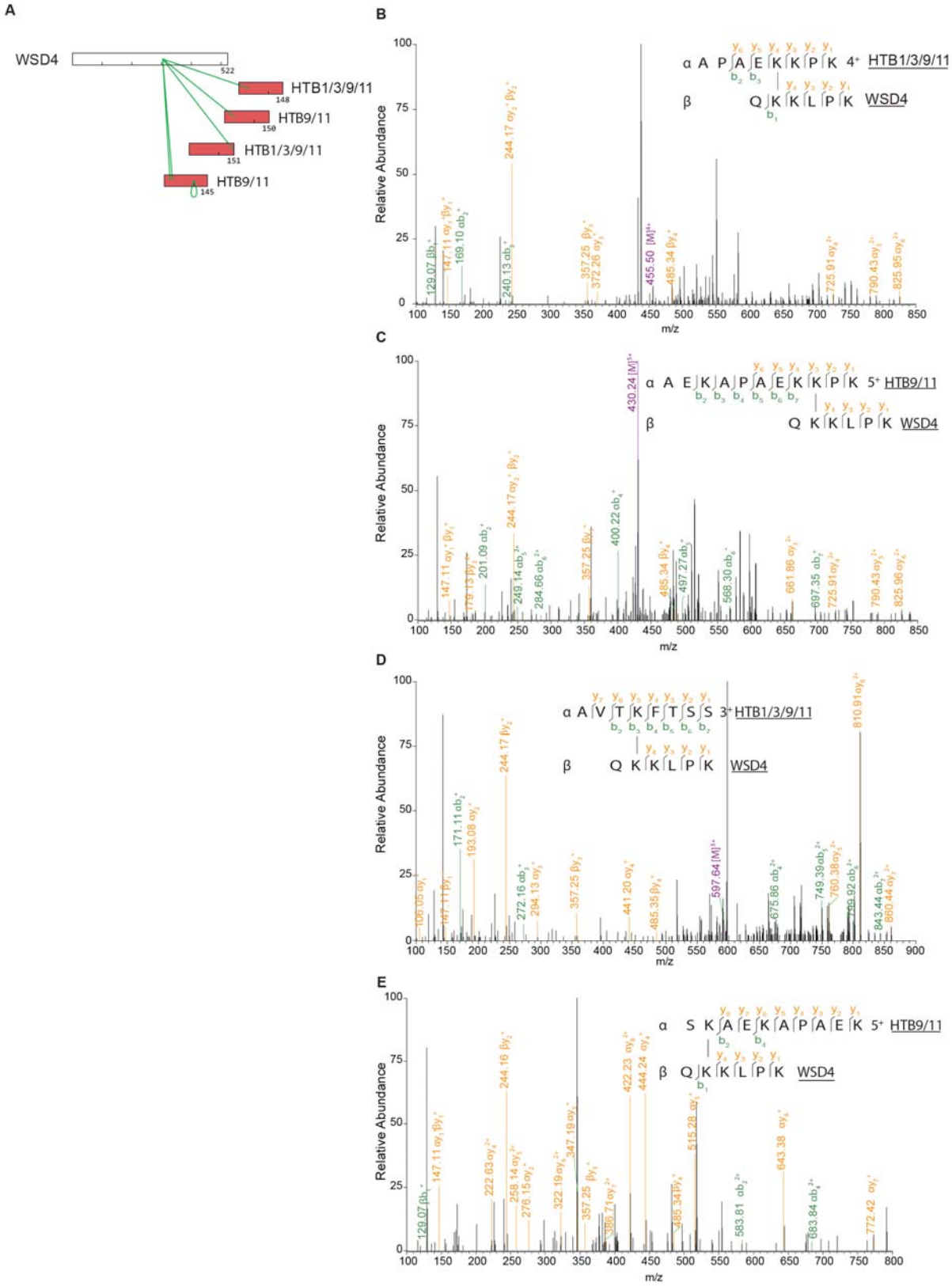
Supplementary Figure S11. Topology of the 80S ribosomal complex.

(**A**) 2D topological arrangement of the 40S and 60S ribosomal subunits based on their connectivity. Links between proteins are colored according to the confidence thresholds of the STRING score.

(**B**) Summary of cross-links and protein-protein interaction (PPI) information from (A).

(**C**) 80S ribosome cross-links mapped onto the high-resolution structure of the tobacco 80S ribosome (PDB: 8B2L). The large ribosomal subunit is colored orange and the small ribosomal subunit is colored red. Links are color-coded according to link type: purple for self-links (intralinks), green for heteromeric links (interlinks), and orange for homomultimeric links.

(**D**) Distribution of cross-link distances from the integrated cross-link model in (C).



Supplementary Fig. S12. XL-MS identifies WSD4 as a novel histone interactor that associates with H2B family members.

(**A**) Cross-links were detected between WSD4 (an O-acyltransferase) and several H2B family members, including HTB1, HTB9, HTB3, and HTB11.
(**B**-**E**) High-confidence cross-link spectra reveal interactions between WSD4 and H2B family members. Series of fragment ions-including those from α- and β-peptides, as well as cross-linker-containing ions-were detected, supporting the reliability of the cross-link identifications.