Mapping Architecture of Protein complexes in Arabidopsis using XL-MS

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

Abstract

Capturing molecular machines in action is essential for understanding protein complex architecture, cellular regulation, and gene function. Here, we establish a cross-linking mass spectrometry (XL-MS) platform using the PhoX cross-linker for proteome-wide interaction mapping in Arabidopsis under semi-native conditions. From cell lysates, chloroplasts, and nuclei, we identified 47,119 unique cross-links, including 3,527 inter-protein cross-links, and constructed a high-confidence protein-protein interaction (PPI) network containing 1,229 proteins and 1,446 PPIs. In silico analysis (STRING) confirmed 637 heteromeric interactions, with the remainder representing novel interactions. Our dataset provides experimental evidence for direct interactions and defines binding interfaces. XL-MS resolved the topology of challenging proteins and revealed both conserved and unique interactions, as exemplified in photosystem complexes, ribosomes, and histone interactomes. By characterizing the histone machinery, we identified a potential novel histone modifying enzyme. This work advances protein annotation and provides new insights into cellular organization and function.

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

47 PhoX. PhoX is a phosphonic acid-based reagent that, like BS3 and DSSO, contains NHS esters
48 and primarily targets lysine residues (11). Its phosphonic acid handle allows efficient enrichment
49 by immobilized metal affinity chromatography (IMAC). With a maximum range of 35Å -
50 consisting of a 20 Å Cα-Cα distance from its rigid structure and an additional 15 Å to account for
51 protein flexibility - PhoX is well suited for high confidence, large-scale structural proteomics
52 (12). To reduce sample complexity and improve cross-linking efficiency, we used subcellular
53 fractionations. To further improve the detection of the low abundance cross-linked peptides, we
54 performed extensive high-pH reversed-phase HPLC after IMAC enrichment. Together, these
55 measures allowed for comprehensive proteome-wide XL-MS mapping.

56 From cell lysates, chloroplasts, and nuclei, we identified a total of 47,119 unique cross-
57 linked peptide pairs, including 43,592 intra-protein (self-link) and 3,527 inter-protein cross-links.
58 These data allowed the construction of a PPI network comprising 1,229 proteins and 1,446
59 interactions, including 1,390 heteromeric and 56 homomultimeric interactions. STRING analysis
60 validated 637 heteromeric PPIs (combined confidence score ≥ 0.4), while the remainder
61 represent previously uncharacterized interactions. Our dataset provides experimental evidence
62 for direct physical binding and defines protein binding interfaces at amino acid resolution for
63 both known and novel interactions.

64 Beyond interaction mapping, our XL-MS data provide structural and functional insights
65 into key cellular processes. By resolving flexible or transient components within the 70S
66 chloroplast ribosome and photosystem complexes, our results complement existing cryo-EM
67 structures. XL-MS also enabled a comprehensive mapping of the photosystem, 80S ribosome,
68 and histone interactomes within heterogeneous protein populations, revealing structural biology
69 and capturing both known and novel interactions. In particular, we identified a metabolic O-

70 acyltransferase that interacts with histones, suggesting a potential new mechanism of chromatin
71 regulation via post-translational modification. Together, these results improve protein annotation,
72 shed light on protein complex architectures, and provide as a valuable resource for the broader
73 research community - laying the foundation for future discoveries.

74 **Results**

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

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

83 Optimization of the PhoX concentration showed a shift to higher molecular weights at 2-
84 5 mM (red arrows, Fig. S1A-D), indicating successful cross-linking, with no additional
85 improvement at higher concentrations. Based on these results, we used 2 and 5 mM PhoX for
86 subsequent experiments (see Methods). After cross-linking, we precipitated the proteins with
87 phenol, digested them with trypsin, and enriched the cross-linked peptides using $GaCl_3$-based
88 IMAC (13). We then fractionated the enriched peptides by high-pH reversed-phase
89 chromatography, analyzed them by LC-MS/MS, and identified cross-links using pLINK2 (14).
90 High-resolution, high-accuracy data for both MS1 (precursor ions) and MS2 (fragment ions)
91 were acquired on an Orbitrap Eclipse mass spectrometer to ensure reliable cross-link
92 identification.

93 To optimize data acquisition, we compared the standard XL workflow with real-time
94 spectral library search (RTLS) (15), both targeting highly charged precursors (+3 and above).
95 RTLS uses a survey MS2 scan to reduce mono-link detection. In cell lysates, RTLS reduced
96 mono- and loop-link identifications by ~50% and ~31%, respectively (Fig. S2), but also
97 significantly reduced intra-link (43%) and inter-link (40%) identifications. Because of this trade-
98 off, we used the standard XL workflow to analyze cross-linked chloroplast and nuclear samples.
99 In total, we identified 47,119 unique cross-linked peptide pairs (intra- and inter-links)
100 representing 5,759 proteins. Specifically, we detected 25,795, 17,090, and 19,612 cross-links
101 from the nucleus, chloroplast, and cell lysate, respectively (Fig. 1B and. S3, Table S1).
102 Subcellular localization analysis using SUBA5 revealed that the cross-linked proteins were
103 distributed in various cellular compartments, including the nucleus, plastid, cytosol,
104 mitochondria and other organelles (Fig. 1C, Table S2).
105 To assess the quality of cross-linked peptide identification, we examined the identified
106 cross-links (Table S1) and found consistent peptide pairs mapping to the same sites or regions.
107 For example, in ribosomal subunits, we detected site-specific cross-links with variable
108 miscleavages, such as those between uL15x-K92 and eL13z-K166/K172 (Fig. S4A). Conserved
109 region cross-links were also observed, including VHA-E3-K49 with both VHA-B1-K437 and
110 VHA-B2-K438 (Fig. S4B), and TOP1A-K672 with ERD14-K38 and COR47-K46 (Fig. S4C).
111 These independent and consistent observations underscore the robustness and reliability of our
112 results.
113 **Validation of intra- and inter-protein cross-links by structural mapping**
114 To evaluate the validity of the cross-links, we first used ComMap (16) to map intra-links to high-
115 resolution protein structures available in the Protein Data Bank (PDB) and predicted structures
116 by AlphaFold. Specifically, 721 Arabidopsis proteins had high-resolution structures in the PDB,

117 and 3,975 were available through AlphaFold (Table S1). A total of 21,981 cross-links were
118 successfully mapped, with 94% of the PDB-mapped cross-links and 89% of the AlphaFold-
119 mapped cross-links satisfying the 35 Å distance constraint (Fig. 2A-B).

120 To evaluate both intra- and inter-protein cross-links, we mapped 174 cross-links onto the
121 L_8S_8 Rubisco assembly modeled from the Arabidopsis L_2S_2 crystal structure (PDB:5IU0) (17).
122 Of these, 169 cross-links were mapped to the large (479 aa) and small (181 aa) subunits, while
123 five were localized to the N-terminal residues (1–11) of RBCL, which are not present in the PDB
124 structure (Fig. 2C). The majority (126, 74.6%) of mapped cross-links fell within the expected
125 Cα-Cα distance of ≤35 Å; the remaining 43 (25.4%) exceeded this range, suggesting regions of
126 structural flexibility (Fig. 2D–E). Similar mapping results were obtained using the hornwort
127 Rubisco cryo-EM structure, which yielded similar results (Fig. S5A-B), further supporting the
128 validity of our cross-linking data.

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

139 **XL-MS systematically identifies direct protein-protein interactions and reveals known and**
140 **novel complexes**

141 We imported 3,527 inter-protein cross-links into Cytoscape (23), revealing 1,446 PPIs
142 among 1,229 proteins, including 1,390 heteromeric and 56 homomultimeric interactions (Fig.
143 3A, Table S1). Proteins were classified using the PANTHER protein class tool, and nodes were
144 colored accordingly (24). Many interacting proteins shared the same classification, supporting
145 the reliability of our PPI map. Notably, 457 of the 1,210 proteins involved in heteromeric
146 interactions lacked PANTHER annotations, contributing to 819 interactions - 369 between
147 unclassified proteins and 450 between classified and unclassified proteins.

148 We also incorporated STRING confidence scores (25) into the network and color-coded
149 the edges accordingly (Fig. 3A). These scores combine evidence from eight sources (Fig.3B)
150 (25). For heteromeric PPIs, 38.2% of PPIs had high STRING confidence: 388 scored ≥ 0.9 and
151 144 scored 0.7-0.9. The remainder included 105 medium (0.4-0.7), 144 low (<0.4), and 609 with
152 no STRING score (Fig.3B). The PANTHER and STRING results indicate that both known and
153 novel PPIs were captured, highlighting the power of systematic XL-MS to reveal previously
154 uncharacterized protein complexes and pathways.

155 Even among PPIs with good STRING combined scores, many lack experimental support
156 for direct or indirect interactions in Arabidopsis or homologs in other species. Among the
157 heteromeric PPIs identified, only 28.9% (n = 402) have experimental support (score ≥ 0.4), while
158 71.1% have low or no scores (Fig. 3B). Our XL-MS data provide evidence for their direct
159 interactions. For example, ORP3B (AT3G09300, oxysterol-binding protein) and PLDRP1
160 (AT5G39570, PLD-regulated protein 1) have a combined score of 0.602 but a low experimental
161 score (0.055); we identified a cross-link between ORP3B-K7 and PLDRP1-K193 (Fig. 3C),

162 suggesting a common role in lipid trafficking and signaling. Similarly, NRPA2 (AT1G29940,
163 RNA polymerase predicted to localize to both the nucleus and chloroplasts) and uL1c
164 (AT3G63490, chloroplast 50S ribosomal protein) have a combined score of 0.588 but no
165 experimental support; a cross-link between NRPA2-K394 and uL1c-K127 suggests a direct
166 interaction and potential coupling of chloroplast transcription and translation (Fig. 3D).

167 **XL-MS reveals unique topology of chloroplast 70S ribosome stalk regions**

168 Using XL-MS, we mapped the topologies of several protein complexes (Fig. S7A-G),
169 including the flexible stalk regions of the chloroplast 70S ribosome (Fig. 4A-C), which have
170 remained difficult to resolve due to the dynamic nature of proteins such as uL1c, uL10c, uL11c,
171 and bL12c (26, 27). In total, we identified 424 cross-links across 34 ribosomal subunits of the
172 70S ribosome, including 352 intraprotein, 64 heteromeric, and 8 homomultimeric cross-links.
173 Notably, many intra- and interprotein cross-links involved these flexible stalk proteins. Although
174 cross-links were mapped to spinach 70S ribosome structures (PDB: 5MMM and 5X8P; Fig.
175 S8A-B), most cross-links involving uL1c, uL10c, uL11c, and bL12c could not be precisely
176 assigned to specific locations within the stalk regions because these proteins are either missing or
177 only partially resolved in the cryo-EM structures.

178 Arabidopsis bL12c has a plant-specific disordered N-terminus, followed by conserved
179 structured N-terminal (NTD) and C-terminal (CTD) domains connected by a flexible hinge (Fig.
180 4B). Essential for translation, bL12c uses hinge-mediated mobility to recruit specific translation
181 factors. XL-MS identified interactions between bL12c (the two highly homologous proteins
182 AT3G27850/bL12cx and AT3G27830/bL12cz) and several ribosomal core proteins, as well as
183 the elongation factor CPEF-G, mainly through its structured CTD (Fig. 4C). Although bL12c
184 dimers are typically mediated by the NTD (28), XL-MS revealed unexpected homodimer cross-

185 links within the CTD - the same region involved in translation factor binding (Fig. 4C). This
186 suggests that CTD dimerization may serve as an interface for the recruitment of interaction
187 partners. Taken together, these results highlight both conserved and plant-specific features of
188 bL12c and demonstrate the power of XL-MS in resolving dynamic aspects of ribosomal
189 architecture.

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

191 Photosystems I and II (PSI and PSII) and their associated light-harvesting complexes
192 (LHCI and LHCII) are essential for light energy harvesting and carbon fixation, but they require
193 continuous repair, especially under high and fluctuating light conditions (29-31). Although
194 progress has been made in understanding certain steps of the PSII repair (32, 33), many aspects,
195 including the detailed mechanism, identification of repair intermediates, and specific
196 contributing factors, remain unresolved.

197 Using XL-MS, we identified extensive interactions involving the PSII repair factor Psb27
198 (AT1G03600) (Fig. 5A). Unlike its cyanobacterial counterpart, which functions in both assembly
199 and repair, Arabidopsis Psb27 appears to be specialized for repair. In contrast, the less
200 homologous Psb27-H1(AT1G05385) is involved in assembly. Notably, Arabidopsis Psb27 lacks
201 a lipid anchor and is structurally more flexible than the cyanobacterial version (34) (Fig. S9A). A
202 partial structure from Chlamydomonas (PDB: 6KAC) captured residues 91-114 of Psb27 (35).
203 Our XL-MS data revealed that Psb27 interacts with both the PSII reaction center and the oxygen-
204 evolving complex (OEC), involving K93 (as in 6KAC) and a second site at K168 (Fig. 5A, Fig.
205 S9B, S10A-H). These distance constraints (Fig. S9B) suggest that Psb27 engages PSII in a more
206 dynamic and flexible manner in plants.

207 XL-MS also revealed extensive interactions between PSI, PSII, LHCI, LHCII, ATP
208 synthase, ATP-dependent proteases (FtsH, SPPA) and transient proteins involved in state
209 transitions, complex assembly and repair (Fig. 5B). Notably, chlorophyll biosynthesis and
210 catabolic proteins (e.g., ACD1, CHLM, FLU) were found in close association with LHCII,
211 suggesting coordinated chlorophyll delivery and reactive oxygen species (ROS) control during
212 LHCII assembly. In addition, many assembly and repair factors were closely associated with
213 LHCII, indicating their readiness for rapid engagement during dynamic photosystem remodeling.

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

215 We identified extensive interactions involving the 80S ribosome, a complex of 60S and
216 40S subunits that synthesizes nuclear-encoded proteins. Ribosomes assemble in the nucleolus,
217 localize to the cytoplasm, and associate with the endoplasmic reticulum (ER). As a dynamic
218 macromolecular machine, the ribosome coordinates translation with numerous proteins and
219 RNAs, while translation errors can generate misfolded proteins that trigger ribosome-associated
220 quality control pathways (36). Using XL-MS, we mapped detailed structural information within
221 this heterogeneous complex, capturing not only intra- and inter-subunit contacts (Fig. S11A-D),
222 but also interactions with proteins involved in ribosome biogenesis, translation, folding, quality
223 control, stress response, and other regulatory pathways (Fig. 6A). These interactions were
224 resolved at the amino acid level, providing detailed insight into the architecture of the functional
225 interface (Fig. 6B, Table S4).

226 While many interactions showed good STRING scores (≥0.4), our analysis also revealed
227 numerous novel interactions (green edges) between specific protein pairs. Although homologs of
228 these proteins have been studied in other organisms for their ribosome-related functions, their
229 specific roles in plants remain largely uncharacterized. Our XL-MS data suggest conserved

230 functions: for example, the UTP7 homologs AT3G10530 and NOG1-1 may contribute to
231 ribosome biogenesis (37, 38), whereas AT5G50840 (α-taxilin-like) and ILITHYIA(ILA) may be
232 involved in protein folding or translation (39, 40). MBF1A, a transcriptional coactivator, may be
233 involved in stress responses at stalled ribosomes, similar to its role in yeast (41) . Other
234 candidates include a methyltransferase potentially involved in rRNA processing and polysome
235 assembly (42, 43) , a RIO kinase associated with ribosome biogenesis (44, 45), and PARP1,
236 which may regulate translation via ADP-ribosylation (46) (Fig. 6C). By mapping these
237 interactions to specific ribosomal proteins, our XL-MS dataset provides a valuable resource for
238 future functional studies of ribosome-associated factors in plants.
239 We examined ubiquitination-related interactors, including the E3 ligases EBF1 and
240 AT4G02760, and the deubiquitinase UBP14. However, we did not identify homologs of known
241 ribosome-targeting E3 ligases from other species. EBF1 regulates ethylene signaling by
242 targeting EIN3/EIL1 for degradation (47), whereas UBP14 is essential for development in
243 Arabidopsis but dispensable in yeast (48). Given the active turnover of defective ribosomes, we
244 confirmed the *in vivo* ubiquitination of 37 large and 27 small ribosomal subunits by ubiquitin
245 diGly proteomic analysis (Fig. 6A, Table S4). Together with observed interactions with the 26S
246 proteasome components RPN8A, PAD1 (Fig.6A), these results suggest the ribosome undergoes
247 ubiquitination and may be targeted for degradation by the 26S proteosome.
248 **Histone interactome reveals conserved and novel chromatin-associated interactions**
249 Our XL-MS analysis revealed extensive interactions between histones and both known and novel
250 interactors, providing new insights into the networks involved in gene expression and chromatin
251 regulation (Fig. 7A). We identified several nuclear complexes associated with chromatin,
252 including nucleosome assembly proteins (NAPs), chromosome structure maintenance proteins
253 (SMCs), nuclear RNA polymerase components (NRPs), chromatin regulators (PDP3, MSI4),

254 transcriptional activator and acetylation reader (AT2G36410 and ACIP1), and subunits of the
255 PAF complex (VIP2, VIP4) (Fig. 7A).
256 Major histone interactors include chaperones (SDJ2, NUC-L1), chromatin remodeling
257 factors (EBS, SYD, SPK1), a transcription factor (SPL3), co-regulators (SNL6), DNA repair
258 proteins (MSH2), and structural proteins such as HMGB7. While many of these proteins are
259 known to influence chromatin architecture and gene regulation, their direct interactions with
260 histones and their roles in plant systems remained largely unexplored. Notably, MOS7, a
261 nucleoporin homologous to human NUP88, cross-linked with several histones, providing direct
262 evidence for nuclear envelope-histone interactions that may contribute to chromatin organization
263 in plants (49).
264 In addition, we identified histone interactions with several enzymatic regulators,
265 including the kinase MLK4 and the poly(ADP-ribose) polymerase PARP1 (both cross-linked to
266 H1), and the ubiquitin-like protein NEDD8 (cross-linked to H2B and H3). While the roles of
267 PARP1 and NEDD8 in histone regulation have not been defined in plants, their involvement in
268 histone modification has been well documented in other organisms (50-53). MLK4, previously
269 shown to phosphorylate H2A to regulate flowering (54), was found to interact with H1,
270 suggesting an expanded role for MLK4 in histone regulation that warrants further investigation.
271 Strikingly, we also identified WSD4, an O-acyltransferase annotated as a wax ester
272 synthase/acyl coenzyme and predicted by TAIR to localize to the nucleus (Fig. 7A-B). Although
273 O-acyltransferases are typically associated with lipid metabolism (55), WSD4 was found to
274 cross-link with four H2B family members - HTB1, HTB9, HTB3 and HTB11- via four
275 independent cross-linked peptides (Fig. S12A-E). These data suggest that WSD4 may catalyze

histone acylation, possibly on serine or threonine residues, rather than the conventional lysine residues modified by other acylation proteins.

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

Discussion

Systematic mapping of PPIs has greatly advanced our understanding of cellular function, but plant PPI networks remain largely unexplored (5, 56). While comprehensive maps exist for humans and other model organisms (57-59), similar efforts are just beginning in plants (60-62). This gap represents a critical opportunity, as plants are central to addressing global challenges such as food security, climate change, and sustainability. Expanding proteome-wide PPI mapping and structural biology studies in plants will accelerate our mechanistic understanding of cellular processes and drive advances in agriculture and biotechnology.

We developed a systems biology approach to capture molecular machines in action by mapping protein-protein interactions (PPIs) and their interfaces using XL-MS. To overcome the challenges of XL-MS in complex biological samples, we established an effective pipeline integrating the PhoX crosslinker, subcellular fractionation, and high-pH reversed-phase HPLC fractionation, enabling the identification of over 47,000 crosslinks. PDB and AlphaFold mapping confirmed the quality of the data.

Our XL-MS analysis identified 1,446 high-confidence PPIs among 1,229 proteins, providing a comprehensive map of protein interactions in Arabidopsis. Functional categorization and STRING analysis revealed that these interactions span diverse biological processes, with 54.2% unannotated in STRING, demonstrating the ability of XL-MS to uncover novel networks.

299 The dataset also revealed the topological organization of protein complexes, including the 70S
300 ribosome, 26S proteasome, photosystems I and II, ATPases, and chromatin-associated factors.
301 XL-MS was particularly effective in resolving spatial relationships within dynamic protein
302 assemblies, such as mapping flexible proteins in the 70S ribosome stalk regions and transient
303 repair proteins in PSII.

304 XL-MS allows the detection of both known and novel PPIs in heterogeneous populations,
305 greatly expanding our understanding of conserved molecular machines such as ribosome and
306 histones (63, 64). While many of the identified interactors have characterized roles in other
307 species, their functions in plants remain unexplored. We discovered several ribosome-associated
308 E3 ligases and confirmed the *in vivo* ubiquitination of most ribosomal proteins, linking ribosome
309 function to ubiquitination-mediated regulation. In addition, our analysis revealed extensive plant-
310 specific PPIs, particularly in chloroplasts, highlighting unique aspects of the plant cellular
311 machinery.

312 Our XL-MS analysis revealed WSD-4, an O-acyltransferase, as a novel factor potentially
313 involved in histone regulation. WSD-4 forms cross-links with several members of the histone
314 H2B family and is predicted to localize in the nucleus. This family of O-acyltransferases is
315 typically involved in lipid metabolism (55); their role in protein modification has not been
316 reported. Interestingly, short-chain lysine acylations have been implicated in the regulation of
317 gene expression (65). Given the emerging evidence that metabolic enzymes such as ACAT1
318 (acetyl-CoA acetyltransferase 1) (66) and PORCN (membrane-bound O-acyltransferase with
319 protein-serine O-palmitoleoytransferase activity) (67, 68) also modify proteins, we propose that
320 WSD-4 may catalyze histone acylation, targeting serine or threonine residues. This provides a
321 novel mechanism of chromatin regulation, suggesting that WSD-4 works in concert with other

322 histone modifying enzymes to influence histone dynamics, chromatin structure, and gene
323 expression.

324 Despite Arabidopsis being a well-studied model organism, many of its genes remain
325 uncharacterized (69). This study highlights how XL-MS can map PPIs in semi-native conditions
326 and accelerate functional annotation, providing new insights into cellular organization. While
327 focused on Arabidopsis, our approach has broad applicability and provides a valuable resource
328 for comparative proteomic and structural studies across species. These findings improve protein
329 annotation, enhance understanding of protein complex architectures, and lay the foundation for
330 future discoveries.

331 **Acknowledgement**

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

340 **Author contributions**

341 S.L.X conceived the project and supervised the study. C.S.T. and R.S. designed and performed
342 most of the XL-MS experiments. R.S. developed the initial XL workflow. W.C.C performed
343 chloroplast XL-MS. S.K. performed ubiquitination enrichment. C.S.T and S.L.X analyzed all

344 the MS data and prepared figures, with assistance from A.V.R, G.L. and K.H. C.S.T and S.L.X
345 interpreted the data and wrote the manuscript with feedback from all authors.

346 **Data Availability**

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

350 **Supplemental Table**

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

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

355 Table S3: Analysis of protein classification using PANTHER.

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

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

359

360 Material and methods

361 Plant growth condition

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

366 Light-grown seedlings: plants were vertically grown in a growth chamber at 22°C under a
367 16-hour light / 8-hour dark cycle for 10-12 days, then flash frozen in liquid nitrogen and stored at
368 -80°C for crosslinking experiments on cell lysates and isolated chloroplasts. Dark-grown

369 seedlings: plants were grown horizontally at 22°C in complete darkness for 3 days in a growth
370 chamber, then flash frozen in liquid nitrogen and stored at -80°C for nucleus isolation. Dark
371 growth was used to reduce Rubisco content.

372 Frozen Arabidopsis tissue was cryo-ground to a fine powder using a Retsch MM400
373 Mixer Mill for downstream processing. Different amounts of tissue were used for cross-linking
374 experiments: 6 g from 10 d seedlings for cell lysates, 195 g from 12d seedlings for chloroplasts,
375 and 32 g (replicate 1 with PhoX) or 14 g (replicate 2 with tbu-PhoX) from 3d dark-grown
376 seedlings for nuclei.

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

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

383 *Chloroplast isolation***:** Frozen, homogenized tissue powder from 12 d old seedlings was
384 mixed with a grinding buffer similar to (70), using HEPES instead of Tricine for crosslinker
385 compatibility. Briefly, 5 g of tissue powder was ground in a mortar and pestle with 30ml of
386 grinding buffer [20mM HEPES-KOH pH 8.0, 330mM sorbitol, 5 mM EDTA, 5 mM EGTA, 5
387 mM $MgCl_2$, and 10 mM $NaHCO_3$] at 4°C, and under green illumination only. The resuspended
388 homogenate was filtered twice through two layers of Miracloth. 30 ml of the filtrate was
389 carefully transferred onto 10 ml of 40% Percoll [20 mM HEPES-KOH pH 8.0, 330 mM sorbitol,
390 40% v/v Percoll, 2.5 mM EDTA, 5 mM $MgCl_2$, and 10 mM $NaHCO_3$] in a 50ml centrifuge tube,
391 and centrifuged at 2,600g for 10 min at 4°C. The bottom pellet, enriched in intact chloroplasts,

392 was collected and resuspended in 30 ml of wash buffer [20 mM HEPES-KOH pH 8.0, 330 mM
393 sorbitol, 2.5 mM EDTA, 2 mM $MgCl_2$, and 10 mM $NaHCO_3$], and centrifuged at 2,600g for 3
394 min at 4°C. Each pellet derived from 5 g of tissue was washed twice with 4.5 ml of 1x PBS
395 supplemented with $MgCl_2$ [137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4,
396 and 2.5 mM $MgCl_2$] and then resuspended in 1mL 1x PBS supplemented with $MgCl_2$ for the
397 cross-linking reaction. This process was repeated across multiple batches, with a total of 195 g
398 of tissue used for cross-linking.

399 *Nuclear isolation*: The method for nuclear isolation was adapted from (71) with slight
400 modification. Powdered dark-growth Arabidopsis tissue of approximately 2 g was completely
401 homogenized with lysis buffer [20 mM HEPES (pH 7.4), 20 mM KCl, 2 mM EDTA, 2.5 mM
402 $MgCl_2$, 25% glycerol, 250 mM sucrose, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl
403 fluoride (PMSF), 1× protease inhibitor cocktail, and 0.2% formaldehyde] at tissue:buffer ratio of
404 1:5 (g/mL, w/v). The homogenate was then filtered through a 40 μm cell strainer and centrifuged
405 at 4°C, 1,500g for 10 min. The pellet was retained as the nuclear fraction. The nuclear pellet was
406 then washed 5 times with 5 mL of nuclear washing buffer [20 mM HEPES (pH 7.4), 2.5 mM
407 $MgCl_2$, 25% glycerol, 0.2% Triton X-100, 1 mM PMSF, and 1× protease inhibitor cocktail] and
408 centrifuged at 4°C, 1,500g for 10 min to collect the pellets after each wash. The washed nuclear
409 pellet was then resuspended in 0.5 mL of nuclear resuspension buffer 1 [20 mM HEPES (pH
410 7.4), 10 mM $MgCl_2$, 250 mM sucrose, 0.2% Triton X-100, 1 mM PMSF, and 1× protease
411 inhibitor cocktail] and carefully laid on top of nuclear resuspension buffer 2 [20 mM HEPES (pH
412 7.4), 10 mM $MgCl_2$, 1.7 M sucrose, 0.2% Triton X-100, 1 mM PMSF, and 1× protease inhibitor
413 cocktail], then centrifuged at 4°C,16,000g for 45 min. The supernatant was carefully removed,
414 and the nuclear pellet was resuspended in 0.5 mL 1 X PBS (pH 7.5) buffer for crosslinking

415 reaction. The intactness and purity of the nuclear pellet was assessed by DAPI staining under a
416 Leica SP8 confocal microscope. Multiple samples were prepared in parallel, and final isolated
417 nuclei were pooled before crosslinking.

418 **Cross-linking reaction**

419 PhoX or tBU-PhoX cross-linker was freshly prepared at a stock concentration of 100 mM in
420 DMSO for each cross-linking experiment (supplementary Fig. S3). A working concentration of
421 2-5 mM cross-linker was added to buffer solution containing whole cell lysate, isolated
422 chloroplasts or isolated nuclei and incubated for 2 h at 4°C (dark condition for chloroplast cross-
423 linking) with end-to-end rotation on a hulamixer. The cross-linking reaction was quenched by
424 adding Tris pH 8.0 to a final concentration of 20 mM and subjected to protein extraction.

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

426 Protein extraction from the cross-linked samples, digestion to peptides, immobilized metal
427 affinity chromatography (IMAC) enrichment, and high pH reverse HPLC fractionation followed
428 the same procedure as described in (13). Multiple batches of PhoX-cross-linked peptides from
429 isolated intact chloroplasts were combined before IMAC enrichment. More information can be
430 found in supplementary Fig. S3. Multiple batches of samples were combined prior to high pH
431 reversed-phase HPLC fractionation (RPLC).

432 To generate selected proteomic libraries, flow-through peptides were retained in the
433 IMAC enrichment process for cell lysate and chloroplast samples. The chloroplast flow-through
434 peptides were further fractionated using the Pierce High pH Reverse-Phase Peptide Fractionation
435 Kit and collected in 8 fractions, measured by LC/MS/MS. The selected cell lysate database
436 contained 7835 abundant proteins, and the selected chloroplast database contained 9798 proteins.

437 **Identification of ubiquitinated proteins by enrichment of ubiquitinated peptides**

438 Proteins were extracted as described in Xu et al. (2017), followed by enrichment using the
439 PTMScan® Ubiquitin Remnant Motif (K-ε-GG) Kit (Cell Signaling Technology) according to
440 the manufacturer's instructions. The resulting peptides were then analyzed by LC-MS/MS.

441 **LC-MS/MS analysis**

442 Cross-linked peptides were analyzed by liquid chromatography-tandem mass spectrometry (LC-
443 MS/MS) using an Easy-nLC 1200 ultra-performance liquid chromatography (UPLC) system
444 (Thermo Fisher) coupled to an Orbitrap Eclipse Tribrid quadrupole Orbitrap mass spectrometer
445 (Thermo Fisher). Peptides were first loaded onto a C18 trap column (Acclaim PepMap 100 C18)
446 and then separated on an analytical C18 column (Aurora Series, 25 cm × 75 µm ID, Ion
447 Opticks). The flow rate was set to 300 nL/min and the separation was performed over a 120-
448 minute gradient. Peptides were eluted using a gradient from 3% to 28% solvent B over 106
449 minutes, followed by an increase from 28% to 44% over 15 minutes and a final wash at 90%
450 solvent B for 14 minutes. Solvent A consisted of 0.1% formic acid and solvent B consisted of
451 80% (v/v) acetonitrile with 0.1% (v/v) formic acid.

452 Data was acquired on an Orbitrap Eclipse mass spectrometer. Full MS scans were
453 acquired over a mass-to-charge (m/z) range of 375–1600 at a resolution of 120,000, with an
454 automatic gain control (AGC) target of 200,000, a maximum injection time of 50 ms, normalized
455 AGC target set to 50%, and RF lens set to 30%. The most intense multiply charged precursors
456 (charge states 2–8) were selected for fragmentation with a cycle time of 3 seconds. MS/MS scans
457 were acquired at a resolution of 15,000 with an AGC target of 5×10^4, maximum injection time
458 of 22 ms, isolation window of 1.4 m/z, and normalized AGC target set to 100%. Fragmentation
459 was performed using higher-energy collisional dissociation (HCD) with a normalized collision

460 energy (NCE) of 27. Dynamic exclusion was enabled for 30 seconds. For cross-linked peptide
461 detection, charge states 3 to 8 were used.
462 Real-time library search (RTLS) was used to reduce mono-link identifications. RTLS
463 included a low-resolution MS2 survey scan performed in the ion trap over an m/z range of 120–
464 500. The resulting MS2 peaks were matched against a custom spectral library containing
465 signature peak patterns characteristic of mono-links and cross-links (signature m/z values:
466 201.1231, 215.1387, 312.0632, and 377.1269, each with characteristic relative intensities for
467 mono- or cross-linked species). Only precursor ions whose survey MS2 spectra matched the
468 cross-link pattern using the "Similarity Search" function were selected for high-resolution MS2
469 acquisition. The RTLS method increases the overall cycle time compared to standard XL-MS
470 acquisition, resulting in a lower overall identification rate.
471 **Data search using pLINK2 and Protein Prospector**
472 Liner peptide searches were searched using Protein Prospector against Arabidopsis TAIR
473 database as described in (72). Precursor and fragment mass tolerance were set up as 10 and 20
474 ppm.
475 Crosslinked peptide searches were performed using pLink 2.3.11 with the following
476 parameters: crosslinker set to PhoX (209.972 Da); peptide length range: 6-60 amino acids;
477 enzyme: trypsin with up to 3 missed cleavages. Fixed modification: Cys carbamidomethylation;
478 variable modifications: protein N-terminal acetylation and Met oxidation. Precursor and
479 fragment mass tolerances were set to 5 ppm and 20 ppm, respectively.
480 MS2 spectra were searched against the TAIR10_pep_20101214, selective cell lysate
481 (7835) or selective chloroplast (9897) databases as described in Supplementary Fig.S3. Results
482 were filtered at 5% FDR for cell lysate and nuclear data sets and 1% FDR for chloroplast data

483 sets, all at the peptide spectrum matches (PSMs) level. FDRs for monolinks, loop-links, intra-
484 and inter-protein cross-links were calculated separately by pLINK2. To compare two acquisition
485 methods (XL-Standard [Standard] vs. Real-time library search [RTLS]), data were acquired on
486 the whole cell lysate XL samples, the search was against TAIR10_pep_20101214 database and
487 compared.

488 **Data compilation and reporting.**

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

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

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

501 **PANTHER protein class analysis and mapping intralinks to PDB and AlphaFold**
502 **structures using ComMap**

503 Arabidopsis gene identifiers (ATG accession numbers) were used for protein class analysis using
504 PANTHER™ Protein Class tool, which contains 210 classification terms (version 19.0, released
505 2024-06-20).

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

511 **Crosslinking data visualization**

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

517 *Xiview:* Protein–protein interactions (PPIs), including both inter- and intra-protein cross-
518 links, were visualized using XiView (https://xiview.org/) (Combe et al., 2024). PPIs were
519 categorized via the built-in STRING data loader (score cutoff 0.4). Cross-linked peptides
520 supporting the same lysine pair were consolidated into a unique K-K linkage, collapsing
521 redundancy (e.g. due to miscleavage events) and reporting each linkage only once.

522 *Mapping of crosslinks to high-resolution structures of proteins/complexes*. Intra- and
523 inter-crosslinks were superimposed on PDB or AlphaFold structures using the built-in function
524 of XiView, which also calculated the corresponding crosslink distances within the high-
525 resolution structures. Alignments of protein orthologs were performed and visualized using
526 Pymol (https://www.pymol.org/).

527 **Figure legend**

528 **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. Liljas A & Sanyal S (2018) The enigmatic ribosomal stalk. *Q Rev Biophys* 51:e12.
29. Komenda J, Sobotka R, & Nixon PJ (2024) The biogenesis and maintenance of PSII: Recent advances and current challenges. *Plant Cell* 36(10):3997-4013.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. Sheng X, *et al.* (2019) Structural insight into light harvesting for photosystem II in green algae. *Nat Plants* 5(12):1320-1330.
36. Martinez-Seidel F, Beine-Golovchuk O, Hsieh YC, & Kopka J (2020) Systematic Review of Plant Ribosome Heterogeneity and Specialization. *Front Plant Sci* 11:948.
37. Chaker-Margot M (2018) Assembly of the small ribosomal subunit in yeast: mechanism and regulation. *RNA* 24(7):881-891.
38. 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.
39. 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.
40. 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.
41. 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.
42. 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.

626 43. Chen J, *et al.* (2024) The Ptch/SPOUT1 methyltransferase deposits an m(3)U
627 modification on 28S rRNA for normal ribosomal function in flies and humans. *Sci Adv*
628 10(50):eadr1743.
629 44. Ghandadi M, Dobi A, & Malhotra SV (2024) A role for RIO kinases in the crosshair of
630 cancer research and therapy. *Biochim Biophys Acta Rev Cancer* 1879(3):189100.
631 45. Messling JE, *et al.* (2022) Targeting RIOK2 ATPase activity leads to decreased protein
632 synthesis and cell death in acute myeloid leukemia. *Blood* 139(2):245-255.
633 46. Kim DS, Challa S, Jones A, & Kraus WL (2020) PARPs and ADP-ribosylation in RNA
634 biology: from RNA expression and processing to protein translation and proteostasis.
635 *Genes Dev* 34(5-6):302-320.
636 47. Binder BM, *et al.* (2007) The Arabidopsis EIN3 binding F-Box proteins EBF1 and EBF2
637 have distinct but overlapping roles in ethylene signaling. *Plant Cell* 19(2):509-523.
638 48. Doelling JH, Yan N, Kurepa J, Walker J, & Vierstra RD (2001) The ubiquitin-specific
639 protease UBP14 is essential for early embryo development in Arabidopsis thaliana. *Plant*
640 *J* 27(5):393-405.
641 49. Van de Vosse DW, Wan Y, Wozniak RW, & Aitchison JD (2011) Role of the nuclear
642 envelope in genome organization and gene expression. *Wiley Interdiscip Rev Syst Biol*
643 *Med* 3(2):147-166.
644 50. Zhang S, Yu Q, Li Z, Zhao Y, & Sun Y (2024) Protein neddylation and its role in health
645 and diseases. *Signal Transduct Target Ther* 9(1):85.
646 51. Alemasova EE & Lavrik OI (2019) Poly(ADP-ribosyl)ation by PARP1: reaction
647 mechanism and regulatory proteins. *Nucleic Acids Res* 47(8):3811-3827.
648 52. Messner S, *et al.* (2010) PARP1 ADP-ribosylates lysine residues of the core histone
649 tails. *Nucleic Acids Res* 38(19):6350-6362.
650 53. Nosella ML, *et al.* (2024) Poly(ADP-ribosyl)ation enhances nucleosome dynamics and
651 organizes DNA damage repair components within biomolecular condensates. *Mol Cell*
652 84(3):429-446 e417.
653 54. Su Y, *et al.* (2017) Phosphorylation of Histone H2A at Serine 95: A Plant-Specific Mark
654 Involved in Flowering Time Regulation and H2A.Z Deposition. *Plant Cell* 29(9):2197-
655 2213.
656 55. Li F, *et al.* (2008) Identification of the wax ester synthase/acyl-coenzyme A:
657 diacylglycerol acyltransferase WSD1 required for stem wax ester biosynthesis in
658 Arabidopsis. *Plant Physiol* 148(1):97-107.
659 56. Vazquez A, Flammini A, Maritan A, & Vespignani A (2003) Global protein function
660 prediction from protein-protein interaction networks. *Nat Biotechnol* 21(6):697-700.
661 57. Giurgiu M, *et al.* (2019) CORUM: the comprehensive resource of mammalian protein
662 complexes-2019. *Nucleic Acids Res* 47(D1):D559-D563.
663 58. Luck K, *et al.* (2020) A reference map of the human binary protein interactome. *Nature*
664 580(7803):402-408.
665 59. Burke DF, *et al.* (2023) Towards a structurally resolved human protein interaction
666 network. *Nat Struct Mol Biol* 30(2):216-225.
667 60. Brandao MM, Dantas LL, & Silva-Filho MC (2009) AtPIN: Arabidopsis thaliana protein
668 interaction network. *BMC Bioinformatics* 10:454.
669 61. Herrera-Ubaldo H, *et al.* (2023) The protein-protein interaction landscape of transcription
670 factors during gynoecium development in Arabidopsis. *Mol Plant* 16(1):260-278.
671 62. Yang X, *et al.* (2020) PlaPPISite: a comprehensive resource for plant protein-protein
672 interaction sites. *BMC Plant Biol* 20(1):61.
673 63. Simsek D, *et al.* (2017) The Mammalian Ribo-interactome Reveals Ribosome Functional
674 Diversity and Heterogeneity. *Cell* 169(6):1051-1065 e1018.
675 64. Peng Y, Markov Y, Goncearenco A, Landsman D, & Panchenko AR (2021) Human
676 Histone Interaction Networks: An Old Concept, New Trends. *J Mol Biol* 433(6):166684.

677 65. Sabari BR, Zhang D, Allis CD, & Zhao Y (2017) Metabolic regulation of gene expression
678 through histone acylations. *Nat Rev Mol Cell Biol* 18(2):90-101.
679 66. Goudarzi A (2019) The recent insights into the function of ACAT1: A possible anti-cancer
680 therapeutic target. *Life Sci* 232:116592.
681 67. Liu Y*, et al.* (2022) Mechanisms and inhibition of Porcupine-mediated Wnt acylation.
682 *Nature* 607(7920):816-822.
683 68. Nile AH & Hannoush RN (2016) Fatty acylation of Wnt proteins. *Nat Chem Biol* 12(2):60-
684 69.
685 69. Xue B & Rhee SY (2023) Status of genome function annotation in model organisms and
686 crops. *Plant Direct* 7(7):e499.
687 70. Casazza AP, Tarantino D, & Soave C (2001) Preparation and functional characterization
688 of thylakoids from Arabidopsis thaliana. *Photosynth Res* 68(2):175-180.
689 71. Shao Z*, et al.* (2024) Nuclear pyruvate dehydrogenase complex regulates histone
690 acetylation and transcriptional regulation in the ethylene response. *Sci Adv*
691 10(30):eado2825.
692 72. Shrestha R*, et al.* (2022) 15N Metabolic Labeling Quantification Workflow in Arabidopsis
693 Using Protein Prospector. *Front. Plant Sci.* 13:832562.
694

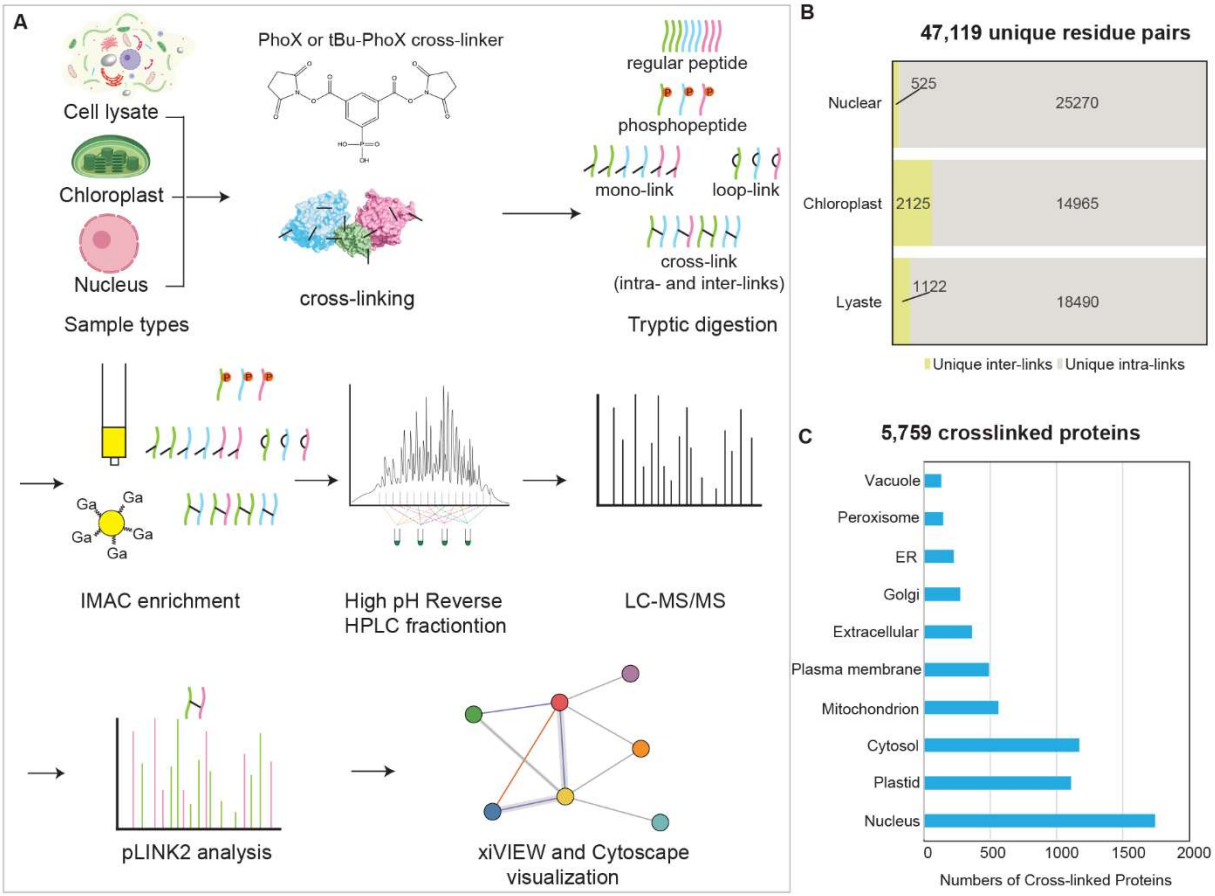


Figure 1. Efficient cross-linking workflow for complex protein samples identifies numerous cross-links in Arabidopsis.

(**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 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.

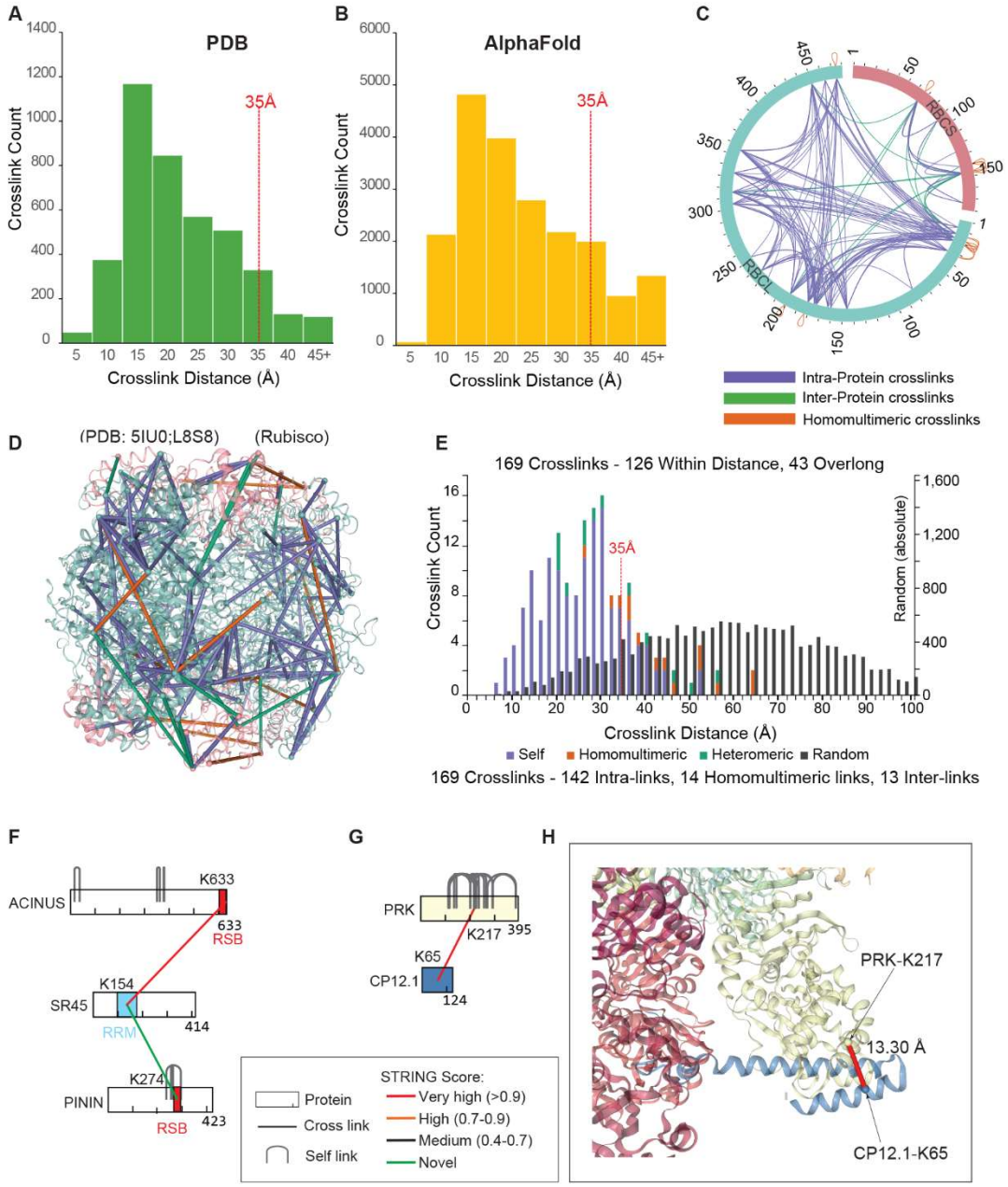


Figure 2. Validation of intra- and inter-protein cross-links by structural mapping.
(**A**-**B**) Histogram of Euclidean Cα-Cα distances of intra-protein cross-links mapped to Arabidopsis structures from PDB (A) or AlphaFold (B), showing 94% and 89% within the 35Å constraint, respectively.
(**C**-**E**) 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 SR45 RRM and RSB domains.
(**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.

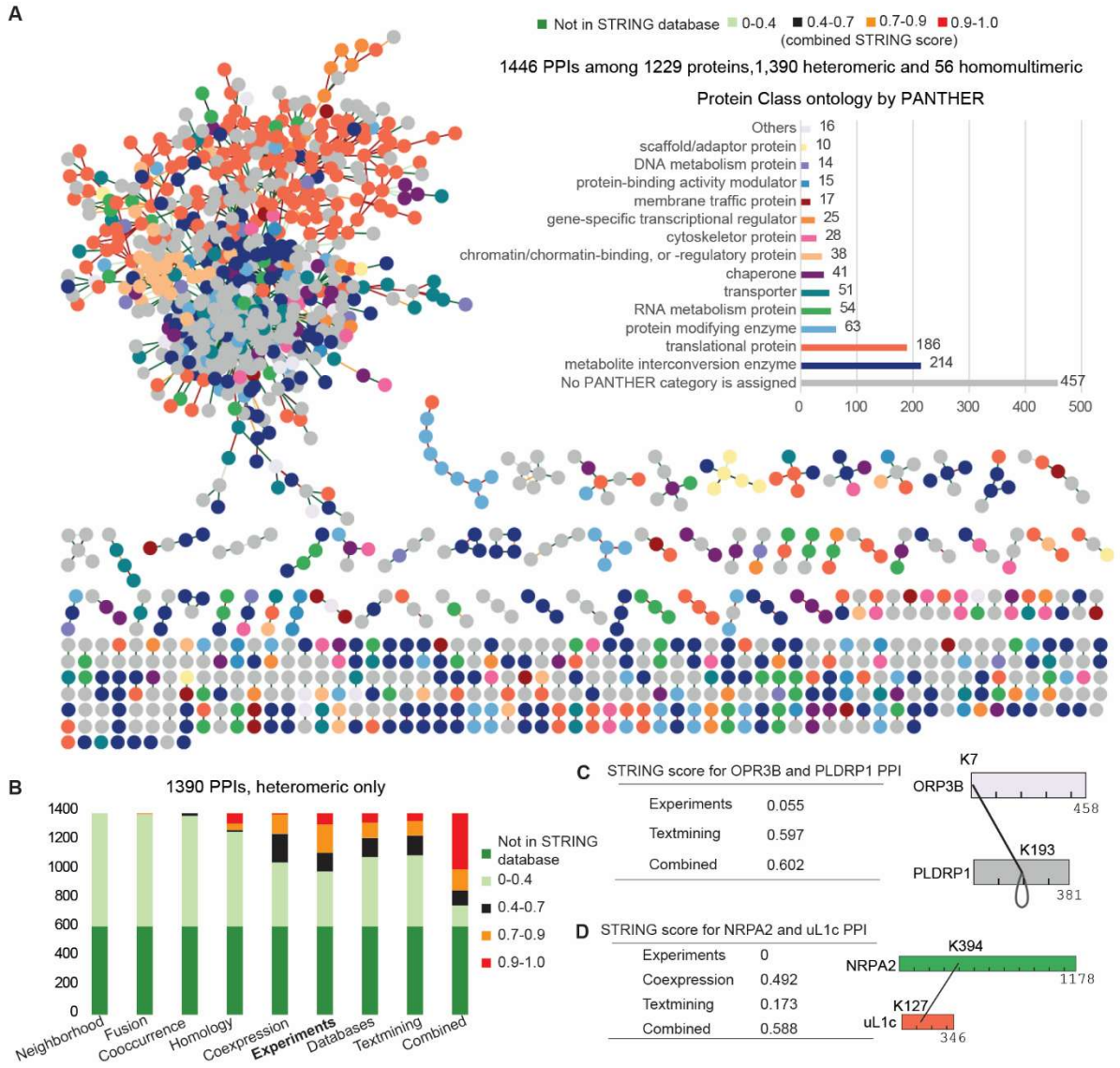


Figure 3. XL-MS identifies direct protein-protein interactions and reveals known and novel complexes.

(**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.

(**B**) The breakdown of the STRING analysis shows that more than 71% of the detected heteromeric PPIs are either of low confidence or have no 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).

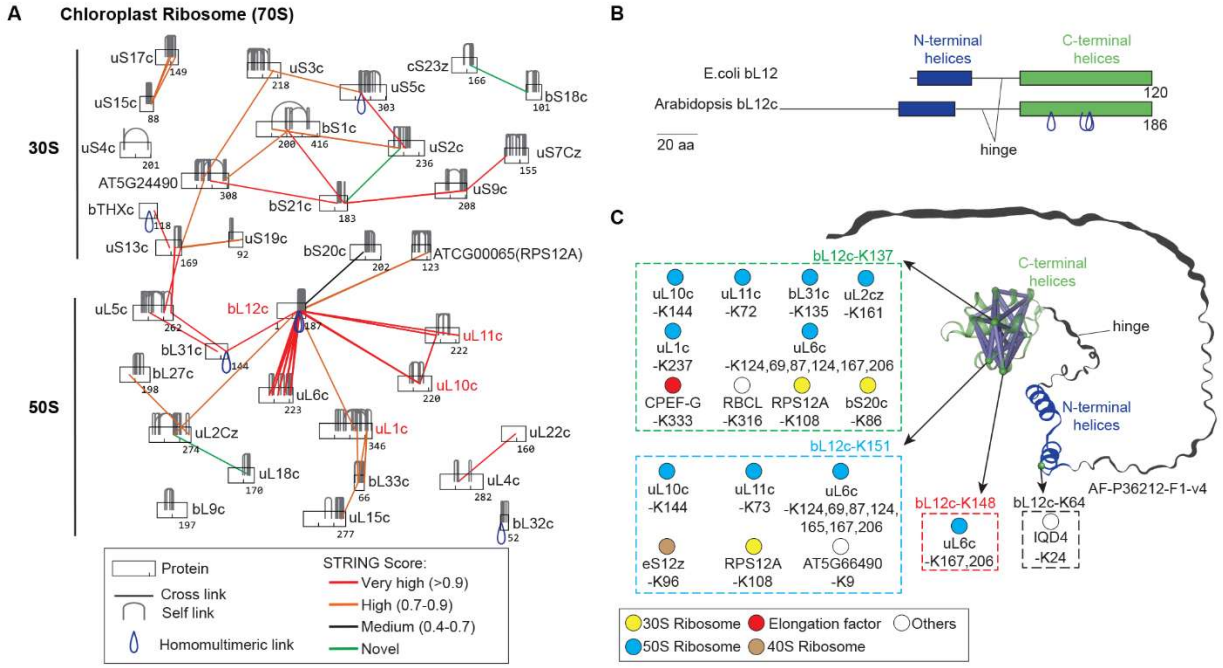


Figure 4. XL-MS reveals unique topology of chloroplast 70S ribosome stalk regions.

(**A**) XL-MS identifies flexible stalk proteins (bL12c, uL1c, uL10c, uL11c), with cross-links color-coded by STRING confidence. Homodimer interfaces are found at the C-terminus of bL12c.

(**B**) Domain comparison between Arabidopsis bL12c and E. coli bL12.

(**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.

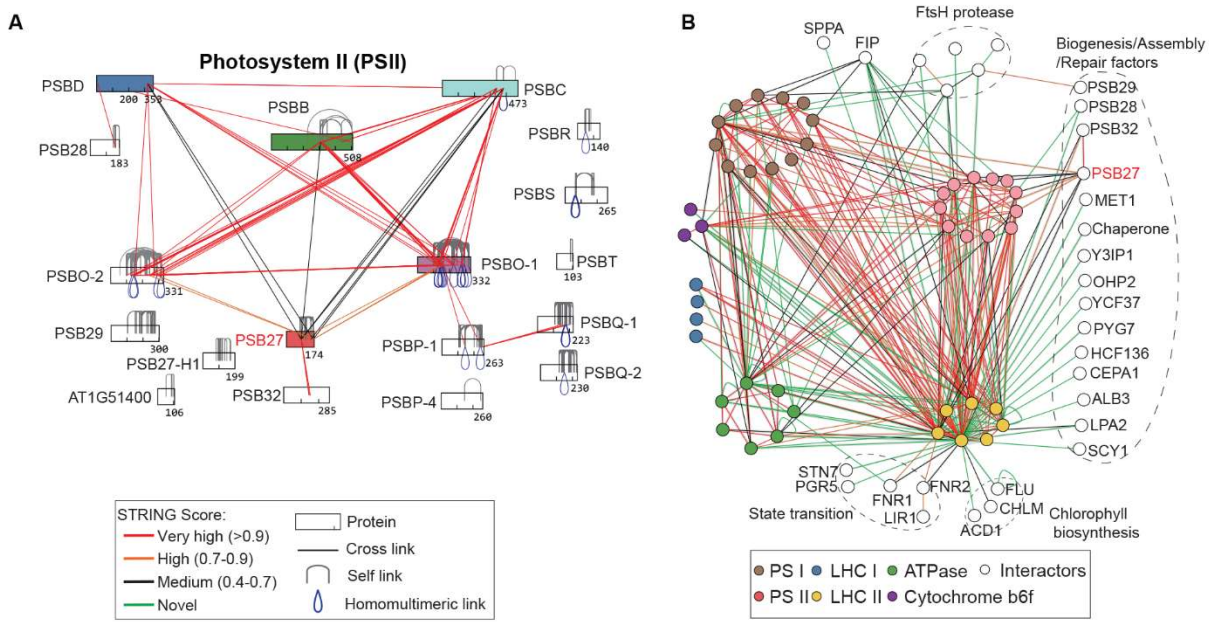
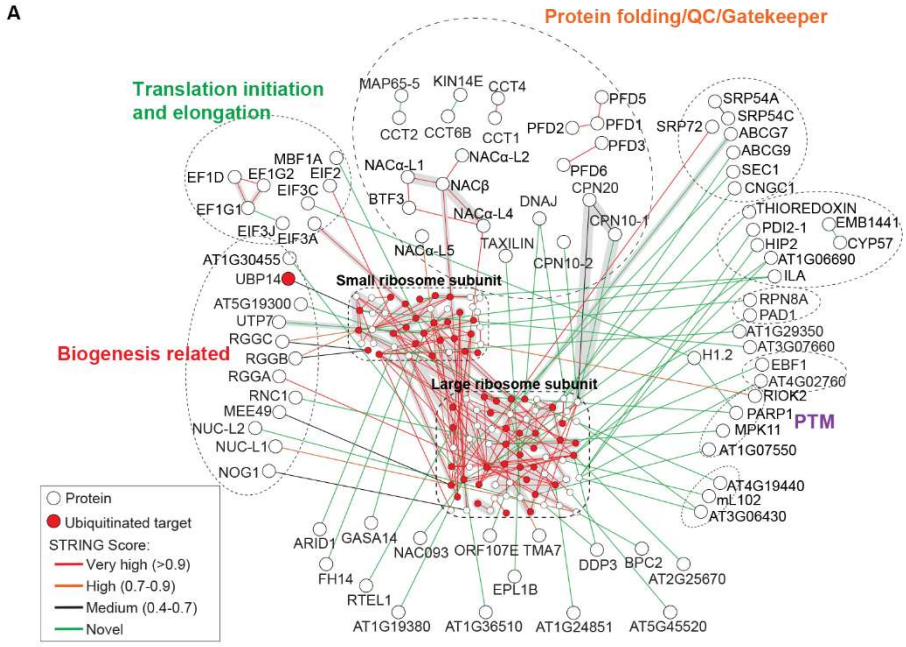


Figure 5. XL-MS reveals dynamic connectivity and unique features in photosystem complexes.

(**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.



B

Interactor	Lys(K)	Ribosome	Lys(K)
MEE49	22	uL29y	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
NAC α-like 5	92	uL29y	43
NACβ	18	eL31z	58
NACβ	18	eL31z	61
NAC α-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
PDIL21	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.

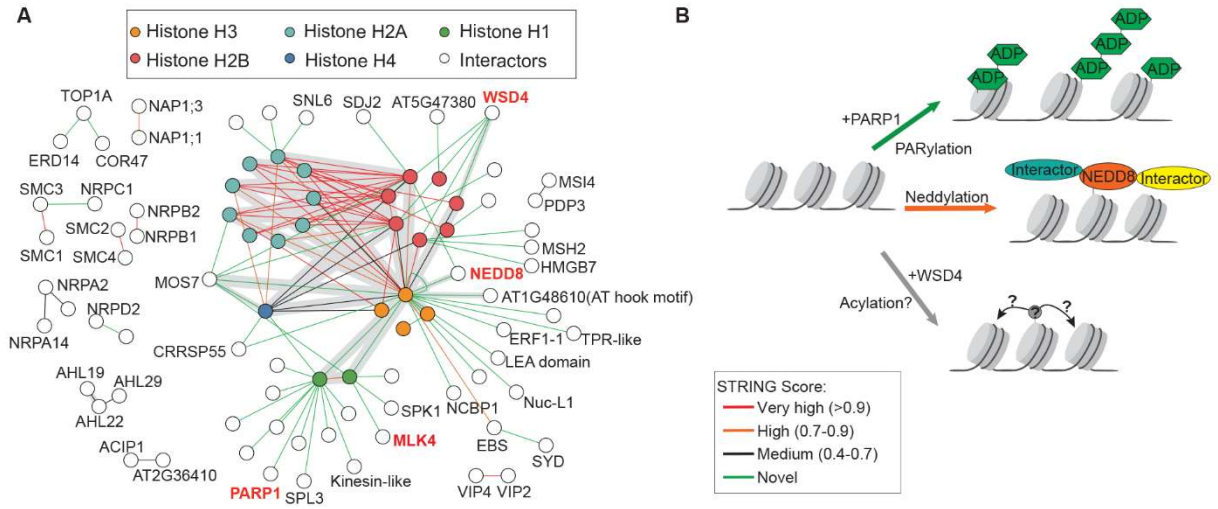
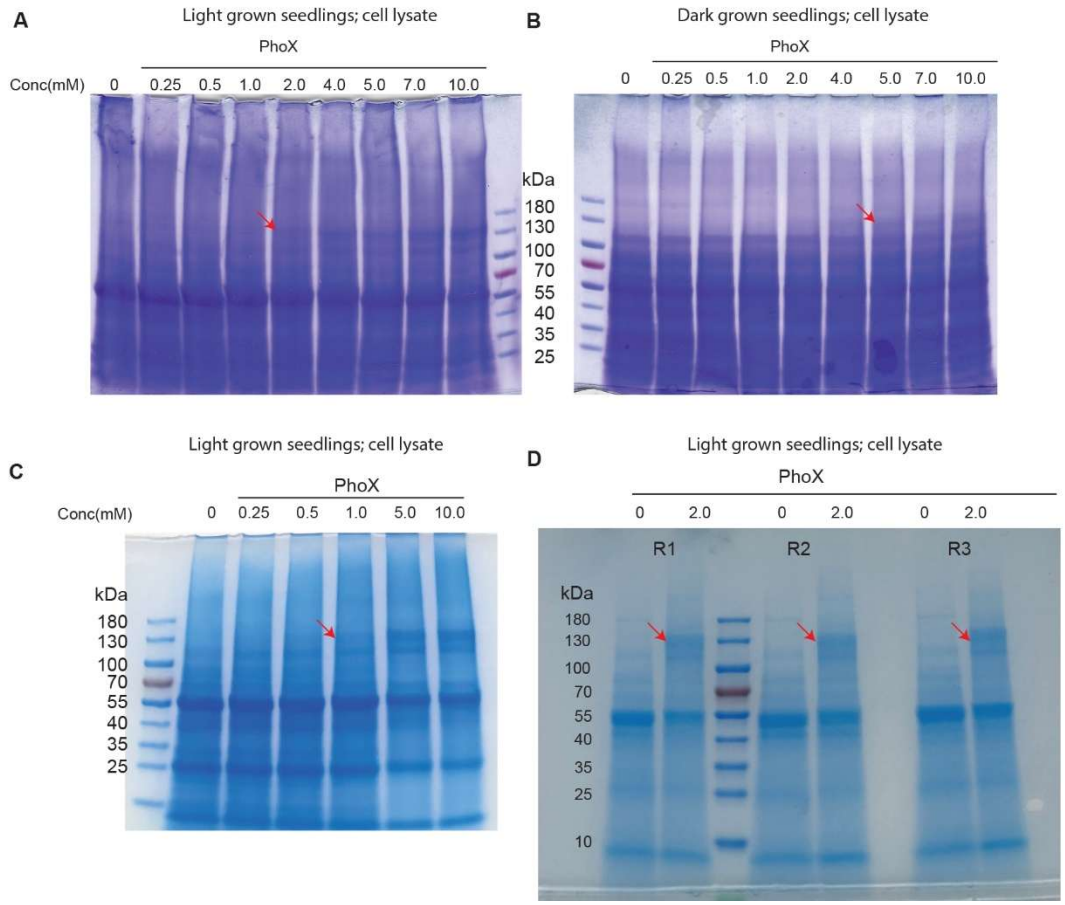


Figure 7. Histone interactome reveals conserved and novel chromatin-associated interactions.

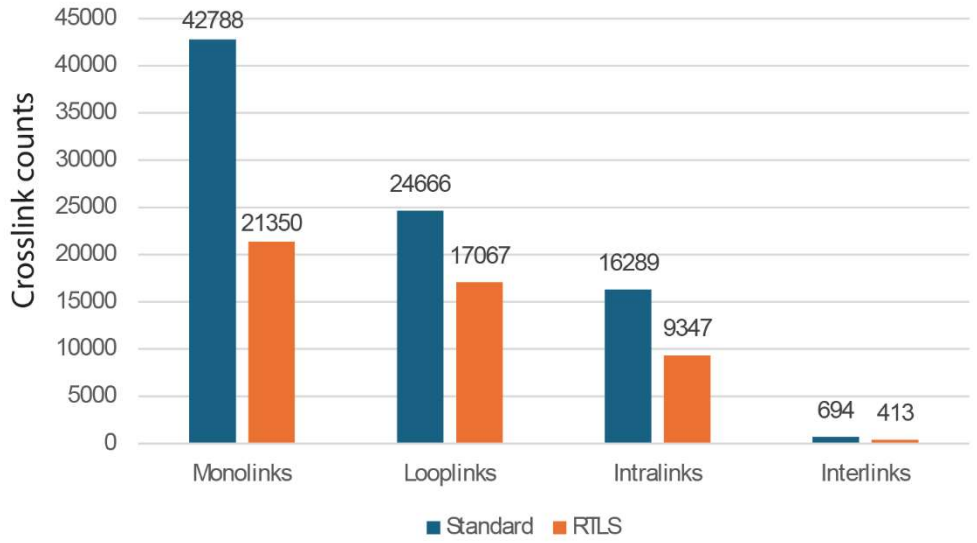
(**A**) XL-MS reveals extensive interactions between histones and both known and previously uncharacterized proteins, including components of chromatin-associated complexes. Enzymatic regulators are highlighted in red.

(**B**) Schematic illustration of histone modifications and polynucleosome organization affected by PARylation and neddylation. A potential role is proposed for an O-acyltransferase - previously associated with metabolic processes - that cross-links with four members of the histone H2B family, suggesting its involvement in histone modification based on XL-MS interaction data.



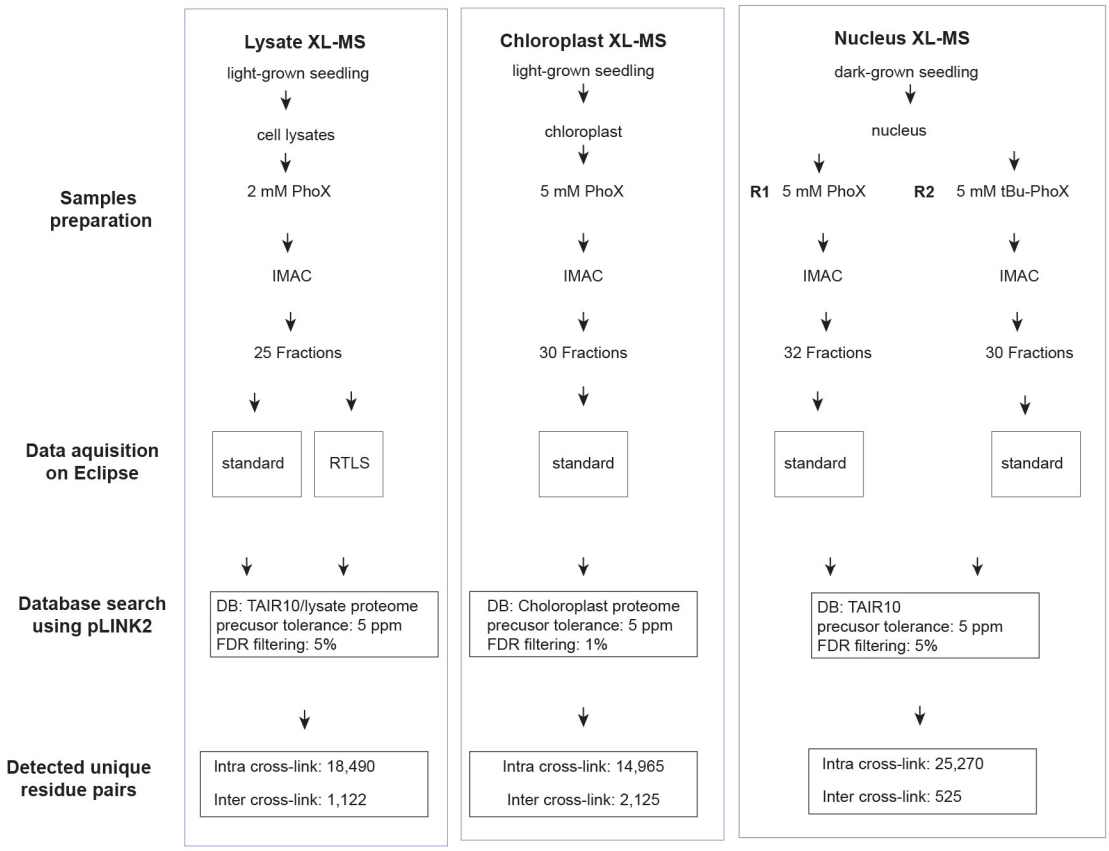
Supplementary Figure S1: Optimization of the PhoX cross-linker concentration for XL-MS.

(**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).



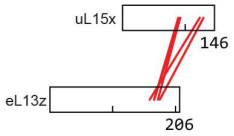
Supplementary Figure S2: **Standard XL acquisition outperforms Real-Time Library Search (RTLS) acquisition.**

Data were acquired on the Orbitrap Eclipse mass spectrometer using PhoX-enriched cross-linked peptides from cell lysate samples and searched against the entire Arabidopsis library. While RTLS acquisition significantly reduced the identification of monolinks (where one NHS ester group is cross-linked to a peptide while the other is hydrolyzed) and loop links (cross-links within a single peptide), it also significantly reduced the detection of inter-peptide cross-links (pairs of peptides covalently linked by the cross-linker), including both intra- and inter-protein cross-links (intralinks and interlinks).



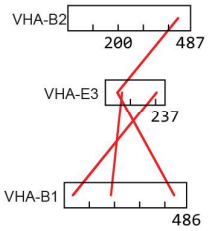
Supplementary Fig. S3: Summary of XL-MS experiments.
Key details of sample preparation, data acquisition using the Orbitrap Eclipse mass spectrometer, and database searching using pLINK2 are provided. A summary of the unique cross-linked peptide residue pairs identified across experiments is also included.

A



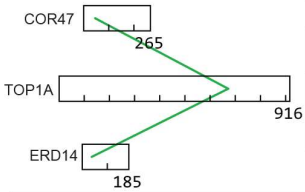
Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
uL15x	K92	ST**K**DNVPLIDVTQHGFFK	eL13z	K166	EKPTMELV**K**LTSEMK	4018.957986
uL15x	K92	AKST**K**DNVPLIDVTQHGFFK	eL13z	K166	EKPTMELV**K**LTSEMK	4218.090052
uL15x	K92	ST**K**DNVPLIDVTQHGFFK	eL13z	K172	LTSEM**K**SFK	3325.58491
uL15x	K92	ST**K**DNVPLIDVTQHGFFK	eL13z	K172	LTSEM**K**SFKAFDK	3786.812323
uL15x	K89	A**K**STKDNVPLIDVTQHGFFK	eL13z	K166	EKPTMELV**K**LTSEMK	4218.090052
uL15x	K123	GHLPENKPFVV**K**AK	eL13z	K159	E**K**PTMELVK	2847.451335
uL15x	K129	LIS**K**TAEKK	eL13z	K175	SF**K**AFDKIR	2338.220553

B



Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
VHA-E3	K67	K**K**IDYSMQLNASR	VHA-B1	K186	QAGLV**K**R	2534.258401
VHA-E3	K204	DG**K**IVCENTLDAR	VHA-B1	K34	TVSGVAGPLVILD**K**VKGPK	3577.833392
VHA-E3	K204	DG**K**IVCENTLDAR	VHA-B1	K36	TVSGVAGPLVILDKV**K**GPK	3577.833392
VHA-E3	K48	LQLVEAE**K**KK	VHA-B1	K437	**K**FVMQGAYDTR	2710.330892
VHA-E3	K49	LQLVEAEK**K**K	VHA-B1	K437	**K**FVMQGAYDTR	2710.330892
VHA-E3	K49	LQLVEAEK**K**K	VHA-B2	K438	**K**FVAQGAYDTR	2650.327524
VHA-E3	K48	LQLVEAE**K**KK	VHA-B2	K438	**K**FVAQGAYDTR	2650.327524

C



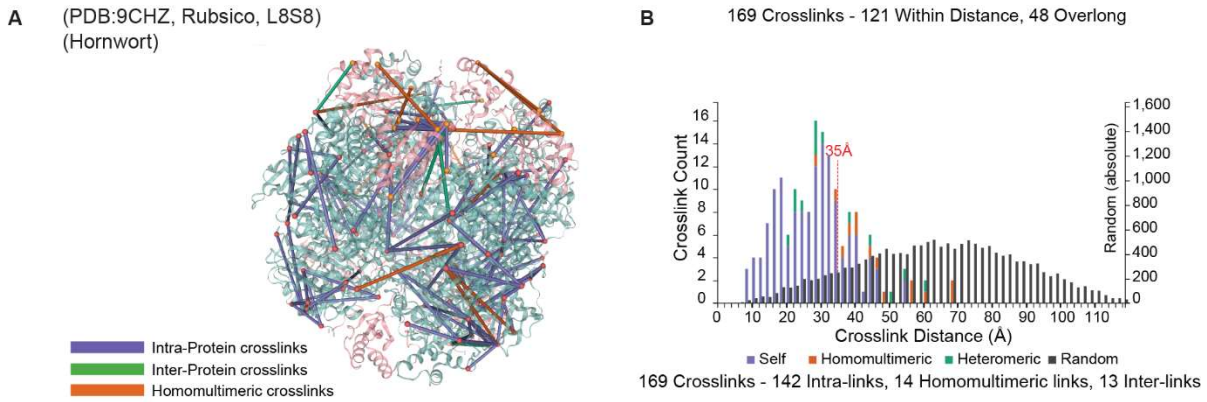
Protein1	Linkage site	PepSeq1	Protein2	Linkage site	PepSeq2	Peptide_Mass
TOP1A	K672	I**K**FDFLGK	EDR14	K38	**K**DETKPEETPIASEFEQK	3282.549239
TOP1A	K672	I**K**FDFLGK	COR47	K46	EEEV**K**PQETTTLESEFDHK	3452.578539

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 were mapped to conserved regions of homologous proteins; for example, the ATPase subunit VHA-E3 (K49) was 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.

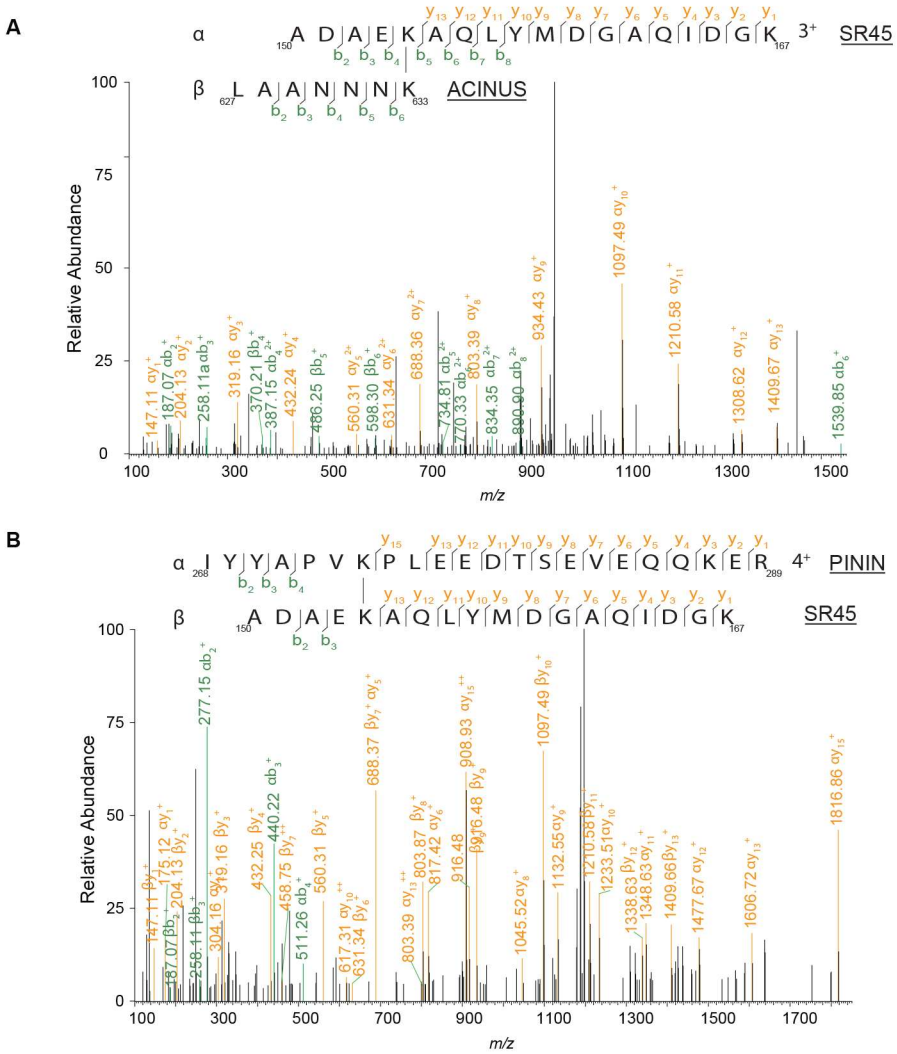


Supplementary Figure S5. Mapping of Rubisco cross-links on a high-resolution orthogonal structure of hornwort Rubisco.

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.

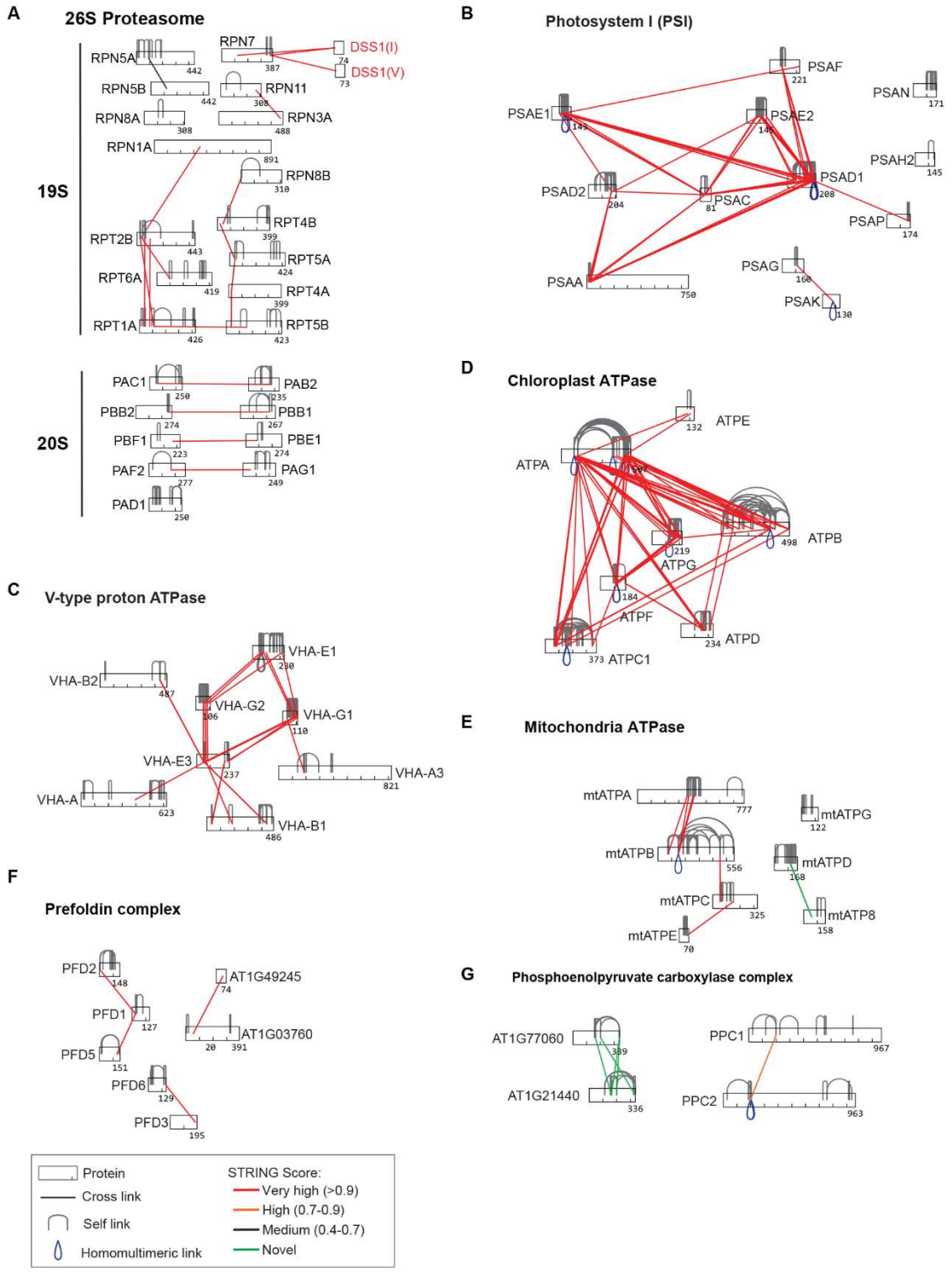


Supplementary Figure S6. High-confidence identifications of cross-linked peptides between ACINUS–SR45 and PININ–SR45.

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.

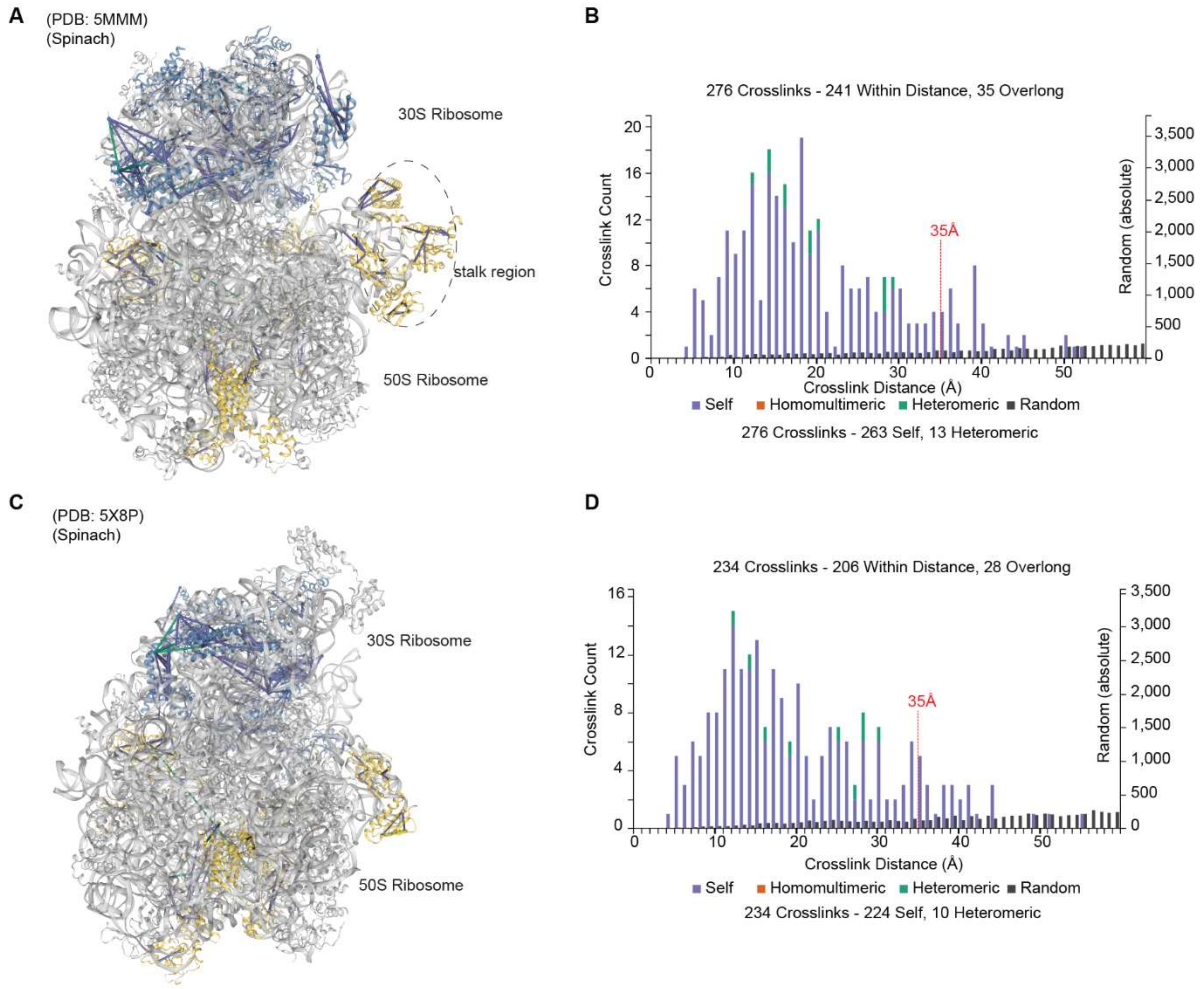
(**A**) Spectrum of the cross-linked peptide between ACINUS-K633 and SR45-K154.

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

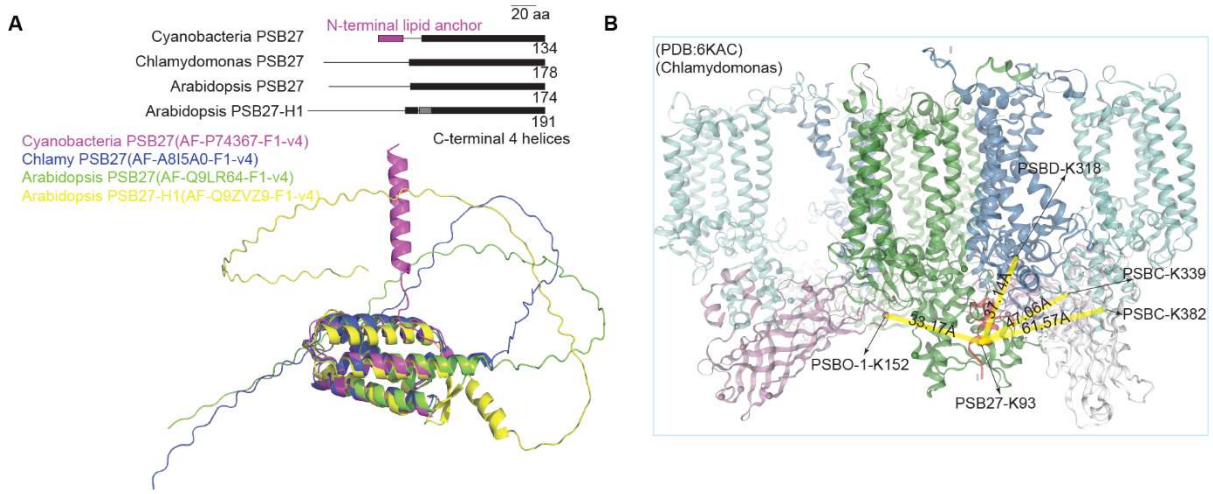


Supplementary Fig. S7. Topological mapping of multiple protein complexes by XL-MS. Proteins are shown as rectangles scaled to their respective lengths.
Cross-links are color-coded based on STRING score confidence thresholds, highlighting both known and novel interactions.

(**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.



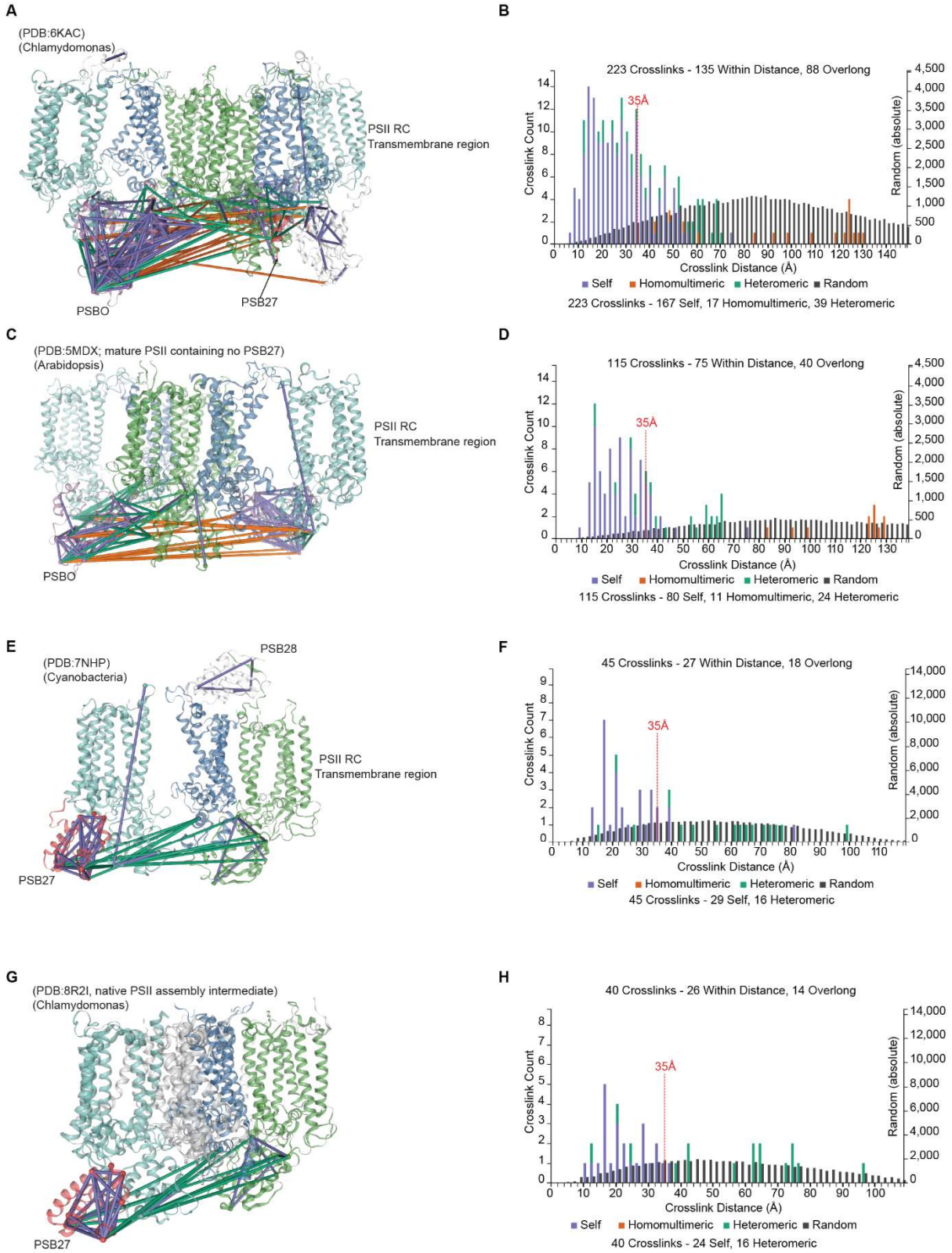
Supplementary Figure S8. Mapping of cross-links in the 70S ribosome complex onto high-resolution structures of spinach 70S ribosomes (PDB: 5MMM and 5X8P).
Cross-links are color-coded by type: purple for intralinks, green for heteromeric inter-protein cross-links, and orange for homomultimeric inter-protein 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.

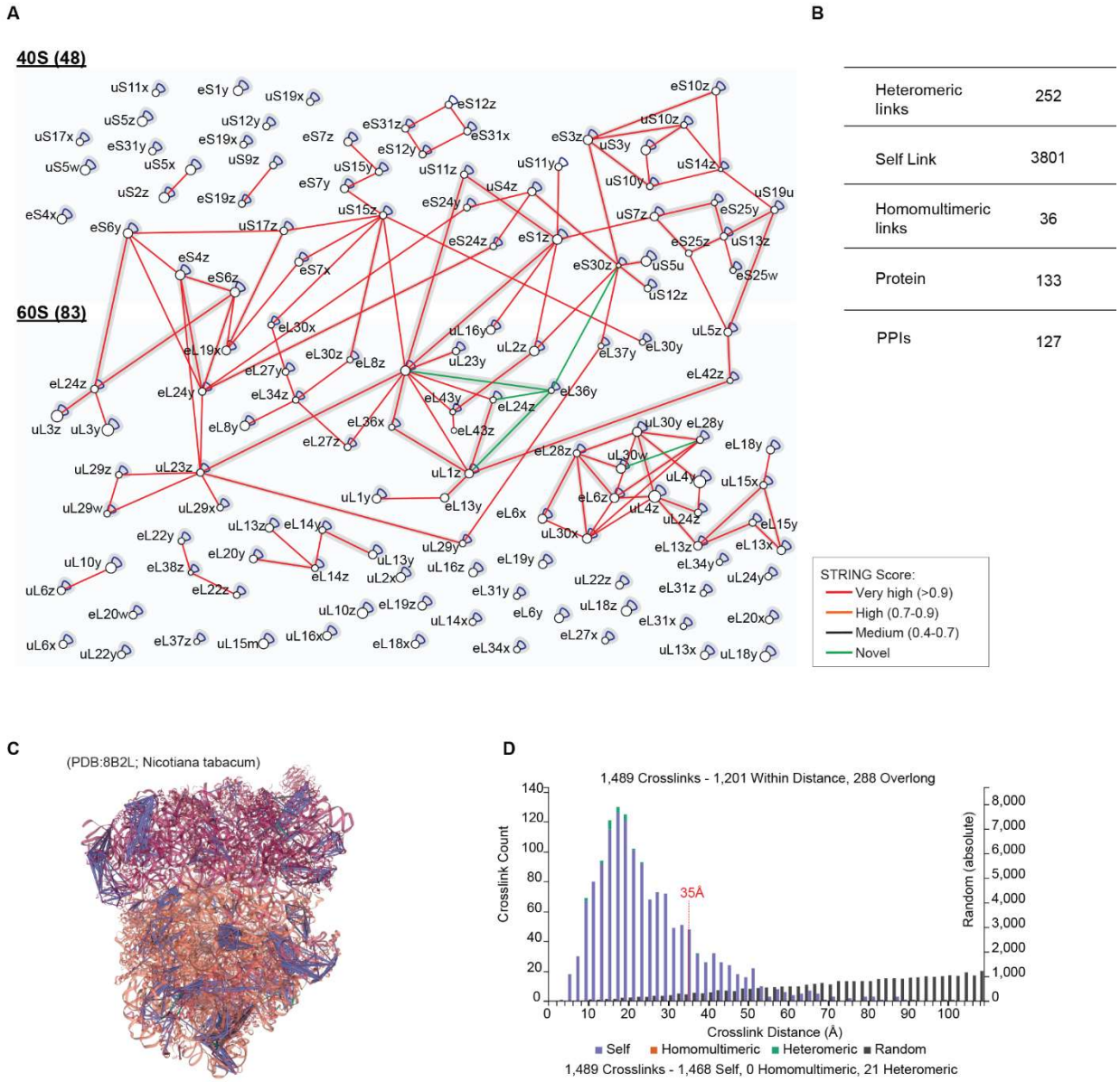
(**A**) Superimposed AlphaFold structures of PSB27 homologs reveal unique features of Arabidopsis PSB27. While it contains a conserved C-terminal four-helix bundle, Arabidopsis PSB27 lacks a lipid anchor and shows divergence in sequence and surface charge.

(**B**) Cross-links between Arabidopsis PSB27 and PSII core/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)** Structure mapping of PSII cross-links onto the cyanobacterial PSII assembly intermediate (PDB: 7NHP), which presents a PSII assembly intermediate containing the PSII reaction center, Psb27 and Psb28.
(**G**-**H**) Structure mapping of PSII cross-links onto the Chlamydomonas PSII assembly intermediate (PDB: 8R2I), which represents a native PSII assembly intermediate containing the PSII reaction center and Psb27.



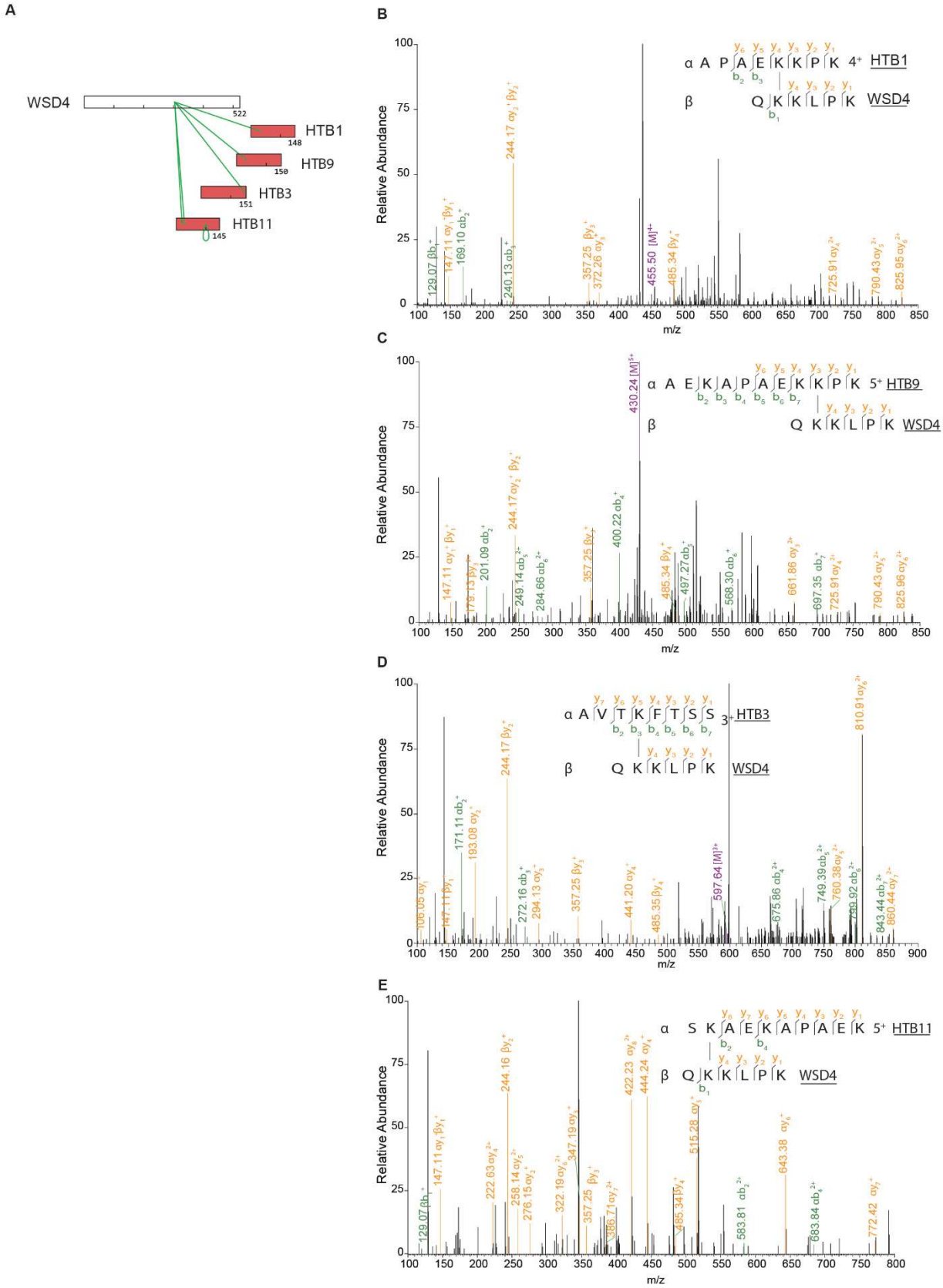
Supplementary Figure S11. Topology of the 80S ribosomal complex.

(**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).


A
WSD4
522
HTB1
148
HTB9
150
HTB3
151
HTB11
145
B
α A P A E K K P K 4+ HTB1
β Q K K L P K WSD4
Relative Abundance
m/z
C
α A E K A P A E K K P K 5+ HTB9
β Q K K L P K WSD4
Relative Abundance
m/z
D
α A V T K F T S S 3+ HTB3
β Q K K L P K WSD4
Relative Abundance
m/z
E
α S K A E K A P A E K 5+ HTB11
β Q K K L P K WSD4
Relative Abundance
m/z

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.