

# A composite filter for low FDR of protein-protein interactions detected by in vivo cross-linking

Luitzen de Jong, Winfried Roseboom and Gertjan Kramer

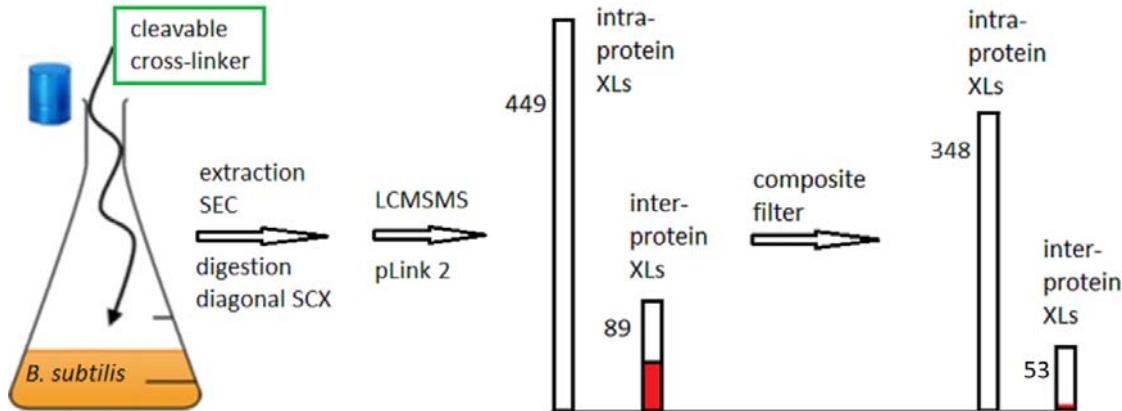
Mass spectrometry of Biomolecules, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands

## Abstract

In vivo chemical cross-linking combined with LCMSMS of digested extracts (in vivo CX-MS) can reveal stable and dynamic protein-protein interactions at a proteome wide-scale and at the peptide level. In vivo CX-MS requires a membrane permeable and cleavable cross-linker that enables isolation of target peptides and a fast and sensitive search engine to identify the linked peptides. Here we explore the use of the search engine pLink 2 for analysis of a previously obtained LCMSMS dataset from exponentially growing *Bacillus subtilis* treated in culture with the cross-linker bis(succinimidyl)-3-azidomethyl-glutarate (BAMG). Cross-linked peptide pairs were identified by pLink 2 in very short time at an overall FDR of < 5%. To also obtain a FDR < 5% for inter-protein cross-linked peptide pairs additional thresholds values were applied for matched fragment intensity and for the numbers of unambiguous y and b ions to be assigned for both composite peptides. Threshold values were based on a set of decoy sequences from yeast and human sequence databases. Also the mass- and charge-dependent retention times of target peptides purified by diagonal strong cation exchange chromatography were used as a criterion to distinguish true from false positives. After this filtering, pLink 2 identified more than 80% of previously reported protein-protein interactions. In addition the use of pLink 2 revealed interesting new inter-protein cross-linked peptide pairs, among others showing interactions between the global transcriptional repressor AbrB and elongation factor Tu and between the essential protein YlaN of unknown function and the ferric uptake repressor Fur.

## Keywords

Cross-linking mass spectrometry, cleavable cross-linker, protein-protein interactions, YlaN, Fur



## Highlights

- Improved protocol for identification of PPIs at low FDR by *in vivo* cross-linking with BAMG
- The use of all intra-protein cross-linked peptide pairs as true positives
- The cytosolic aminopeptidase (AMPA\_BACSU) interacts with the 50S ribosomal protein L17
- The transition state regulator AbrB interacts with elongation factor Tu
- The essential protein YlaN of unknown function interacts with the iron uptake repressor Fur

## Significance

Important for reliable identification of PPIs by chemical cross-linking *in vivo* is a low FDR of non-redundant inter-protein peptide pairs. Here we describe how to recognize the presence of spurious interactions in a dataset of cross-linked peptide pairs enriched by 2D strong cation exchange chromatography and identified by LCMSMS by taking into account chromatographic behavior of cross-linked peptide pairs and protein abundance of corresponding peptides. Based on these criteria we assessed that the FDR of the fraction of non-redundant inter-protein cross-linked peptide pairs was approx. 20-25% by interrogating an entire species specific database at an overall FDR of 5% or 0.1% with a search engine that otherwise scores best in sensitivity among other search engines. We have defined a composite filter to decrease this high FDR of inter-protein cross-linked peptide pairs to only about 2%.

1        1. Introduction

2        Specific protein-protein interactions (PPIs) are crucial for the regulation of biochemical processes.

3        Large scale approaches like affinity purification combined with mass spectrometry[1], proximity-

4        dependent labeling[2], the use of co-elution profiles[3] and the yeast two-hybrid system[4] can

5        reveal which proteins interact with each other under different experimental and physiological

6        conditions. In vivo chemical cross-linking followed by mass spectrometry and database searching to

7        identify cross-linked peptide pairs (CX-MS) has the potency to reveal cellular protein-protein

8        interactions at a proteome-wide scale in a single experiment and in a short time[5–7]. Besides stable

9        protein complexes also dynamic assemblies that may dissociate upon cell extraction can be trapped

10       by cross-linking in vivo, implying that PPIs may be encountered that have thus far escaped detection

11       by in vitro approaches. Most importantly, the spatial arrangement of proteins in a complex can be

12       assessed by CX-MS by virtue of the identity of linked amino acid residues in the protein sequences as

13       determined by MS and the known length of the spacer of the cross-link. The presence of cross-links

14       formed between two proteins in a complex facilitates modeling of the overall structure if the

15       coordinates of the atoms in the two interacting proteins are known. Such models may lead to

16       hypotheses about the functional significance of the interaction. Thus, CX-MS in vivo is a useful

17       approach in structural and systems biology.

18

19       In vivo CX-MS requires a membrane permeable cross-linker designed to facilitate mass spectrometric

20       identification of cross-linked sites. Peptide identification is achieved by liquid chromatography

21       coupled to mass spectrometry (LCMSMS). Mass and charge of the peptides are determined in the

22       MS1 stage. In the MS2 stage peptides are selected for fragmentation by cleavages of the peptide

23       bonds induced by collision with gas molecules (collision-induced dissociation, CID) in a data

24       dependent way, i.e., selection is dependent on signal intensity, mass and charge, and whether a

25       particular precursor ion has been selected before in a given time window. This results in an MS2

26 fragment spectrum that is characteristic for the peptide sequence. With an MS1MS2 dataset,  
27 unmodified peptides, or peptides with a defined modification, can be identified by searching in a  
28 sequence database. However, for cross-linked peptides this approach is challenging for two main  
29 reasons. In the first place the MS1 data give no information about the masses of the composite  
30 peptides, which hampers their identification by database searching. The lack of peptide mass  
31 knowledge can be circumvented using a gas phase cleavable cross-linker yielding fragment ions from  
32 which the masses of the composite peptides can be deduced. In the second place cross-linked  
33 peptides are present in sub-stoichiometric amounts as compared with unmodified peptides, so that  
34 selection in the mass spectrometer in the process of data dependent acquisition of MS1MS2 spectra  
35 is hampered. This problem requires enrichment of the rare cross-linked species out of the bulk of  
36 unmodified peptides.

37 To meet these requirements and challenges we previously synthesized bis(succinimidyl)-3-  
38 azidomethyl-glutarate (BAMG)[8], a membrane permeable cross-linker[9]. The short spacer of BAMG  
39 gives relatively high resolution cross-link maps, while the azido group in the spacer can be modified  
40 in different ways. This reactive versatility enables three different cross-link analysis strategies[10–  
41 14]. For CX-MS *in vivo* we have made use of a TCEP-induced reduction of the azido group to enrich  
42 cross-linked peptide pairs obtained from growing *Bacillus subtilis* cells treated in culture with  
43 BAMG[9]. Besides isolation of target peptides, reduction of the azido group also facilitates mass  
44 spectrometric identification of the linked peptide pair. This is due to the fact that the two cross-link  
45 amide bonds of a peptide pair are cleavable in the gas phase by CID [11]. The cleavage of the cross-  
46 link amide bond can occur along with a peptide bond cleavage. The principle was demonstrated  
47 with protein complexes from *Bacillus subtilis* in the mass range 400 kDa to 1-2 MDa obtained by size  
48 exclusion chromatography (SEC) after *in vivo* cross-linking. We identified several cross-linked  
49 peptides revealing transient and stable PPIs of high biological significance and low FDR by searching  
50 MS1MS2 data from the entire *B. subtilis* sequence database[9]. This analysis was supported by the

51 use of in-house developed scripts called Raeng and YeunYan in combination with the search engine  
52 MASCOT.

53 This procedure, further called the Raeng/Mascot approach, is relatively laborious, making it  
54 unattractive for general use. This holds in particular in cases where large datasets are to be expected  
55 by the use of state of the art HPLC coupled to sensitive and fast mass spectrometers like the recently  
56 introduced trapped ion mobility time-of flight mass spectrometer (timsTOF Pro) enabling online  
57 parallel accumulation-serial fragmentation (PASEF)[15]. Here we explore the use of the recently  
58 launched second generation search engine, pLink 2[16]. In comparison with many other search  
59 engines developed for cross-link analysis, pLink performs best with respect to number of identified  
60 cross-links with good FDR at cross-link spectrum matches (CSM) level [16,17]. Moreover it is  
61 extremely fast, and is also suitable for analysis using cleavable cross-linkers like BAMG[16]. With an  
62 existing LCMSMS dataset we observed a large overlap with the results obtained with the  
63 Raeng/Mascot approach, but also found evidence that the FDR for inter-protein peptide pairs is  
64 higher for inter-protein peptide pairs than for intra-protein peptide pairs. Here we describe an  
65 approach to further lower the FDR for the latter type of cross-links that reliably led to interesting  
66 new protein-protein interactions.

67

68 **2. Materials and methods.**

69 *2.1. Acquisition of the LCMSMS dataset to explore the use of pLink 2.*

70 The LCMSMS dataset was acquired as described previously in detail [9]. In short, exponentially  
71 growing *B. subtilis* (strain 168) was treated in culture with 2 mM BAMG at 37° C for 5 min. The  
72 soluble extract, obtained after harvest and sonication of cross-linker-treated cells, was subjected to  
73 SEC. A tryptic digest of protein complexes in the size range 400 kDa to 1-2 MDa was selected for  
74 further analysis. First the peptide mixture was subjected to diagonal strong cation exchange (SCX)

75 chromatography to sequester the cross-linked peptides from the bulk of unmodified peptides.  
76 Fractions enriched in cross-linked peptide pairs were analyzed by LCMSMS using an Eksigent Expert  
77 nanoLC 425 system connected to the Nano spray source of a TripleTOF 5600+ mass spectrometer.  
78 Data processing was as described [9]. For analysis by pLink 2 the original mgf files were combined  
79 into 13 mgf files. An overview of the origin and size of the combined mgf files along with the mass  
80 accuracy used for pLink 2 searches is shown in Supplementary Table S1. The original mgf files are  
81 available via ProteomeXchange with identifier PXD006287.

82 *2.2. Identification of cross-linked peptides*

83 BAMG-cross-linked peptides become cleavable in the gas phase when the azido group in the spacer  
84 of the cross-link has been reduced to an amino group. For cross-link identification we used pLink 2 as  
85 a search engine operating in the stepped-HCD mode for MS-cleavable cross-linkers [16]. This mode  
86 has been developed for the cross-linkers DSBU[18] and DSSO[19]. In the stepped-HCD mode  
87 cleavage of the DSSO- or DSBU-cross-link and generation of peptide fragments occurs with two  
88 different CID energies after which the data obtained in the two steps are combined in one file entry.  
89 Cleavage of the BAMG-cross-link amide bonds and cleavage of the peptide bonds resulting in y and b  
90 ions occurs in one step under defined CID conditions (Fig 1). Therefore both DSSO-, DSBU- and  
91 BAMG-cross-linked  $\alpha$ - $\beta$  peptide pairs yield similar MSMS spectra containing the signals from intact  
92 cleaved peptides along with the signals from peptide bond cleavages. Like DSSO and DSBU, BAMG is  
93 scissile at two identical sites. This implies that a single cleavage event in the members of an  
94 ensemble of identical cross-linked peptide pairs results in a mixture of two pairs of products with a  
95 characteristic mass difference. The short versions of the cleaved peptides are denoted  $\alpha$ S and  $\beta$ S and  
96 the long version are denoted  $\alpha$ L and  $\beta$ L. For BAMG the mass difference between  $\alpha$ S and  $\alpha$ L and  
97 between  $\beta$ S and  $\beta$ L is 125.048 Da,  $\alpha$ L and  $\beta$ L being modified by the remnant of the cross-linker in the  
98 form of a  $\gamma$ -lactam, while  $\alpha$ S and  $\beta$ S are unmodified after the cleavage (Fig 1). The presence of four  
99 such cleavage products directly reveals the masses of the composite peptides. Also double cleavage

100 events usually take place, by which cleavage of a cross-link amide bond occurs along with a peptide  
101 bond cleavage. Often signals from only one pair of cleavage products are detectable, usually from  
102 the shortest composite peptide, the other pair of cleavage products being completely fragmented by  
103 secondary cleavages. If only one pair of peaks with a 125 Da mass difference is present in the mass  
104 spectrum, the mass of the other peptide can be calculated by subtracting the mass of the first  
105 peptide from the mass of the precursor ion. We have shown that 83% of the mass spectra from 401  
106 different cross-linked peptide pairs display at least one pair of cleavage products from either the  $\alpha$ -  
107 peptide or the  $\beta$ -peptide. In principle the presence of 2 cleavage products in the combinations  $\alpha S$   
108 and  $\beta S$ ,  $\alpha S$  and  $\beta L$ ,  $\alpha L$  and  $\beta S$  and  $\alpha L$  and  $\beta L$  can also be used to deduce the masses of  $\alpha$  and  $\beta$ . These  
109 combinations occur in an additional 5% of the instances [11]. Peptide bond cleavages can also result  
110 in ion pairs differing 125.048 Da, denoted  $\alpha S y$  or  $\alpha S b$  for the short and  $\alpha L y$  or  $\alpha L b$  for the long (+  
111 125.048 Da) versions of y or b ions from the  $\alpha$  peptide. For the  $\beta$  peptide short and long version of y  
112 and b ions are denoted  $\beta S y$ ,  $\beta L y$ ,  $\beta S b$  and  $\beta L b$ . pLink 2 was adapted for BAMG cross-link  
113 identification by considering BAMG-specific fragment ions. It is anticipated that the presence of y-  
114 ion or b-ion pairs with mass differences of 125.024 Da will not prevent assessment of the correct  
115 masses of  $\alpha$  and  $\beta$  by pLink 2, although the presence of these ion pairs may result in the calculation  
116 of more than one candidate for the masses of  $\alpha$  and  $\beta$  of a given precursor ion.

117 Four different protein databases were used for identification; (i) the Uniprot protein database of  
118 *Bacillus subtilis* (4260 entries); (ii) a hybrid database composed of the proteins from database (i) and  
119 the Uniprot protein database of *Saccharomyces cerevisiae* (6043 entries) [20]; (iii) a database of the  
120 673 proteins identified in the primary SCX fractions and (iv) a hybrid database composed of the  
121 proteins from database (iii) and 1085 human proteins identified in a SEC fraction of a HeLa cell  
122 nuclear extract[12]. The yeast and human proteins in the hybrid databases are used as a source of  
123 decoy sequences for FDR estimations.

124 *2.3. Mass spectrometric criteria for cross-link identity assignment*

125 The following parameters were used for pLink 2 searches of tryptic peptides using the mgf dataset as  
126 shown in Supplementary Table S1; mass accuracy for precursor and fragment ions: 25-75 ppm,  
127 depending on the mgf file; up to two missed cleavages allowed; masses: 600-6000; lengths for  $\alpha$  and  
128  $\beta$  peptides: 6-60 amino acids; a carbamidomethyl group at C as fixed modification; oxidation at M as  
129 variable modification; 5% or 0.1% FDR for cross-linked spectrum matches. With these settings intra-  
130 protein cross-links were assigned as nominated by pLink 2. For assignment of inter-protein cross-  
131 links the following additional requirements were taken into account: the assignment of least 4  
132 unambiguous y ions for  $\alpha$  and  $\beta$  peptides with a length of 12 amino acids or more, and at least 3  
133 unambiguous y ions for peptides consisting of 11 amino acids or less along with a matched intensity  
134 scoring higher than 35%. For a matched intensity score of more than 50%, at least 2 unambiguous y  
135 ions are sufficient for assignment of peptides consisting of 11 amino acids or less, provided that also  
136 at least 2 unambiguous b ions can be assigned to the peptide. A y or b ion is considered ambiguous if  
137 it can also be assigned to one or more other fragments. A yS and yL or bS and bL ion pair with the  
138 125 Da mass difference is counted only once for the requirement with respect to the minimal  
139 number of unambiguous y or b ions for validation and assignment. Doubly charged y or b ions at m/z  
140  $\leq 700$  are not taken into account. The ignorance of b ions as a selection criterion, except for short  
141 peptides and at high matched intensity as described above, is based on their relatively low  
142 occurrence as compared with y ions [9].

143 Matched intensity and numbers of unambiguously assigned y and b ions are mentioned in result  
144 tables only once for candidates of which more than one MSMS result was put forward as a result of  
145 multiple selections for MSMS by the mass spectrometer. Usually the MSMS spectrum with the  
146 highest matched intensity was chosen, provided that also the other criteria for assignment were  
147 met.

148 *2.3. Elution time window during SCX as a criterion for cross-link identity assignment*

149 A further filter concerned the mass of a cross-linked peptide pair and the calculated charge state in  
150 relation to the elution time. The charge state is calculated at pH 3, i.e., the conditions of strong  
151 cation exchange chromatography, assuming protonation of all acid and basic groups. To assess the  
152 distribution of mass and charge in relation of elution time in SCX chromatography of true positives  
153 we took into account all intra-protein cross-linked peptide pairs identified by pLink 2 and only the  
154 inter-protein peptide pairs that were identified by both Raeng and pLink 2. Identical peptide pairs  
155 that eluted in different SCX fractions were also taken into account.

156 *2.4. FDR estimation; the use of different decoy sequences for inter- and intra-protein peptide pairs;  
157 selectivity and sensitivity*

158 The overall FDR is defined as  $FDR = d/(d + t) \times 100\%$ , in which d is the number of decoy  $\alpha$ - $\beta$  sequence  
159 hits and t is the number of identified target  $\alpha$ - $\beta$  sequences. A decoy  $\alpha$ - $\beta$  peptide pair consists of two  
160 human or yeast sequences or of one human or yeast and one target (*B. subtilis*) sequence. For intra-  
161 protein cross-linked peptide pairs, decoy sequences for  $\alpha$  and  $\beta$  are from the same yeast or human  
162 protein, whereas for inter-protein peptide pairs the decoy sequences are from different proteins.  
163 Also target-reversed and reversed-reversed peptide pairs with different sequences of  $\alpha$  and  $\beta$   
164 peptides from the same proteins are decoy sequences for intra-protein cross-linked peptide pairs.  
165 However, previously we showed that no reversed-reversed versions and only three target-reversed  
166 versions of these intra-protein decoy sequences were put forward on a total of 1288 decoy hits[11].

167 **3. Results and discussion**

168 *3.1. Protein composition of the extract obtained after *in vivo* cross-linking of exponentially growing  
169 cells with BAMG*

170 Previously we have identified cross-linked peptide pairs from protein complexes in the size range  
171 400 kDa to 1-2 MDa after SEC of an extract obtained from an exponentially growing *B. subtilis*  
172 culture treated with BAMG. A size exclusion chromatogram of the protein extract is shown in Fig 2.

173 In the selected SEC fractions 4-7 we identified 673 proteins [9]. The known MW of most of these  
174 proteins is much smaller than the size range of 400 kD to 1-2 MDa. Besides specific cross-linking in  
175 stable and transient protein complexes also cross-linking during random diffusional encounters  
176 between proteins in the concentrated cytosol may attribute to this high molecular weight shift and  
177 size heterogeneity. It is anticipated that only specific interactions will be found, assuming that in a  
178 digest the many particular non-specific cross-links are so rare that they will remain largely  
179 undetected by LCMSMS.

180 *3.2. Isolation of cross-linked peptides by diagonal strong cation exchange (SCX) chromatography*

181 The principle of diagonal chromatography to isolate BAMG-cross-linked peptides from the bulk of  
182 unmodified peptides [12] is depicted in Fig 3. Isolation of peptides with a specific reactivity by two-  
183 dimensional liquid chromatography was introduced in 2002 in order to increase the numbers of  
184 proteins identified by LCMSMS [21]. In this approach, coined diagonal chromatography, target  
185 peptides in chromatographic fractions are subjected to a reaction that modifies their  
186 chromatographic retention time. In the second dimension peptides in treated fractions sequester  
187 from the bulk of unmodified peptides that elute at the same retention time as in the primary run.  
188 Here we use reduction of the azido group to an amino group on the spacer of a crosslink in peptides  
189 to increase the charge state. Also peptides with an internal cross-link, called loop-linked peptides  
190 and peptides with one modified lysine residue, called mono-linked peptides, the second reactive  
191 ester of the cross-linker being hydrolyzed, are enriched by diagonal SCX chromatography.

192 An additional advantage of purification of target peptides by diagonal SCX chromatography is the  
193 characteristic relationship between de elution time and the mass range and positive charge range of  
194 cross-linked peptide pairs at pH 3.0 of the mobile phase (Table 1). Table 1 has been composed from  
195 the data of 626 cross-linked peptide pairs (section 2.4), showing that (i) highly charged cross-linked  
196 peptide pairs tend to elute late and (ii) at each charge state 4<sup>+</sup>, 5<sup>+</sup> and 6<sup>+</sup> the mass range decreases at  
197 increasing elution time. Three 5<sup>+</sup> charged peptides from fraction 10 and one 6<sup>+</sup> charge peptide from

198 fraction 12 form the only exceptions to the otherwise consistent distribution. This consistency  
199 enables the use of elution time in relation to the charge and mass of a candidate cross-link as one of  
200 the criteria to discriminate between true and false positives.

201 3.3. *pLink 2 is a fast search engine for identification of BAMG-cross-linked peptides, but for a FDR*  
202 *< 5% for inter-protein cross-linked peptide pairs further filtering is required.* We tested the suitability  
203 of pLink 2 as a search engine to identify BAMG-cross-linked peptides using an existing LCMSMS  
204 dataset. The dataset consisted of mgf files from LCMSMS analysis of 10 fractions obtained by  
205 diagonal SCX chromatography of trypsin-digested material present in SEC fraction 4-7. Previously we  
206 identified with the Raeng/Mascot approach several inter-protein cross-linked peptide pairs that  
207 fulfill the requirements of a minimal peptide length of 6 amino acids for identification by pLink 2  
208 (Table 2). These inter-protein cross-linked peptide pairs revealed 41 interactions between different  
209 proteins. The dataset also contained several homo-dimeric peptide pairs and intra-protein peptide  
210 pairs. Although the results were obtained by searching the entire *B. subtilis* sequence database, all  
211 inter-protein cross-linked peptide pairs belong to the set of 673 actually identified proteins in the  
212 analyzed SEC material. Only a few homo-dimeric and intra-protein peptide pairs did not belong to  
213 the 673 identified proteins.

214 The 13 MSMS data files (supplementary Table S1) in mgf format with an average size of about 14.4  
215 MB were separately analyzed by pLink in the very short time of about 3 min on average per mgf file.  
216 Nearly 10,000 spectra were identified by pLink 2 at less than 5% FDR (Table 3), distributed over intra-  
217 protein and inter-protein cross-linked peptide pairs, mono-linked peptides, loop-linked peptides and  
218 regular peptides. Table 2 depicts the numbers of the non-redundant intra-protein cross-linked  
219 peptide pairs, inter-protein peptide pairs and homo-dimeric peptide-pairs identified by pLink 2. Less  
220 than 3% of the intra-protein cross-linked peptide pairs did not belong to the 673 independently  
221 identified proteins (Supplementary Table S2, column F, salmon-pink highlighted). However, no less  
222 than 24 out of the 89 inter-protein cross-linked peptide pairs nominated by plink 2 did not belong to

223 the 673 identified proteins (Supplementary Table S2, column F, pink and cyan highlighted).  
224 Moreover, 16 of the 89 nominated species violated the expected elution time in CX chromatography  
225 based on their mass and calculated charge at the pH (pH 3.0) of the mobile phase in SCX  
226 chromatography (Table 1 and Supplementary Table S2, column D, blue highlighted). The presence of  
227 a relatively large number of inter-protein peptide pair candidates not belonging to the 673 most  
228 abundant proteins and the violation by several inter-protein peptide pairs of the mass and charge  
229 rules of the SCX elution times strongly suggests the presence of false positive nominations among  
230 these inter-protein cross-linked species. This observation is not unexpected, since it is well  
231 documented that by searching a large sequence database at a given FDR at CSM level, practically all  
232 false positives are confined to inter-protein peptide pairs [9–11]. Based on a total of 1856 spectra of  
233 cross-linked peptide pairs (Table 3) we roughly calculate a FDR at CSM level of 1.7% assuming that  
234 the 32 total inter-peptide cross-linked peptide pairs with aberrant SCX elution time, or of which at  
235 least one peptide does not belong to the 673 independently proteins, are false positives. This would  
236 imply a FDR of 11.3% for the 252 inter-protein peptide pairs or, with 30 non-redundant decoy hits,  
237 about 25% for the 89 non-redundant inter-protein peptide pairs put forward by pLink 2 (Tabel 2).  
238 The FDR would be even larger if related to PPIs. These trends are in agreement with a previous  
239 discussion on FDR estimations on different levels[22].

240 *3.4. A target-decoy database to find criteria for a low FDR of inter-protein peptide pairs*

241 A target-decoy database is used by pLink 2 for FDR calculation. The decoy database consists of the  
242 reversed sequence of the target database. Since reversed sequences are not reported by pLink we  
243 decided to use a hybrid database in which the entire *B. subtilis* Uniprot protein database (4260  
244 entries) was combined with the entire *Saccharomyces cerevisiae* Uniprot protein database (6043  
245 entries). In this way we could get insight in the nature of false positives, i.e. nominations with one or  
246 two yeast sequences, in order to find criteria to discriminate between true and false positives.

247 The result of a pLink search with the hybrid database is shown in Table 2. With the increase of the  
248 search space while keeping the pLink 2 settings at 5% FDR, the numbers of inter-protein and intra-  
249 protein peptide pairs overlapping with the Raeng cross-links dropped by about 10% (Table 2). Also  
250 the number of homodimeric peptide pairs decreased. In addition, pLink 2 nominated several decoy  
251 sequences. All decoy sequences contain peptide pair sequences from different proteins  
252 (supplementary Table S3, column F, salmon-pink highlighted). Since no decoy sequences are present  
253 with peptide pairs from the same yeast protein, the FDR for intra-protein target peptide pairs must  
254 be very low. Previously we also noticed that only about 0.3% of all decoy hits consisted of intra-  
255 protein target and reversed sequences in a complex sample of cross-linked human proteins [11]. This  
256 justifies the inclusion of all intra-protein cross-links, along with the common set of inter-protein  
257 cross-links put forward by Rang and pLink 2, in the dataset to calculate the mass and charge  
258 distribution patterns in relation with the elution time in SCX chromatography (Table 1). It appeared  
259 that some decoy sequences put forward by pLink 2 eluted in an unexpected time window. We use an  
260 anomalous elution time during SCX chromatography as one of the criteria for filtering the dataset to  
261 decrease the FDR for inter-protein cross-linked peptide pairs.

262 We also noticed that many decoy sequences showed relatively few y ions assigned in one or both  
263 composite peptides. Furthermore ambiguity in the assignment of y or b ions occurred sometimes. A  
264 relatively low matched intensity of MS/MS spectra was also not uncommon. Therefore we applied  
265 thresholds for matched intensity and for the required number of unambiguously assigned y ions and  
266 sometimes b ions as described in sections 2.3 and 2.4.

267 Another source of nominated decoy sequences are cases in which pLink 2 put forward 2 candidates  
268 for the same precursor ion or for two different precursor ions with the same mass and eluting in the  
269 LCMSMS run within a time window of 5 sec. In a few other instances alternative candidates were  
270 also possible for a given precursor ion when certain post-translational modifications were taken into  
271 account, namely formylation in two instances and carbamidomethylation at the N-terminus in three

272 instances. In supplementary Table S4 these double nominations are listed. In all cases only the best  
273 scoring candidate was assigned, or, in case of equal scores, none of the candidates. In  
274 supplementary Tables S2, S3, S5 and S6 these co-called ambiguous nominations that are not  
275 assigned are cyan highlighted in column C. In future experiments, the otherwise rare  
276 carbamidomethylation at the N-terminus can be circumvented by preventing the presence of  
277 iodoacetamide during digestion by trypsin. Formylation is probably unavoidable [23] in the sample  
278 preparation work flow where relatively high concentrations of ammonium formate are used to elute  
279 peptides during SCX chromatography. To keep the formylation level as low as possible, the SCX  
280 fractions were snap frozen in liquid nitrogen immediately after elution followed by lyophilisation

281 The aim of retention time window determination, application of thresholds for matched intensity  
282 and for the numbers of unambiguously assigned y and b ions a, and exclusion of double nominations  
283 for the same precursor ions is to lower substantially the FDR, while a high percentage of true  
284 positive cross-linked pairs should survive the stringent criteria. Since intra-protein cross-links were  
285 identified under conditions that no intra-protein decoy sequences were detected, these target cross-  
286 links can be considered as true positives. Therefore intra-protein peptide pairs can be used to assess  
287 the effect of filtering on the sensitivity of assignment of true positive inter-protein cross-linked  
288 peptide pairs.

289 *3.5. Application of mass spectrometric and SCX chromatographic criteria to obtain a low FDR for*  
290 *inter-peptide protein pairs*

291 The effect of application of the criteria as formulated in the Material and Methods sections 2.3 and  
292 2.4 and discussed in section 3.4 is shown in Table 2 (column assigned target peptide pairs and  
293 column assigned decoy sequences). For details see columns H-J in supplementary Tables S2 and S3).  
294 Only one decoy sequence remains that fulfils the criteria, corresponding to a FDR of about 2% for  
295 non-redundant inter-protein peptide pairs (Supplementary Table S3). With one exception the  
296 nominated inter-protein peptide pairs of which at least one of the  $\alpha$  or  $\beta$  peptide did not belong to

297 the 673 independently identified proteins in the sample were all rejected by the filtering  
298 (supplementary Table S2). This shows that the composite filter that we developed is very effective to  
299 prevent suspicious candidates from assignment.

300 Application of the composite filter to the intra-protein cross-linked peptide pairs identified by pLink  
301 2 from the entire *B. subtilis* database shows that 77% fulfilled the criteria for assignment of inter-  
302 protein peptide pairs (supplementary Table S2, columns G, H and I). This shows that also true  
303 positive cross-linked peptide pairs revealing protein-protein interactions can be detected at a very  
304 low FDR and high sensitivity.

305 The small drop in the number of identified peptides when the size of the database increase  
306 prompted to interrogate a much smaller database with pLink 2, in the expectation to identify some  
307 more cross-linked peptide pairs. In this case the protein database to be searched was composed of  
308 the 673 independently identified proteins in the sample. For decoy sequences 1085 human proteins  
309 were added to the *B. subtilis* set of proteins. Nominated and identified cross-links and decoy  
310 sequences are listed in Supplementary Tables S5 and S6. The results are summarized in Table 2,  
311 showing a small increase in the number of assigned inter-protein and intra-protein cross-linked  
312 peptide pairs as compared with the results obtained by interrogation of the entire *B. subtilis*  
313 sequence database.

314 *3.6. Resemblances and differences of results and approaches between pLink 2 and Raeng/Mascot*  
315 The combined searches showed that the overlap between the identifications by Raeng/Mascot and  
316 pLink 2 is about 80%. In supplementary figure 1 pLabel-generated mass spectra are depicted with  
317 input of the mgf files of the inter-protein cross-linked peptide pairs that had escaped detection by  
318 pLink 2. All spectra fulfilled the criteria for assignment. On the other hand pLink identified several  
319 cross-linked peptide pairs that had been overlooked by Raeng/Mascot (see for 4 examples the mass  
320 spectra in supplementary figure 2). The amount of cross-links put forward by pLink 2 and Raeng may

321 depend on how exhaustive pLink 2 searches for the right peptide masses of  $\alpha$  and  $\beta$ , sometimes  
322 against a background of secondary fragment pairs differing 125.048 Da, and how exhaustive it  
323 searches for their identities, while in the Raeng/Mascot approach the suitability of Mascot to find  
324 candidate sequences for  $\alpha$  and  $\beta$  under the given conditions may be limiting.

325 *3.7. New PPIs detected by pLink 2*

326 The new cross-links identified by pLink 2 revealed 6 PPIs not identified in the Raeng/Mascot  
327 approach. One cross-linked peptide pair points to the interaction by which enzyme I (PT1\_BACSU)  
328 from the sugar PTS system transfers a phosphoryl group from phosphoenolpyruvate (PEP) to the  
329 phosphoryl carrier protein HPr (PTHP\_BACSU) [24]. Another cross-link reveals an interaction  
330 between the RNA chaperone CspB (CSBP\_BACSU) and ribosomes, in particular with the 50S  
331 ribosomal protein L7/L12 (RL7\_BACSU). Previously we found already a cross-link between CspB and  
332 the 30S ribosomal protein S2 [9]. This shows that CspB acts in close proximity of ribosomes,  
333 corroborating other observations, although a direct interaction with ribosomes could not be  
334 demonstrated before [25]. Four cross-links point to novel PPIs (Table 4). In supplementary Fig. 2 the  
335 corresponding mass spectra are depicted. One PPI not reported up to now as far as we know is  
336 revealed by a cross-linked peptide pair from the transition state regulatory protein AbrB  
337 (ABRB\_BACSU) and translation elongation factor Tu (EFTU\_BACSU). AbrB is known for its DNA  
338 binding activity of a large number of sites [26,27]. It can be phosphorylated by several protein  
339 kinases and it interacts with the anti-repressor AbbA [28–30]. The physiological significance of the  
340 interaction of AbrB with elongation factor Tu is not clear. AMPA\_BACSU is a cytosolic  
341 aminopeptidase with broad specificity [31]. The interaction with the 50S ribosomal protein L17  
342 (RL17\_BACSU), which is located in close vicinity of the site where the nascent polypeptide emerges  
343 from the ribosome at the end of the tunnel [32] raises the question whether AmpA plays a role in  
344 either the removal of aminoterminal (formyl)-methionine during protein synthesis along with the  
345 canonical methionine aminopeptidases MAP and YflG [33,34] or in another post-translational

346 modification. The only inter-protein cross-linked peptide pair of which one of the proteins (YopJ) is  
347 not a member of the 673 independently identified proteins, revealed an interaction with the  $\alpha$ -  
348 subunit of RNA polymerase. The function of YopJ, a SPbeta prophage-derived protein, is not known.  
349 An intriguing interaction is found between the essential protein YlaN (YLAN\_BACSU) [35] and the  
350 ferric uptake repressor Fur (FUR\_BACSU) [36]. The function of YlaN has long been an enigma, but  
351 recent data point to a role in FeS cluster biogenesis [37]. The interaction of Ylan with Fur as shown  
352 here may provide a clue to understand how the effect of Ylan on FeS cluster biogenesis is brought  
353 about.

354 *3.8. Contributions of the different assignment criteria to obtain a low FDR for inter-protein peptide  
355 pairs.*

356 In Table 5 the contributions are listed of the different criteria used to discriminate between true and  
357 false positives. Details are depicted in Supplementary Tables S7 and S8. In the majority of cases the  
358 acceptance of assignment of a decoy sequence can only be prevented based on a single criterion,  
359 most decoy candidates passing the thresholds for the other criteria. The most discriminating  
360 criterion is the number of unambiguously assigned y ions. Unambiguousness in this respect is crucial,  
361 since it can significantly lower the FDR as shown in Table 5. The dependence of the elution time  
362 window during SCX chromatography on mass and charge of cross-linked peptide pairs is also a  
363 powerful criterion to identify a significant fraction of false positive inter-protein peptide pairs. SCX  
364 fractionation is often used for cleavable and non-cleavable cross-linkers using 1D [34–36]  
365 or 2D approaches [41]. The usually large population of intra-protein cross-linked peptide pairs can  
366 be used to prepare a set of references for SCX elution time windows. Here and previously [11,12] we  
367 have shown that the chance that intra-protein cross-link pairs are found by accident by searching an  
368 entire species specific sequence database is so small that false positives will be rare in this category  
369 of cross-links .

370

371 3.9. A search at 0.1% FDR also requires additional filtering to obtain a low FDR for inter-protein  
372 peptide pairs

373 Finally the dataset was searched at 0.1% FDR by pLink 2 against the entire *B. subtilis* database. The  
374 data are depicted in Table 2. Slightly less candidates were nominated for both inter-protein peptide  
375 pairs, inter-protein peptide pairs and homo-dimer peptide pairs than in the search at 5% FDR.  
376 However, compared with the search at 5% FDR, exactly the same number and same identity of inter-  
377 protein peptide pairs were assigned after application of the composite filter. Only a few less  
378 candidates with aberrant SCX elution times or with one or both peptides not belonging to the 673  
379 independently proteins were rejected by applying the composite filter than in the 5% FDR search.  
380 The relatively large number of rejected spurious candidates in a search with an overall FDR as low as  
381 0.1%, underscores the usefulness of the composite filter to obtain a low FDR for inter-protein  
382 peptide pairs during interrogation of a large sequence database.

383

384 **4. Conclusions**

385 Here we show that pLink 2 efficiently nominated cross-linked peptide pairs from complex protein  
386 extracts after in vivo treatment of exponentially growing cells with BAMG. pLink 2 is also extremely  
387 fast as compared with use of our in house developed program Raeng for nomination of BAMG-cross-  
388 linked peptides. However, at an overall FDR of 5% or 0.1%, additional filtering is required to obtain a  
389 low FDR for inter-protein cross-linked peptide pair identifications. This is due to the notion that false  
390 positives are practically all confined to inter-protein peptide pairs when a search space as large as an  
391 entire species specific database is interrogated at a given FDR, if equal criteria for assignment of  
392 intra-protein and inter-protein peptides are used [11]. This can be circumvented by applying more  
393 stringent criteria for inter-protein cross-linked peptide pairs than for intra-protein cross-links. Here  
394 we show that matched intensity, the number of y ions to be assigned for both  $\alpha$  and  $\beta$  and the

395 unambiguousness of y ions, and to a lesser extent, b ions are useful criteria to diminish the FDR for  
396 inter-protein cross-linked peptide pairs. For future use of pLink 2 as a search engine for BAMG-cross-  
397 linked peptides pairs it would be useful if these criteria could be implemented in the pLink 2 code for  
398 cross-link approval. Also the number of assigned b ions may be included in the filter criteria, but with  
399 a correction factor for the relatively low average abundance as compared with the abundance of y  
400 ions. The isolation method used for BAMG-cross-linked peptides, diagonal SCX chromatography,  
401 offers an additional criterion for filtering based on their mass and charge at the pH of the  
402 chromatographic mobile phase in relation to the elution time.

403 About 20% high scoring candidates identified previously with Raeng had escaped detection by pLink  
404 2, and vice versa. While pLink 2 is fast and efficient, identifying approximately the same number of  
405 cross-linked-peptide pairs as compared with our previous approach, it would be worthwhile to  
406 understand the reasons for the slightly difference in output between these two approaches for the  
407 benefit of future research with BAMG and other cleavable cross-linkers.

#### 408 **Acknowledgements**

409 The authors thank PhD student Zhen-Lin Chen for adapting pLink 2 to enable the use of BAMG in the  
410 stepped HCD mode and for valuable comments during the preparation of the manuscript.

#### 411 **References**

- 412 [1] W.H. Dunham, M. Mullin, A.-C. Gingras, Affinity-purification coupled to mass  
413 spectrometry: Basic principles and strategies, *PROTEOMICS*. 12 (2012) 1576–1590.  
414 <https://doi.org/10.1002/pmic.201100523>.
- 415 [2] R. Varnaité, S.A. MacNeill, Meet the neighbors: Mapping local protein interactomes by  
416 proximity-dependent labeling with BioID, *Proteomics*. 16 (2016) 2503–2518.  
417 <https://doi.org/10.1002/pmic.201600123>.
- 418 [3] D. Salas, R.G. Stacey, M. Akinlaja, L.J. Foster, Next-generation Interactomics:  
419 Considerations for the Use of Co-elution to Measure Protein Interaction Networks,  
420 *Molecular & Cellular Proteomics*. 19 (2020) 1–10.  
421 <https://doi.org/10.1074/mcp.R119.001803>.
- 422 [4] A. Paiano, A. Margiotta, M. De Luca, C. Bucci, Yeast Two-Hybrid Assay to Identify  
423 Interacting Proteins, *Curr Protoc Protein Sci*. 95 (2019) e70.  
424 <https://doi.org/10.1002/cpps.70>.

425 [5] J.D. Chavez, J.P. Mohr, M. Mathay, X. Zhong, A. Keller, J.E. Bruce, Systems structural  
426 biology measurements by in vivo cross-linking with mass spectrometry, *Nat Protoc.* 14  
427 (2019) 2318–2343. <https://doi.org/10.1038/s41596-019-0181-3>.

428 [6] C. Iacobucci, M. Götze, A. Sinz, Cross-linking/mass spectrometry to get a closer view  
429 on protein interaction networks, *Curr. Opin. Biotechnol.* 63 (2019) 48–53.  
430 <https://doi.org/10.1016/j.copbio.2019.12.009>.

431 [7] C. Yu, L. Huang, Cross-Linking Mass Spectrometry: An Emerging Technology for  
432 Interactomics and Structural Biology, *Anal. Chem.* 90 (2018) 144–165.  
433 <https://doi.org/10.1021/acs.analchem.7b04431>.

434 [8] P.T. Kasper, J.W. Back, M. Vitale, A.F. Hartog, W. Roseboom, L.J. de Koning, J.H. van  
435 Maarseveen, A.O. Muijsers, C.G. de Koster, L. de Jong, An aptly positioned azido  
436 group in the spacer of a protein cross-linker for facile mapping of lysines in close  
437 proximity, *Chembiochem.* 8 (2007) 1281–1292.  
438 <https://doi.org/10.1002/cbic.200700150>.

439 [9] L. de Jong, E.A. de Koning, W. Roseboom, H. Buncherd, M.J. Wanner, I. Dapic, P.J.  
440 Jansen, J.H. van Maarseveen, G.L. Corthals, P.J. Lewis, L.W. Hamoen, C.G. de Koster,  
441 In-Culture Cross-Linking of Bacterial Cells Reveals Large-Scale Dynamic Protein-  
442 Protein Interactions at the Peptide Level, *J. Proteome Res.* 16 (2017) 2457–2471.  
443 <https://doi.org/10.1021/acs.jproteome.7b00068>.

444 [10] H. Buncherd, M.A. Nessen, N. Nouse, S.K. Stelder, W. Roseboom, H.L. Dekker, J.C.  
445 Arents, L.E. Smeenk, M.J. Wanner, J.H. van Maarseveen, X. Yang, P.J. Lewis, L.J. de  
446 Koning, C.G. de Koster, L. de Jong, Selective enrichment and identification of cross-  
447 linked peptides to study 3-D structures of protein complexes by mass spectrometry, *J  
448 Proteomics.* 75 (2012) 2205–2215. <https://doi.org/10.1016/j.jprot.2012.01.025>.

449 [11] H. Buncherd, W. Roseboom, L.J. de Koning, C.G. de Koster, L. de Jong, A gas phase  
450 cleavage reaction of cross-linked peptides for protein complex topology studies by  
451 peptide fragment fingerprinting from large sequence database, *J Proteomics.* 108 (2014)  
452 65–77. <https://doi.org/10.1016/j.jprot.2014.05.003>.

453 [12] H. Buncherd, W. Roseboom, B. Ghavim, W. Du, L.J. de Koning, C.G. de Koster, L. de  
454 Jong, Isolation of cross-linked peptides by diagonal strong cation exchange  
455 chromatography for protein complex topology studies by peptide fragment  
456 fingerprinting from large sequence databases, *J Chromatogr A.* 1348 (2014) 34–46.  
457 <https://doi.org/10.1016/j.chroma.2014.04.083>.

458 [13] M. Glas, H.B. van den Berg van Saparoea, S.H. McLaughlin, W. Roseboom, F. Liu,  
459 G.M. Koningstein, A. Fish, T. den Blaauwen, A.J.R. Heck, L. de Jong, W. Bitter, I.J.P.  
460 de Esch, J. Luijink, The Soluble Periplasmic Domains of *Escherichia coli* Cell Division  
461 Proteins FtsQ/FtsB/FtsL Form a Trimeric Complex with Submicromolar Affinity, *J.  
462 Biol. Chem.* 290 (2015) 21498–21509. <https://doi.org/10.1074/jbc.M115.654756>.

463 [14] W. Roseboom, M.G. Nazir, N.Y. Meiresonne, T. Mohammadi, J. Verheul, H. Buncherd,  
464 A.M.J.J. Bonvin, L.J. de Koning, C.G. de Koster, L. de Jong, T. den Blaauwen,  
465 Mapping the Contact Sites of the *Escherichia coli* Division-Initiating Proteins FtsZ and  
466 ZapA by BAMG Cross-Linking and Site-Directed Mutagenesis, *Int J Mol Sci.* 19  
467 (2018). <https://doi.org/10.3390/ijms19102928>.

468 [15] F. Meier, A.-D. Brunner, S. Koch, H. Koch, M. Lubeck, M. Krause, N. Goedecke, J.  
469 Decker, T. Kosinski, M.A. Park, N. Bache, O. Hoerning, J. Cox, O. Räther, M. Mann,  
470 Online Parallel Accumulation-Serial Fragmentation (PASEF) with a Novel Trapped Ion  
471 Mobility Mass Spectrometer, *Mol. Cell Proteomics.* 17 (2018) 2534–2545.  
472 <https://doi.org/10.1074/mcp.TIR118.000900>.

473 [16] Z.-L. Chen, J.-M. Meng, Y. Cao, J.-L. Yin, R.-Q. Fang, S.-B. Fan, C. Liu, W.-F. Zeng,  
474 Y.-H. Ding, D. Tan, L. Wu, W.-J. Zhou, H. Chi, R.-X. Sun, M.-Q. Dong, S.-M. He, A

475 high-speed search engine pLink 2 with systematic evaluation for proteome-scale  
476 identification of cross-linked peptides, *Nat Commun.* 10 (2019) 3404.  
477 <https://doi.org/10.1038/s41467-019-11337-z>.

478 [17] R. Beveridge, J. Stadlmann, J.M. Penninger, K. Mechtler, A synthetic peptide library for  
479 benchmarking crosslinking-mass spectrometry search engines for proteins and protein  
480 complexes, *Nat Commun.* 11 (2020) 742. <https://doi.org/10.1038/s41467-020-14608-2>.

481 [18] M.Q. Müller, F. Dreicker, C.H. Ihling, M. Schäfer, A. Sinz, Cleavable cross-linker for  
482 protein structure analysis: reliable identification of cross-linking products by tandem  
483 MS, *Anal. Chem.* 82 (2010) 6958–6968. <https://doi.org/10.1021/ac101241t>.

484 [19] A. Kao, C. Chiu, D. Vellucci, Y. Yang, V.R. Patel, S. Guan, A. Randall, P. Baldi, S.D.  
485 Rychnovsky, L. Huang, Development of a novel cross-linking strategy for fast and  
486 accurate identification of cross-linked peptides of protein complexes, *Mol. Cell*  
487 *Proteomics.* 10 (2011) M110.002212. <https://doi.org/10.1074/mcp.M110.002212>.

488 [20] UniProt: a worldwide hub of protein knowledge, *Nucleic Acids Res.* 47 (2019) D506–  
489 D515. <https://doi.org/10.1093/nar/gky1049>.

490 [21] K. Gevaert, J. Van Damme, M. Goethals, G.R. Thomas, B. Hoorelbeke, H. Demol, L.  
491 Martens, M. Puype, A. Staes, J. Vandekerckhove, Chromatographic isolation of  
492 methionine-containing peptides for gel-free proteome analysis: identification of more  
493 than 800 *Escherichia coli* proteins, *Mol. Cell Proteomics.* 1 (2002) 896–903.  
494 <https://doi.org/10.1074/mcp.m200061-mcp200>.

495 [22] L. Fischer, J. Rappaport, Quirks of Error Estimation in Cross-Linking/Mass  
496 Spectrometry, *Anal. Chem.* 89 (2017) 3829–3833.  
497 <https://doi.org/10.1021/acs.analchem.6b03745>.

498 [23] S. Zheng, A.A. Doucette, Preventing N- and O-formylation of proteins when incubated  
499 in concentrated formic acid, *Proteomics.* 16 (2016) 1059–1068.  
500 <https://doi.org/10.1002/pmic.201500366>.

501 [24] W. Kundig, S. Ghosh, S. Roseman, PHOSPHATE BOUND TO HISTIDINE IN A  
502 PROTEIN AS AN INTERMEDIATE IN A NOVEL PHOSPHO-TRANSFERASE  
503 SYSTEM\*, *Proc Natl Acad Sci U S A.* 52 (1964) 1067–1074.

504 [25] K. Hunger, C.L. Beckering, F. Wiegeshoff, P.L. Graumann, M.A. Marahiel, Cold-  
505 induced putative DEAD box RNA helicases CshA and CshB are essential for cold  
506 adaptation and interact with cold shock protein B in *Bacillus subtilis*, *J. Bacteriol.* 188  
507 (2006) 240–248. <https://doi.org/10.1128/JB.188.1.240-248.2006>.

508 [26] O. Chumsakul, H. Takahashi, T. Oshima, T. Hishimoto, S. Kanaya, N. Ogasawara, S.  
509 Ishikawa, Genome-wide binding profiles of the *Bacillus subtilis* transition state  
510 regulator AbrB and its homolog Abh reveals their interactive role in transcriptional  
511 regulation, *Nucleic Acids Res.* 39 (2011) 414–428. <https://doi.org/10.1093/nar/gkq780>.

512 [27] T.C. Rösch, L.M. Oviedo-Bocanegra, G. Fritz, P.L. Graumann, SMTracker: a tool for  
513 quantitative analysis, exploration and visualization of single-molecule tracking data  
514 reveals highly dynamic binding of *B. subtilis* global repressor AbrB throughout the  
515 genome, *Sci Rep.* 8 (2018) 15747. <https://doi.org/10.1038/s41598-018-33842-9>.

516 [28] A. Kobir, S. Poncet, V. Bidnenko, O. Delumeau, C. Jers, S. Zouhir, R. Grenha, S.  
517 Nessler, P. Noirot, I. Mijakovic, Phosphorylation of *Bacillus subtilis* gene regulator  
518 AbrB modulates its DNA-binding properties, *Mol. Microbiol.* 92 (2014) 1129–1141.  
519 <https://doi.org/10.1111/mmi.12617>.

520 [29] A.T. Tucker, B.G. Bobay, A.V. Banse, A.L. Olson, E.J. Soderblom, M.A. Moseley, R.J.  
521 Thompson, K.M. Varney, R. Losick, J. Cavanagh, A DNA mimic: the structure and  
522 mechanism of action for the anti-repressor protein AbbA, *J. Mol. Biol.* 426 (2014)  
523 1911–1924. <https://doi.org/10.1016/j.jmb.2014.02.010>.

524 [30] B. Zhu, J. Stölke, SubtiWiki in 2018: from genes and proteins to functional network  
525 annotation of the model organism *Bacillus subtilis*, *Nucleic Acids Res.* 46 (2018) D743–  
526 D748. <https://doi.org/10.1093/nar/gkx908>.

527 [31] A. Kale, T. Pijning, T. Sonke, B.W. Dijkstra, A.-M.W.H. Thunnissen, Crystal Structure  
528 of the Leucine Aminopeptidase from *Pseudomonas putida* Reveals the Molecular Basis  
529 for its Enantioselectivity and Broad Substrate Specificity, *Journal of Molecular Biology*.  
530 398 (2010) 703–714. <https://doi.org/10.1016/j.jmb.2010.03.042>.

531 [32] J.A. Lake, Evolving ribosome structure: domains in archaeabacteria, eubacteria, eocytes  
532 and eukaryotes, *Annu. Rev. Biochem.* 54 (1985) 507–530.  
533 <https://doi.org/10.1146/annurev.bi.54.070185.002451>.

534 [33] C. You, H. Lu, A. Sekowska, G. Fang, Y. Wang, A.-M. Gilles, A. Danchin, The two  
535 authentic methionine aminopeptidase genes are differentially expressed in *Bacillus*  
536 *subtilis*, *BMC Microbiol.* 5 (2005) 57. <https://doi.org/10.1186/1471-2180-5-57>.

537 [34] C. Giglione, S. Fieulaine, T. Meinnel, N-terminal protein modifications: Bringing back  
538 into play the ribosome, *Biochimie*. 114 (2015) 134–146.  
539 <https://doi.org/10.1016/j.biochi.2014.11.008>.

540 [35] L. Xu, S.E. Sedelnikova, P.J. Baker, A. Hunt, J. Errington, D.W. Rice, Crystal structure  
541 of *S. aureus* YlaN, an essential leucine rich protein involved in the control of cell shape,  
542 *Proteins*. 68 (2007) 438–445. <https://doi.org/10.1002/prot.21377>.

543 [36] H. Pi, J.D. Helmann, Sequential induction of Fur-regulated genes in response to iron  
544 limitation in *Bacillus subtilis*, *Proc. Natl. Acad. Sci. U.S.A.* 114 (2017) 12785–12790.  
545 <https://doi.org/10.1073/pnas.1713008114>.

546 [37] J.M. Peters, A. Colavin, H. Shi, T.L. Czarny, M.H. Larson, S. Wong, J.S. Hawkins,  
547 C.H.S. Lu, B.-M. Koo, E. Marta, A.L. Shiver, E.H. Whitehead, J.S. Weissman, E.D.  
548 Brown, L.S. Qi, K.C. Huang, C.A. Gross, A Comprehensive, CRISPR-based Functional  
549 Analysis of Essential Genes in Bacteria, *Cell*. 165 (2016) 1493–1506.  
550 <https://doi.org/10.1016/j.cell.2016.05.003>.

551 [38] O. Rinner, J. Seebacher, T. Walzthoeni, L.N. Mueller, M. Beck, A. Schmidt, M.  
552 Mueller, R. Aebersold, Identification of cross-linked peptides from large sequence  
553 databases, *Nat. Methods*. 5 (2008) 315–318. <https://doi.org/10.1038/nmeth.1192>.

554 [39] F. Liu, D.T.S. Rijkers, H. Post, A.J.R. Heck, Proteome-wide profiling of protein  
555 assemblies by cross-linking mass spectrometry, *Nat. Methods*. 12 (2015) 1179–1184.  
556 <https://doi.org/10.1038/nmeth.3603>.

557 [40] K.A.T. Makepeace, Y. Mohammed, E.L. Rudashevskaya, E.V. Petrotchenko, F.-N.  
558 Vögtle, C. Meisinger, A. Sickmann, C.H. Borchers, Improving Identification of In-  
559 organello Protein-Protein Interactions Using an Affinity-enrichable, Isotopically Coded,  
560 and Mass Spectrometry-cleavable Chemical Crosslinker, *Mol Cell Proteomics*. 19  
561 (2020) 624–639. <https://doi.org/10.1074/mcp.RA119.001839>.

562 [41] V. Tinnefeld, A.S. Venne, A. Sickmann, R.P. Zahedi, Enrichment of Cross-Linked  
563 Peptides Using Charge-Based Fractional Diagonal Chromatography (ChaFRADIC), *J.*  
564 *Proteome Res.* 16 (2017) 459–469. <https://doi.org/10.1021/acs.jproteome.6b00587>.

565

566

567

568 Table 1. Charge at pH 3 and mass distributions of cross-linked peptide pairs in SCX fractions

SCX frac- tion	XL pep- tide pairs (n)	Mass range (Da) at charge +3	XL pep- tide pairs (n)	Mass range (Da) at charge +4	XL pep- tide pairs (n)	Mass range (Da) at charge +5	XL pep- tide pairs (n)	Mass range (Da) at charge 6+	XL pep- tide pairs (n)
7	30	1793-1868	2	2274-4505	28				
8	62	1869-2109	2	2274-4387	59	4377	1		
9	212			1756-3890	201	3273-4672	11		
10	72			1641-3033	69	2554-3394	3		
11	83			1734-2858	45	2861-5181	36	4055-6092	2
12	57			1513-2568	17	2540-4462	39	2661	1
13	52			1562-2090	5	2262-3828	46	3884	1
14	39			1435-1779	4	1967-3351	24	3419-4919	11
15	2					1862-2595	2	3405	1
16	17					1862-3510	13	3313-3870	4

569

570

571

572 Table 2. Overview of non-redundant spectral counts of nominated and assigned cross-linked peptide pairs

database	cross-linked peptide pair type	non-redundant spectral counts					
		Raeng		pLink 2			
		assigned target peptide pairs	nominated target peptide pairs	overlap with Raeng	assigned target peptide pairs	nominated decoy sequences	assigned decoy sequences
target; <i>B. subtilis</i> , 4620 proteins; 5% FDR	inter-protein	55	89	40	53		
	intra-protein	335	449	284	449		
	homodimer	24	13	11	13		
target/decoy; <i>B. subtilis</i> 4620 proteins/ <i>S. cerevisiae</i> 6049 proteins; 5% FDR	inter-protein		59	36	46	29	1
	intra-protein		401	278	401	0	0
	homodimer		6	4	6	0	0
target; <i>B. subtilis</i> , 673 proteins; 5% FDR	inter-protein		88	44	58		
	intra-protein		463	n.d.	463		
	homodimer		7	n.d.	7		
target/decoy; <i>B. subtilis</i> 673 proteins/ <i>H. sapiens</i> 1085 proteins; 5% FDR	inter-protein		50	40	45	30	0
	intra-protein		450	n.d.	450	0	0
	homodimer		7	n.d.	7	0	0
target; <i>B. subtilis</i> , 4620 proteins, 0.1% FDR	inter-protein		85	40	53		
	intra-protein		444	n.d.	444		
	homodimer		10	n.d.	10		

573 n.d., not determined

574

575 Table 3. Overview of spectral counts identified by pLink 2 at <5% FDR

SCX fraction	Counts					Total spectra	
	Cross-linked spectra		Loop-linked spectra	Mono-link spectra	Regular peptide spectra		
	Intra-protein	Inter-protein					
7	94	3	183	566	1185	2031	
8	124	16	143	478	779	1540	
9.2	12	9	2	30	138	191	
9.3	268	39	50	251	161	769	
9.4	254	19	103	305	185	867	
9.5	32	2	74	136	105	349	
10	174	23	53	150	404	803	
11	229	10	33	172	441	885	
12	173	17	18	130	268	606	
13	146	99	64	133	392	834	
14	66	7	37	47	320	477	
15	6	1	0	4	172	183	
16	26	7	10	22	257	322	
total	1604	252	770	2424	4807	9857	
percent	16.3	2.6	7.8	24.6	48.8		

576

577

578

579

580 Table 4. Cross-linked peptide pairs identified by pLink 2 revealing new protein-protein interactions

SCX fract- ion	peptide sequences; numbering of linked K between brackets	mass (Da)	ID of linked proteins; numbering of linked K between brackets	mass spect- rum	M.I. (%)	Y ions α	Y ions β
9.3	KEWEDVVGLVGK(1)- SVVEKMI TLGK(5)	2704	sp O32106 AMPA_BACSU (250)- sp P20277 RL17_BACSU (46)	Suppl. Fig. 2a	77	7	6
9.4	KLLDYAEAGDNIGALLR(1)- KVDELGR(1)	2773	sp P33166 EFTU_BACSU (266)- sp P08874 ABRB_BACSU (11)	Suppl. Fig. 2b	84	11	5
9.4	FAKDIAEEVYYSLK(3)- KSLEEVK(1)	2632	sp O31928 YOPJ_BACSU (127)- sp P20429 RPOA_BACSU (294)	Suppl. Fig. 2c	74	6	5
9.3	VVDKINFQDGVS(4)- VDAEKILK(5)	2445	sp P54574 FUR_BACSU (74)- sp O07638 YLAN_BACSU (23)	Suppl. Fig. 2d	72	10	3

581 M.I., matched intensity

582

583

584

585 Table 5. Contribution of different criteria used to obtain a low FDR for inter-protein cross-linked  
586 peptide pairs

Criterion	Decisive contribution to rejection of hits			
	decoy hits		target hits	
	n	percent	n	percent
Matched intensity	1	1.7	5	6.9
Required number of y ions	16	27.1	18	25.0
Required number of b ions	2	3.4	3	4.2
Unambiguous assignment to y or b ions	6	10.2	8	11.1
Expected SCX elution time window	5	8.5	7	9.7
Combination of two or more of the above criteria	22	37.3	26	36.1
Better scoring candidate with the same mass and elution time in LCMSMS	3	5.1	3	4.2
Correction for formylation or carbamidomethylation	3	5.1	2	2.8
Assigned decoy sequence	1	1.7		
Total	59		72	

587

588 **Legends to Figures**

589

590 Figure 1. Gas phase cleavage reactions of BAMG-cross-linked peptides in which the azido group has  
591 been reduced to an amine group. Upper left corner, structure of BAMG. Middle part, collision  
592 induced dissociation (CID) of a cross-linked peptide pair leads to cleavages of the two cross-link  
593 amide bonds along with cleavages of peptide bonds resulting in y and b ions. Cleavage of an amide  
594 bond probably occurs by nucleophilic attack of the amine in the spacer of the cross-link to a  
595 protonated carbonyl group of the amide bond. This leads to formation of an unmodified peptide or  
596 short version of the cleavage product ( $\alpha$ S or  $\beta$ S), the other peptide being modified by the remnant of  
597 the cross-linker in the form of a y lactam adding 125.048 Da to the mass of the peptide. This is the  
598 longer version of the cleavage product ( $\alpha$ L or  $\beta$ L). Amino acids are depicted as colored candies. The  
599 indicated gas phase charge states of the cross-linked peptide and the cleavage products are  
600 arbitrarily. The lower part is a cartoon of a fragment mass spectrum with two pairs of cleavage  
601 products with the characteristic 125.048 Da mass difference (purple sticks) and some peaks of b  
602 (green) and y (red) ions.

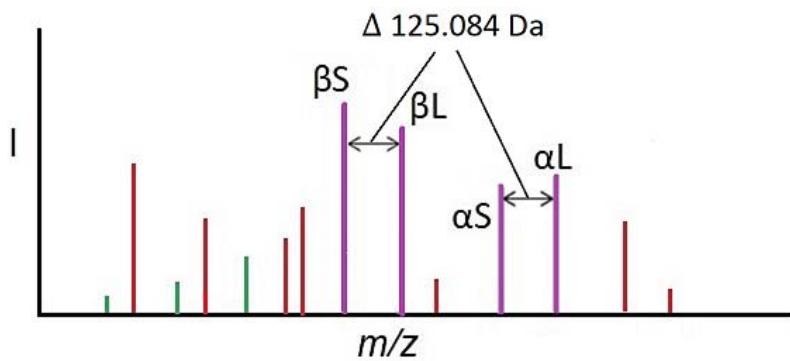
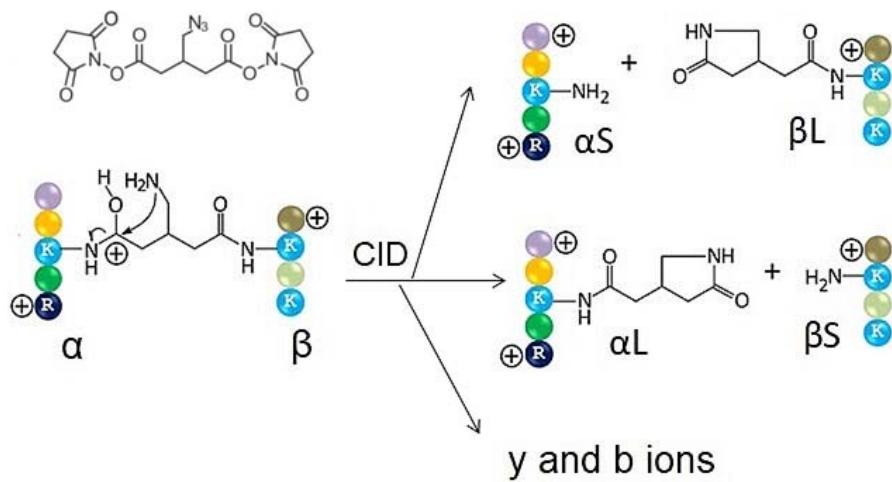
603

604 Figure 2. Size exclusion chromatogram on a Superose 6 10/300 GL column of a soluble extract from  
605 *Bacillus subtilis* cross-linked in vivo with BAMG. MW markers, 443 kDa, apoferritin (horse spleen);  
606 240 kDa catalase (bovine liver); 150 kDa, alcohol dehydrogenase (yeast); 66 kDa, bovine serum  
607 albumine, 30 kDa carbonic anhydrase (bovine erythrocytes). Fractions are indicated by green  
608 vertical lines. Duplicate chromatograms are shown by red and blue lines. Material in fraction 4-7 was  
609 used in this study.

610

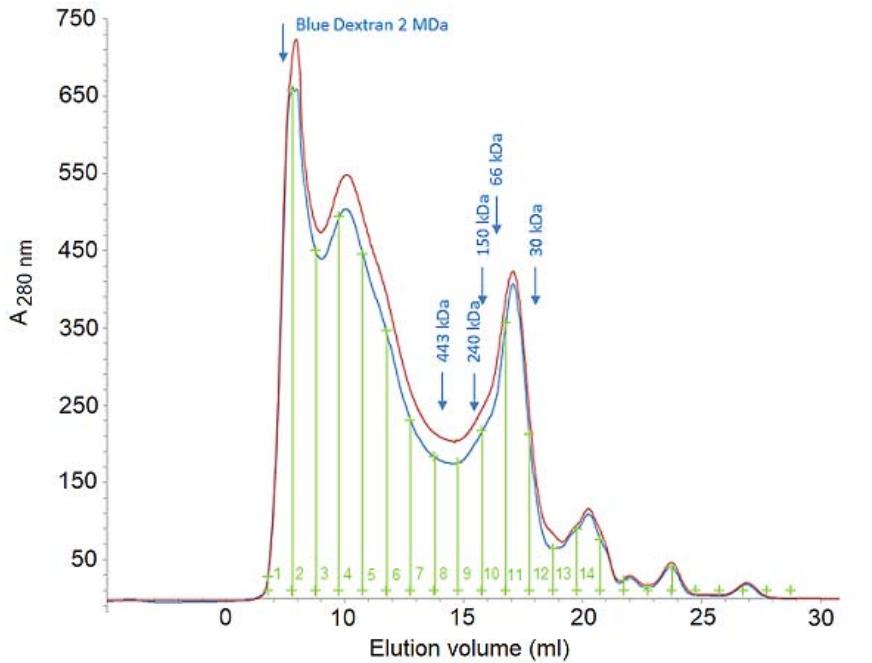
611 Figure 3. Workflow from in vivo cross-linking with BAMG to LCMSMS. Left part, 5 min after addition  
612 of BAMG to an exponentially growing *Bacillus subtilis* culture the cross-linker is quenched and cells  
613 are harvested and sonicated. (1), the cross-linked protein extract is subjected to SEC and then

614 digested to obtain a peptide mixture with cross-linked  $\alpha$ - $\beta$  peptide pairs. Right part, the peptide  
615 mixture is fractionated by strong cation exchange chromatography (first dimension SCX), using a  
616 mobile phase of pH 3 and a salt gradient (red lines) of ammonium formate to elute bound peptides.  
617 Grey, regular peptides; cyan, cross-linked peptides. (2), reduction by TCEP of the azido group in the  
618 spacer of the cross-linker to an amine group in selected SCX fractions. At the pH of the mobile phase  
619 of SCX chromatography the amino group is protonated adding an extra positive charge to cross-  
620 linked peptides. The TCEP-treated primary fractions are then separately subjected to the secondary  
621 runs of SCX. Here target peptides are sequestered from the bulk of unmodified peptides that elute at  
622 the same time as in the primary run, while the extra charge state of the cross-link peptides leads to  
623 elution at a later time. Depicted peptide charge states after (1) and (2) are calculated for pH 3,  
624 assuming full protonation of the two amino-termini plus 2 basic amino-acid side chains, carboxylic  
625 acid side chains being uncharged under these conditions.



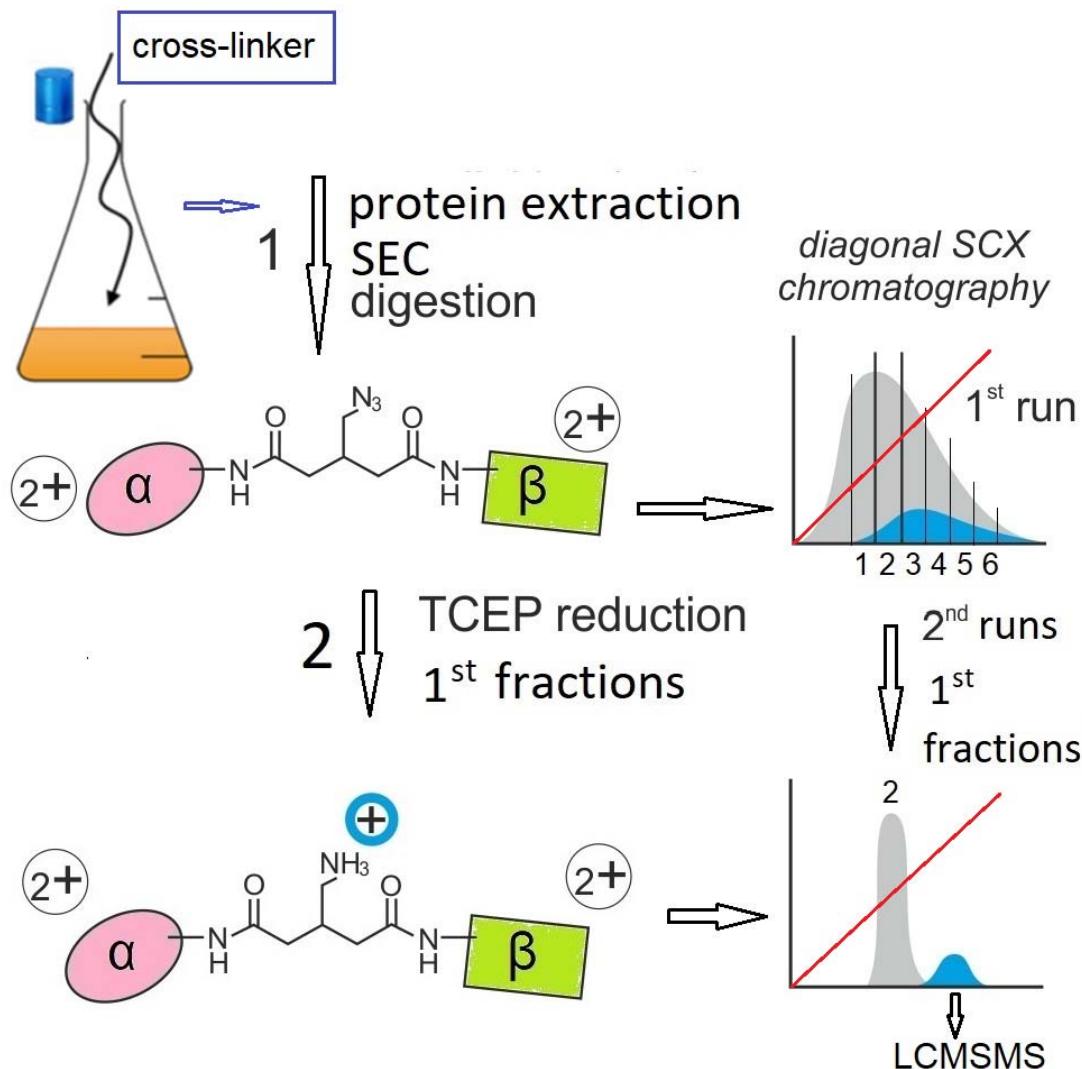
626

627 Fig 1.



628

629 Fig. 2



630

631 Fig 3.

632

633 **Supplementary material**

634 Table S1

635 Legends to Supplementary figures

636 Supplementary Figure 1

637 Supplementary Figure 2

638

639

640 Supplementary Table S1. Mgf files used in this study

mgf file used in this study	combined from original mgf files	primary SCX fraction	Size (MB)	mass accuracy (ppm) for both precursor and fragments ions
7.mgf	7.1.1, 7.2.1 and 7.3.1	7	31.9	25
8.mgf	8.1.1, 8.2.1. and 8.3.1	8	21.7	25
9.2. mgf		9	6.0	75
9.3.mgf		9	19.2	50
9.4.mgf		9	25.5	75
9.5.mgf		9	16.1	30
10.mgf	10.1.1, 10.2.1 and 10.3.1	0	17.5	25
11.mgf	11.1.1, 11.2.1 and 11.3.1	1	18.3	25
12.mgf	12.1.1, 12.2.1 and 12.3.1	2	10.4	25
13.mgf	13.1.1, 13.2.1 and 13.3.1	3	12.7	25
14.mgf	14.1.1, 14.2.1 and 14.3.1	4	8.2	25
15.mgf	15.1.1, 15.2.1 and 15.3.1	5	0.8	25
16.mgf	16.1.1, 16.2.1 and 16.3.1	6	4.4	25

641

642

643

644 **Legends to supplementary figures**

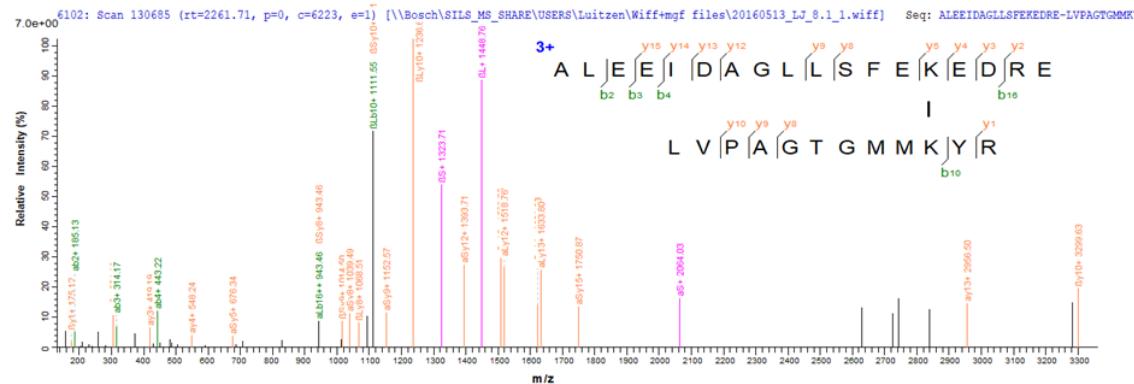
645 Supplementary figure 1. Mass spectra with high scores of inter-protein cross-linked pairs identified  
646 by Raeng/Mascot, but not pLink 2 showing interactions between a, RPOZ\_BACSU and RPOC\_BACSU;  
647 b, SIGA\_BACSU and RPOB\_BACSU; c, RPOB\_BACSU and SIGA\_BACSU; d, GREA\_BACSU and  
648 RPOC\_BACSU; e, RL11\_BACSU and RL7\_BACSU; f, RL11\_BACSU and RL7\_BACSU; g, RL31\_BACSU and  
649 RS13\_BACSU; h, GLTA\_BACSU and GLTB\_BACSU; i, ODPA\_BACSU and ODPB\_BACSU; j, RS10\_BACSU  
650 and RS18\_BACSU

651 Supplementary figure 2. Mass spectra of cross-linked peptide pairs identified by pLink 2 revealing  
652 new protein-protein interactions a, AMPA\_BACSU-RL7\_BACSU; b, EFTU\_BACSU-ABRB\_BACSU; c,  
653 YOPJ\_BACSU-RPOA\_BACSU; d, FUR\_BACSU-YLAN\_BACSU

654

655 Supplementary Figure 1. Mass spectra of intra-protein cross-linked pairs identified by  
656 Raeng?Mascot, but not by pLink 2.

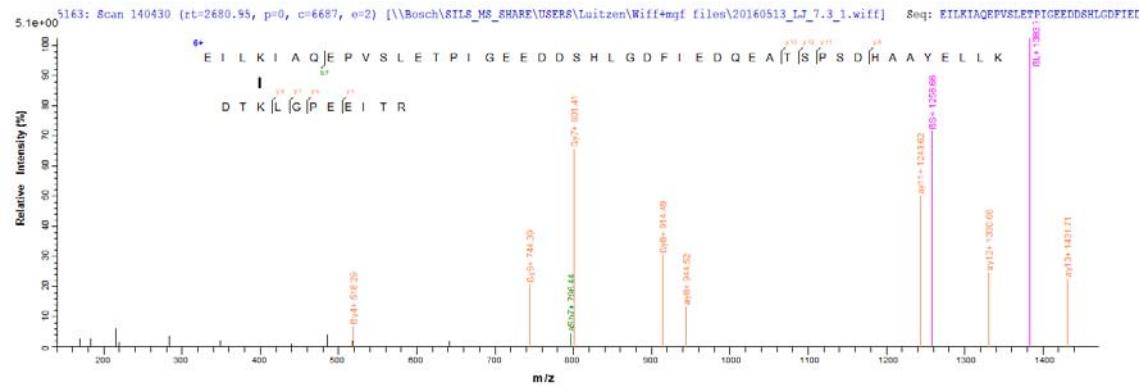
657 a



658

659 Cross-link between RPOZ\_BACSU and RPOC\_BACSU; matched intensity 83.1%

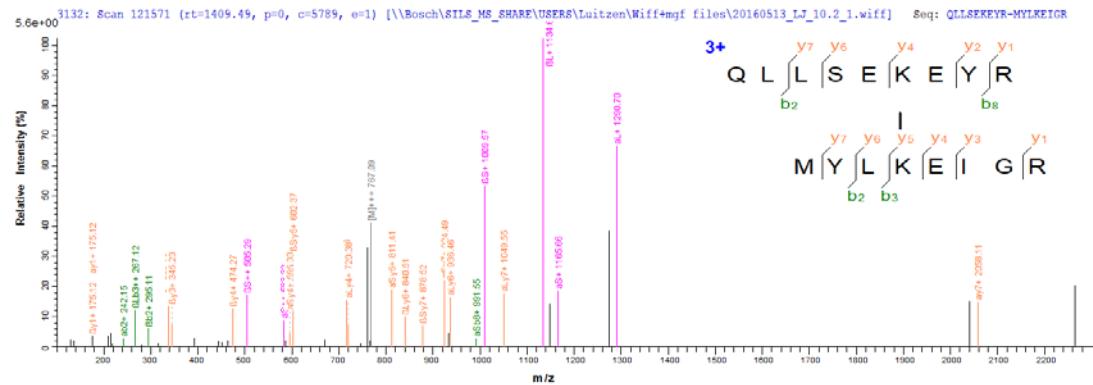
660



661

662 Cross-link between SIGA\_BACSU and RPOB\_BACSU; matched intensity 93.8%

663

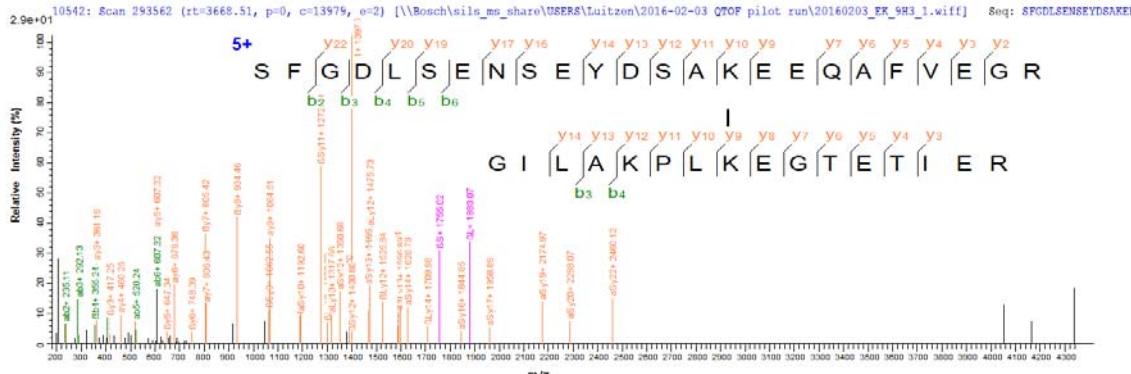


664

665 Cross-link between RPOC\_BACSU and SIGA\_BACSU; matched intensity 58.1%

666

667 d



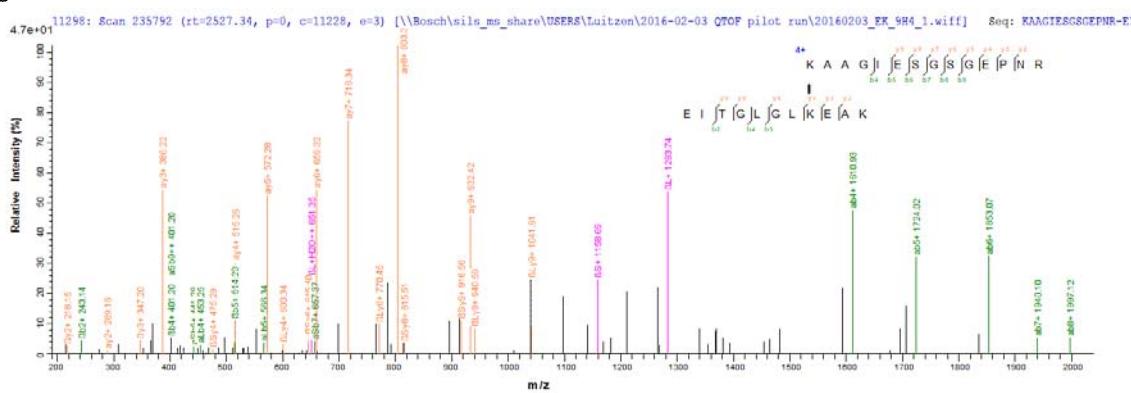
668

669

Cross-link between GRE\_A\_BACSU and RPOC\_BACSU; matched intensity 82.3%

670

e



671

672

Cross-link between RL11\_BACSU and RL7\_BACSU; matched intensity 62.0%

<img alt="Mass spectrum and sequence diagram for peptide KAAAGI[ES]GS[GEP]N R. The x-axis is m/z from 200 to 1900, and the y-axis is relative intensity from 0 to 100. The spectrum shows several peaks with labels: 263.13, 367.25, 499-500.21, 514-515.25, 516-517.25, 624-625.25, 634-635.25, 671-672.25, 692-693.25, 712-713.25, 732-733.25, 752-753.25, 772-773.25, 792-793.25, 812-813.25, 832-833.25, 852-853.25, 872-873.25, 892-893.25, 912-913.25, 932-933.25, 952-953.25, 972-973.25, 992-993.25, 1012-1013.25, 1032-1033.25, 1052-1053.25, 1072-1073.25, 1092-1093.25, 1112-1113.25, 1132-1133.25, 1152-1153.25, 1172-1173.25, 1192-1193.25, 1212-1213.25, 1232-1233.25, 1252-1253.25, 1272-1273.25, 1292-1293.25, 1312-1313.25, 1332-1333.25, 1352-1353.25, 1372-1373.25, 1392-1393.25, 1412-1413.25, 1432-1433.25, 1452-1453.25, 1472-1473.25, 1492-1493.25, 1512-1513.25, 1532-1533.25, 1552-1553.25, 1572-1573.25, 1592-1593.25, 1612-1613.25, 1632-1633.25, 1652-1653.25, 1672-1673.25, 1692-1693.25, 1712-1713.25, 1732-1733.25, 1752-1753.25, 1772-1773.25, 1792-1793.25, 1812-1813.25, 1832-1833.25, 1852-1853.25, 1872-1873.25, 1892-1893.25, 1912-1913.25, 1932-1933.25, 1952-1953.25, 1972-1973.25, 1992-1993.25, 2012-2013.25, 2032-2033.25, 2052-2053.25, 2072-2073.25, 2092-2093.25, 2112-2113.25, 2132-2133.25, 2152-2153.25, 2172-2173.25, 2192-2193.25, 2212-2213.25, 2232-2233.25, 2252-2253.25, 2272-2273.25, 2292-2293.25, 2312-2313.25, 2332-2333.25, 2352-2353.25, 2372-2373.25, 2392-2393.25, 2412-2413.25, 2432-2433.25, 2452-2453.25, 2472-2473.25, 2492-2493.25, 2512-2513.25, 2532-2533.25, 2552-2553.25, 2572-2573.25, 2592-2593.25, 2612-2613.25, 2632-2633.25, 2652-2653.25, 2672-2673.25, 2692-2693.25, 2712-2713.25, 2732-2733.25, 2752-2753.25, 2772-2773.25, 2792-2793.25, 2812-2813.25, 2832-2833.25, 2852-2853.25, 2872-2873.25, 2892-2893.25, 2912-2913.25, 2932-2933.25, 2952-2953.25, 2972-2973.25, 2992-2993.25, 3012-3013.25, 3032-3033.25, 3052-3053.25, 3072-3073.25, 3092-3093.25, 3112-3113.25, 3132-3133.25, 3152-3153.25, 3172-3173.25, 3192-3193.25, 3212-3213.25, 3232-3233.25, 3252-3253.25, 3272-3273.25, 3292-3293.25, 3312-3313.25, 3332-3333.25, 3352-3353.25, 3372-3373.25, 3392-3393.25, 3412-3413.25, 3432-3433.25, 3452-3453.25, 3472-3473.25, 3492-3493.25, 3512-3513.25, 3532-3533.25, 3552-3553.25, 3572-3573.25, 3592-3593.25, 3612-3613.25, 3632-3633.25, 3652-3653.25, 3672-3673.25, 3692-3693.25, 3712-3713.25, 3732-3733.25, 3752-3753.25, 3772-3773.25, 3792-3793.25, 3812-3813.25, 3832-3833.25, 3852-3853.25, 3872-3873.25, 3892-3893.25, 3912-3913.25, 3932-3933.25, 3952-3953.25, 3972-3973.25, 3992-3993.25, 4012-4013.25, 4032-4033.25, 4052-4053.25, 4072-4073.25, 4092-4093.25, 4112-4113.25, 4132-4133.25, 4152-4153.25, 4172-4173.25, 4192-4193.25, 4212-4213.25, 4232-4233.25, 4252-4253.25, 4272-4273.25, 4292-4293.25, 4312-4313.25, 4332-4333.25, 4352-4353.25, 4372-4373.25, 4392-4393.25, 4412-4413.25, 4432-4433.25, 4452-4453.25, 4472-4473.25, 4492-4493.25, 4512-4513.25, 4532-4533.25, 4552-4553.25, 4572-4573.25, 4592-4593.25, 4612-4613.25, 4632-4633.25, 4652-4653.25, 4672-4673.25, 4692-4693.25, 4712-4713.25, 4732-4733.25, 4752-4753.25, 4772-4773.25, 4792-4793.25, 4812-4813.25, 4832-4833.25, 4852-4853.25, 4872-4873.25, 4892-4893.25, 4912-4913.25, 4932-4933.25, 4952-4953.25, 4972-4973.25, 4992-4993.25, 5012-5013.25, 5032-5033.25, 5052-5053.25, 5072-5073.25, 5092-5093.25, 5112-5113.25, 5132-5133.25, 5152-5153.25, 5172-5173.25, 5192-5193.25, 5212-5213.25, 5232-5233.25, 5252-5253.25, 5272-5273.25, 5292-5293.25, 5312-5313.25, 5332-5333.25, 5352-5353.25, 5372-5373.25, 5392-5393.25, 5412-5413.25, 5432-5433.25, 5452-5453.25, 5472-5473.25, 5492-5493.25, 5512-5513.25, 5532-5533.25, 5552-5553.25, 5572-5573.25, 5592-5593.25, 5612-5613.25, 5632-5633.25, 5652-5653.25, 5672-5673.25, 5692-5693.25, 5712-5713.25, 5732-5733.25, 5752-5753.25, 5772-5773.25, 5792-5793.25, 5812-5813.25, 5832-5833.25, 5852-5853.25, 5872-5873.25, 5892-5893.25, 5912-5913.25, 5932-5933.25, 5952-5953.25, 5972-5973.25, 5992-5993.25, 6012-6013.25, 6032-6033.25, 6052-6053.25, 6072-6073.25, 6092-6093.25, 6112-6113.25, 6132-6133.25, 6152-6153.25, 6172-6173.25, 6192-6193.25, 6212-6213.25, 6232-6233.25, 6252-6253.25, 6272-6273.25, 6292-6293.25, 6312-6313.25, 6332-6333.25, 6352-6353.25, 6372-6373.25, 6392-6393.25, 6412-6413.25, 6432-6433.25, 6452-6453.25, 6472-6473.25, 6492-6493.25, 6512-6513.25, 6532-6533.25, 6552-6553.25, 6572-6573.25, 6592-6593.25, 6612-6613.25, 6632-6633.25, 6652-6653.25, 6672-6673.25, 6692-6693.25, 6712-6713.25, 6732-6733.25, 6752-6753.25, 6772-6773.25, 6792-6793.25, 6812-6813.25, 6832-6833.25, 6852-6853.25, 6872-6873.25, 6892-6893.25, 6912-6913.25, 6932-6933.25, 6952-6953.25, 6972-6973.25, 6992-6993.25, 7012-7013.25, 7032-7033.25, 7052-7053.25, 7072-7073.25, 7092-7093.25, 7112-7113.25, 7132-7133.25, 7152-7153.25, 7172-7173.25, 7192-7193.25, 7212-7213.25, 7232-7233.25, 7252-7253.25, 7272-7273.25, 7292-7293.25, 7312-7313.25, 7332-7333.25, 7352-7353.25, 7372-7373.25, 7392-7393.25, 7412-7413.25, 7432-7433.25, 7452-7453.25, 7472-7473.25, 7492-7493.25, 7512-7513.25, 7532-7533.25, 7552-7553.25, 7572-7573.25, 7592-7593.25, 7612-7613.25, 7632-7633.25, 7652-7653.25, 7672-7673.25, 7692-7693.25, 7712-7713.25, 7732-7733.25, 7752-7753.25, 7772-7773.25, 7792-7793.25, 7812-7813.25, 7832-7833.25, 7852-7853.25, 7872-7873.25, 7892-7893.25, 7912-7913.25, 7932-7933.25, 7952-7953.25, 7972-7973.25, 7992-7993.25, 8012-8013.25, 8032-8033.25, 8052-8053.25, 8072-8073.25, 8092-8093.25, 8112-8113.25, 8132-8133.25, 8152-8153.25, 8172-8173.25, 8192-8193.25, 8212-8213.25, 8232-8233.25, 8252-8253.25, 8272-8273.25, 8292-8293.25, 8312-8313.25, 8332-8333.25, 8352-8353.25, 8372-8373.25, 8392-8393.25, 8412-8413.25, 8432-8433.25, 8452-8453.25, 8472-8473.25, 8492-8493.25, 8512-8513.25, 8532-8533.25, 8552-8553.25, 8572-8573.25, 8592-8593.25, 8612-8613.25, 8632-8633.25, 8652-8653.25, 8672-8673.25, 8692-8693.25, 8712-8713.25, 8732-8733.25, 8752-8753.25, 8772-8773.25, 8792-8793.25, 8812-8813.25, 8832-8833.25, 8852-8853.25, 8872-8873.25, 8892-8893.25, 8912-8913.25, 8932-8933.25, 8952-8953.25, 8972-8973.25, 8992-8993.25, 9012-9013.25, 9032-9033.25, 9052-9053.25, 9072-9073.25, 9092-9093.25, 9112-9113.25, 9132-9133.25, 9152-9153.25, 9172-9173.25, 9192-9193.25, 9212-9213.25, 9232-9233.25, 9252-9253.25, 9272-9273.25, 9292-9293.25, 9312-9313.25, 9332-9333.25, 9352-9353.25, 9372-9373.25, 9392-9393.25, 9412-9413.25, 9432-9433.25, 9452-9453.25, 9472-9473.25, 9492-9493.25, 9512-9513.25, 9532-9533.25, 9552-9553.25, 9572-9573.25, 9592-9593.25, 9612-9613.25, 9632-9633.25, 9652-9653.25, 9672-9673.25, 9692-9693.25, 9712-9713.25, 9732-9733.25, 9752-9753.25, 9772-9773.25, 9792-9793.25, 9812-9813.25, 9832-9833.25, 9852-9853.25, 9872-9873.25, 9892-9893.25, 9912-9913.25, 9932-9933.25, 9952-9953.25, 9972-9973.25, 9992-9993.25, 10012-10013.25, 10032-10033.25, 10052-10053.25, 10072-10073.25, 10092-10093.25, 10112-10113.25, 10132-10133.25, 10152-10153.25, 10172-10173.25, 10192-10193.25, 10212-10213.25, 10232-10233.25, 10252-10253.25, 10272-10273.25, 10292-10293.25, 10312-10313.25, 10332-10333.25, 10352-10353.25, 10372-10373.25, 10392-10393.25, 10412-10413.25, 10432-10433.25, 10452-10453.25, 10472-10473.25, 10492-10493.25, 10512-10513.25, 10532-10533.25, 10552-10553.25, 10572-10573.25, 10592-10593.25, 10612-10613.25, 10632-10633.25, 10652-10653.25, 10672-10673.25, 10692-10693.25, 10712-10713.25, 10732-10733.25, 10752-10753.25, 10772-10773.25, 10792-10793.25, 10812-10813.25, 10832-10833.25, 10852-10853.25, 10872-10873.25, 10892-10893.25, 10912-10913.25, 10932-10933.25, 10952-10953.25, 10972-10973.25, 10992-10993.25, 11012-11013.25, 11032-11033.25, 11052-11053.25, 11072-11073.25, 11092-11093.25, 11112-11113.25, 11132-11133.25, 11152-11153.25, 11172-11173.25, 11192-11193.25, 11212-11213.25, 11232-11233.25, 11252-11253.25, 11272-11273.25, 11292-11293.25, 11312-11313.25, 11332-11333.25, 11352-11353.25, 11372-11373.25, 11392-11393.25, 11412-11413.25, 11432-11433.25, 11452-11453.25, 11472-11473.25, 11492-11493.25, 11512-11513.25, 11532-11533.25, 11552-11553.25, 11572-11573.25, 11592-11593.25, 11612-11613.25, 11632-11633.25, 11652-11653.25, 11672-11673.25, 11692-11693.25, 11712-11713.25, 11732-11733.25, 11752-11753.25, 11772-11773.25, 11792-11793.25, 11812-11813.25, 11832-11833.25, 11852-11853.25, 11872-11873.25, 11892-11893.25, 11912-11913.25, 11932-11933.25, 11952-11953.25, 11972-11973.25, 11992-11993.25, 12012-12013.25, 12032-12033.25, 12052-12053.25, 12072-12073.25, 12092-12093.25, 12112-12113.25, 12132-12133.25, 12152-12153.25, 12172-12173.25, 12192-12193.25, 12212-12213.25, 12232-12233.25, 12252-12253.25, 12272-12273.25, 12292-12293.25, 12312-12313.25, 12332-12333.25, 12352-12353.25, 12372-12373.25, 12392-12393.25, 12412-12413.25, 12432-12433.25, 12452-12453.25, 12472-12473.25, 12492-12493.25, 12512-12513.25, 12532-12533.25, 12552-12553.25, 12572-12573.25, 12592-12593.25, 12612-12613.25, 12632-12633.25, 12652-12653.25, 12672-12673.25, 12692-12693.25, 12712-12713.25, 12732-12733.25, 12752-12753.25, 12772-12773.25, 12792-12793.25, 12812-12813.25, 12832-12833.25, 12852-12853.25, 12872-12873.25, 12892-12893.25, 12912-12913.25, 12932-12933.25, 12952-12953.25, 12972-12973.25, 12992-12993.25, 13012-13013.25, 13032-13033.25, 13052-13053.25, 13072-13073.25, 13092-13093.25, 13112-13113.25, 13132-13133.25, 13152-13153.25, 13172-13173.25, 13192-13193.25, 13212-13213.25, 13232-13233.25, 13252-13253.25, 13272-13273.25, 13292-13293.25, 13312-13313.25, 13332-13333.25, 13352-13353.25, 13372-13373.25, 13392-13393.25, 13412-13413.25, 13432-13433.25, 13452-13453.25, 13472-13473.25, 13492-13493.25, 13512-13513.25, 13532-13533.25, 13552-13553.25, 13572-13573.25, 13592-13593.25, 13612-13613.25, 13632-13633.25, 13652-13653.25, 13672-13673.25, 13692-13693.25, 13712-13713.25, 13732-13733.25, 13752-13753.25, 13772-13773.25, 13792-13793.25, 13812-13813.25, 13832-13833.25, 13852-13853.25, 13872-13873.25, 13892-13893.25, 13912-13913.25, 13932-13933.25, 13952-13953.25, 13972-13973.25, 13992-13993.25, 14012-14013.25, 14032-14033.25, 14052-14053.25, 14072-14073.25, 14092-14093.25, 14112-14113.25, 14132-14133.25, 14152-14153.25, 14172-14173.25, 14192-14193.25, 14212-14213.25, 14232-14233.25, 14252-14253.25, 14272-14273.25, 14292-14293.25, 14312-14313.25, 14332-14333.25, 14352-14353.25, 14372-14373.25, 14392-14393.25, 14412-14413.25, 14432-14433.25, 14452-14453.25, 14472-14473.25, 14492-14493.25, 14512-14513.25, 14532-14533.25, 14552-14553.25, 14572-14573.25, 14592-14593.25, 14612-14613.25, 14632-14633.25, 14652-14653.25, 14672-14673.25, 14692-14693.25, 14712-14713.25, 14732-14733.25, 14752-14753.25, 14772-14773.25, 14792-14793.25, 14812-14813.25, 14832-14833.25, 14852-14853.25, 14872-14873.25, 14892-14893.25, 14912-14913.25, 14932-14933.25, 14952-14953.25, 14972-14973.25, 14992-14993.25, 15012-15013.25, 15032-15033.25, 15052-15053.25, 15072-15073.25, 15092-15093.25, 15112-15113.25, 15132-15133.25, 15152-15153.25, 15172-15173.25, 15192-15193.25, 15212-15213.25, 15232-15233.25, 15252-15253.25, 15272-15273.25, 15292-15293.25, 15312-15313.25, 15332-15333.25, 15352-15353.25, 15372-15373.25, 15392-15393.25, 15412-15413.25, 15432-15433.25, 15452-15453.25, 15472-15473.25, 15492-15493.25, 15512-15513.25, 15532-15533.25, 15552-15553.25, 15572-15573.25, 15592-15593.25, 15612-15613.25, 15632-15633.25, 15652-15653.25, 15672-15673.25, 15692-15693.25, 15712-15713.25, 15732-15733.25, 15752-15753.25, 15772-15773.25, 15792-15793.25, 15812-15813.25, 15832-15833.25, 15852-15853.25, 15872-15873.25, 15892-15893.25, 15912-15913.25, 15932-15933.25, 15952-15953.25, 15972-15973.25, 15992-15993.25, 16012-16013.25, 16032-16033.25, 16052-16053.25, 16072-16073.25, 16092-16093.25, 16112-16113.25, 16132-16133.25, 16152-16153.25, 16172-16173.25, 16192-16193.25, 16212-16213.25, 16232-16233.25, 16252-16253.25, 16272-16273.25, 16292-16293.25, 16312-16313.25, 16332-16333.25, 16352-16353.25, 16372-16373.25, 16392-16393.25, 16412-16413.25, 16432-16433.25, 16452-16453.25, 16472-16473.25, 16492-16493.25, 16512-16513.25, 16532-16533.25, 16552-16553.25, 16572-16573.25, 16592-16593.25, 16612-16613.25, 16632-16633.25, 16652-16653.25, 16672-16673.25, 16692-16693.25, 16712-16713.25, 16732-16733.25, 16752-16753.25, 16772-16773.25, 16792-16793.25, 16812-16813.25, 16832-16833.25, 16852-16853.25, 16872-16873.25, 16892-16893.25, 16912-16913.25, 16932-16933.25, 16952-16953.25, 16972-16973.25, 16992-16993.25, 17012-17013.25, 17032-17033.25, 17052-17053.25, 17072-17073.25, 17092-17093.25, 17112-17113.25, 17132-17133.25,

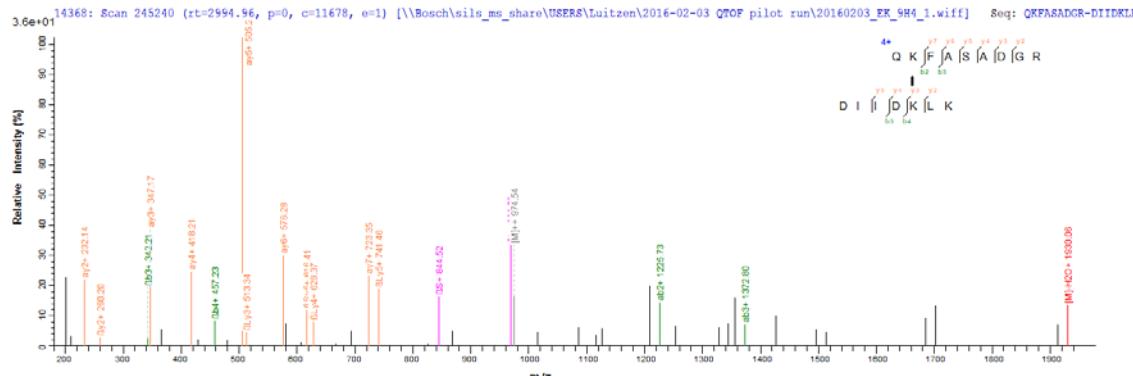
673

674

Cross-link between RL11\_BACSU and RL7\_BACSU; matched intensity 70.1%

676

g

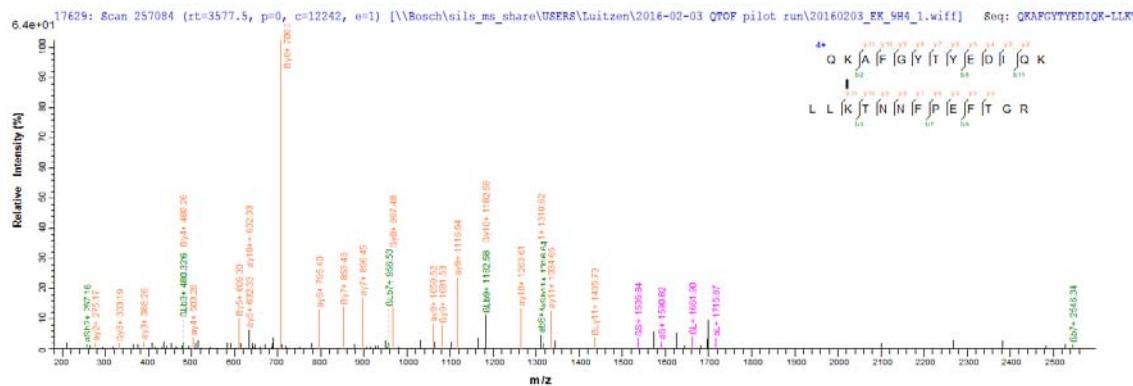


677

678 Cross-link between RL31\_BACSU and RS13\_BACSU; matched intensity 65.2%

679

h

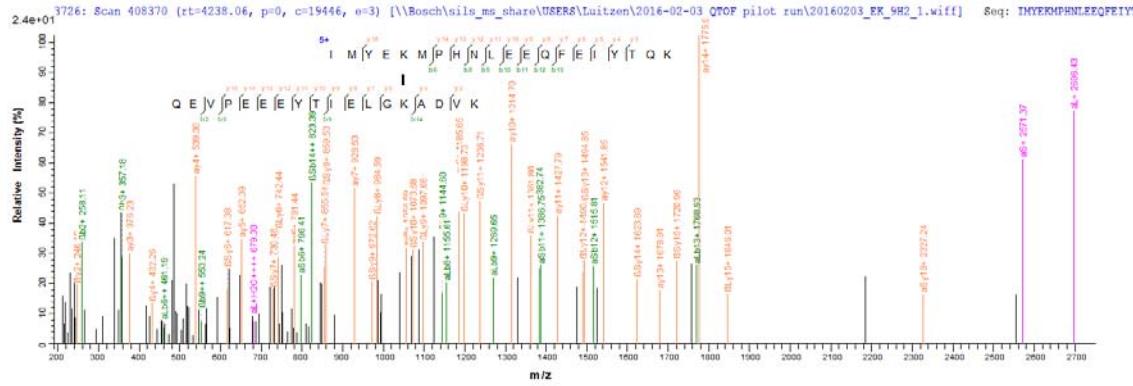


680

681 Cross-link between GLTA\_BACSU and GLTB\_BACSU; matched intensity 74.5%

682

i

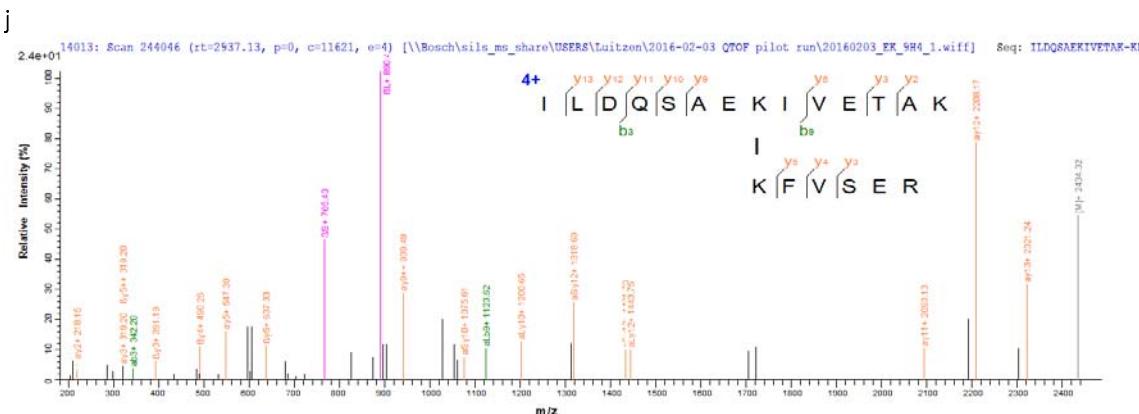


683

684 Cross-link between ODPA\_BACSU and ODPB\_BACSU; matched intensity 58.1%

685

686



687

688 Cross-link between RS10\_BACSU and RS18\_BACSU; matched intensity 67.2%

689

690 Supplementary Figure 3. Mass spectra of cross-linked peptide pairs revealing new protein-protein  
 691 interaction a, AMPA\_BACSU-RL7\_BACSU; b, EFTU\_BACSU-ABRB\_BACSU; c, YOPJ\_BACSU-  
 692 RPOA\_BACSU; d, FUR\_BACSU-YLAN\_BACSU; see also Table 4.

693 a

694

695

696

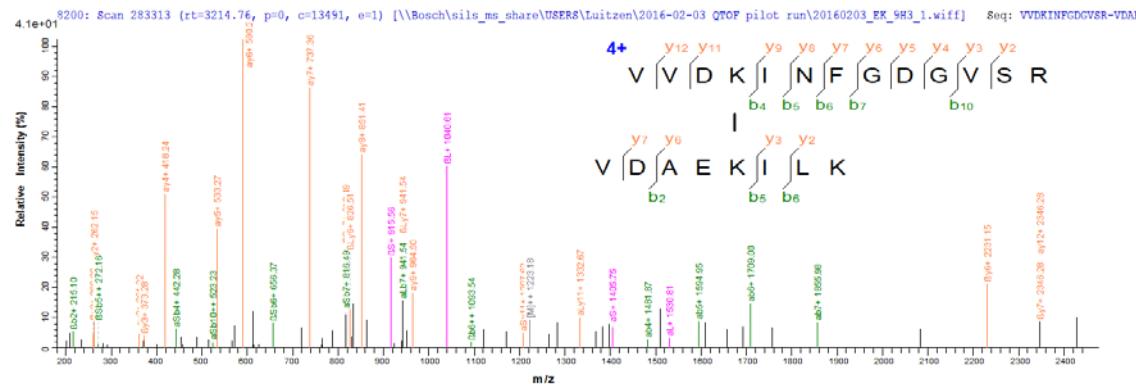
697

Mass spectrum and fragmentation diagram for the peptide FAKDIELSLEVK at 4+ charge. The x-axis is  $m/z$  from 200 to 1900, and the y-axis is Relative Intensity (%) from 0 to 2.0e+01. The base peak is at  $m/z$  597.29. The fragmentation diagram shows the peptide backbone with various cleavage sites and resulting ions labeled with y- and z- ions.

698

699

700 d



701

702

703