

Supporting Information

The First Community-Wide, Comparative Cross-linking Mass Spectrometry Study

Claudio Iacobucci^{1#}, Christine Piotrowski^{1#}, Ruedi Aebersold^{2,3}, Bruno C. Amaral⁴, Philip Andrews⁵, Katja Bernfur¹⁷, Christoph Borchers^{6,7,8,9}, Nicolas I. Brodie⁶, James E. Bruce¹⁰, Yong Cao¹⁵, Stéphane Chaignepain¹¹, Juan D. Chavez¹⁰, Stéphane Claverol¹², Jürgen Cox¹³, Trisha Davis⁴¹, Gianluca Degliesposti¹⁴, Meng-Qiu Dong¹⁵, Nufar Edinger¹⁶, Cecilia Emanuelsson¹⁷, Marina Gay¹⁸, Michael Götze¹⁹, Francisco Gomes-Neto⁴⁴, Fabio C. Gozzo⁴, Craig Gutierrez²⁰, Caroline Haupt²¹, Albert J. R. Heck²², Franz Herzog²³, Lan Huang²⁰, Michael R. Hoopmann²⁴, Nir Kalisman¹⁶, Oleg Klykov²², Zdeněk Kukačka²⁵, Fan Liu²⁶, Michael J. MacCoss²⁷, Karl Mechtler²⁸, Ravit Mesika¹⁶, Robert L. Moritz²⁴, Nagarjuna Nagaraj²⁹, Victor Nesati³⁰, Ana G. C. Neves-Ferreira⁴⁴, Robert Ninnis³⁰, Petr Novák²⁵, Francis J O'Reilly³¹, Matthias Pelzing³⁰, Evgeniy Petrotchenko⁶, Lolita Piersimoni⁵, Manolo Plasencia⁵, Tara Pukala³², Kasper D. Rand³³, Juri Rappsilber^{31,34}, Dana Reichmann¹⁶, Carolin Sailer³⁵, Chris P. Sarnowski^{2,36}, Richard A. Scheltema²², Carla Schmidt²¹, David C. Schriemer³⁷, Yi Shi⁴⁰, J. Mark Skehel¹⁴, Moriya Slavin¹⁶, Frank Sobott^{42,43}, Victor Solis-Mezarino²³, Heike Stephanowitz²⁶, Florian Stengel³⁵, Christian E. Stieger²⁸, Esben Trabjerg³³, Michael Trnka³⁸, Marta Vilaseca¹⁸, Rosa Viner³⁹, Yufei Xiang⁴⁰, Sule Yilmaz¹³, Alex Zelter⁴¹, Daniel Ziemianowicz³⁷, Alexander Leitner^{2*}, Andrea Sinz^{1*}

¹ Department of Pharmaceutical Chemistry and Bioanalytics, Institute of Pharmacy, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany.

² Department of Biology, Institute of Molecular Systems Biology, ETH Zurich, Otto-Stern-Weg 3, 8093 Zurich, Switzerland.

³ Faculty of Science, University of Zurich, 8006 Zurich, Switzerland.

⁴ Institute of Chemistry, University of Campinas, Campinas SP, 13083-970, Brazil.

⁵ Departments of Biological Chemistry, Bioinformatics, and Chemistry, University of Michigan, Ann Arbor Michigan, MI 48109, USA.

⁶ University of Victoria -Genome British Columbia Proteomics Centre, Vancouver Island Technology Park, Victoria, British Columbia V8Z 7X8, Canada.

⁷ Department of Biochemistry and Microbiology, University of Victoria, Petch Building, Room 270d, 3800 Finnerty Road, Victoria, British Columbia V8P 5C2, Canada.

⁸ Gerald Bronfman Department of Oncology, Jewish General Hospital, McGill University, 3755 Côte Ste-Catherine Road, Montréal, Quebec H3T 1E2, Canada.

⁹ Proteomics Centre, Segal Cancer Centre, Lady Davis Institute, Jewish General Hospital, McGill University, 3755 Côte Ste-Catherine Road, Montréal, Quebec H3T 1E2, Canada.

¹⁰ Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA.

¹¹ CBMN, UMR 5248, CNRS, Université de Bordeaux, INP Bordeaux, Pessac 33607, France.

¹² Centre de Génomique Fonctionnelle, Plateforme Protéome, Université de Bordeaux, Bordeaux33000, France.

- ¹³ Computational Systems Biochemistry Research Group, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany.
- ¹⁴ MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Ave, Cambridge CB2 0QH, UK.
- ¹⁵ National Institute of Biological Sciences, Beijing, 7 Science Park Road, ZGC Life Science Park, 102206 Beijing, China.
- ¹⁶ Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, Safra Campus Givat Ram, The Hebrew University of Jerusalem, Jerusalem 91904, Israel.
- ¹⁷ Department of Biochemistry and Structural Biology, Center for Molecular Protein Science, Lund University, 221 00 Lund, Sweden.
- ¹⁸ Institute for Research in Biomedicine (IRB Barcelona), The Barcelona Institute of Science and Technology (BIST), Baldiri Reixac 10, 08028 Barcelona, Spain.
- ¹⁹ Institute for Biochemistry and Biotechnology, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany.
- ²⁰ Department of Physiology & Biophysics, University of California, Irvine, CA 92697, USA.
- ²¹ Interdisciplinary Research Center HALOmem, Institute for Biochemistry and Biotechnology, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, 06120 Halle/Saale, Germany.
- ²² Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, University of Utrecht and Netherlands Proteomics Centre, Padualaan 8, 3584 CH Utrecht, The Netherlands.
- ²³ Gene Center Munich, Department of Biochemistry, Faculty of Chemistry and Pharmacy, Ludwig Maximilians University of Munich, Feodor-Lynen-Str. 25, 81377 Munich, Germany.
- ²⁴ Institute for Systems Biology, 401 Terry Ave N, Seattle, WA 98109, USA.
- ²⁵ Institute of Microbiology, BIOCEV, Prumyslova 595, 252 50 Vestec, Czech Republic.
- ²⁶ Leibniz Institute of Molecular Pharmacology (FMP), Robert-Rössle-Str. 10, 13125 Berlin, Germany.
- ²⁷ Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA.
- ²⁸ Protein Chemistry Facility, Research Institute of Molecular Pathology (IMP) and Institute of Molecular Biotechnology (IMBA), Vienna Biocenter (VBC), Dr. Bohr-Gasse 3, 1030 Vienna, Austria.
- ²⁹ Biochemistry Core Facility, Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany;
- ³⁰ Analytical Biochemistry, CSL Limited, Bio21 Institute, 30 Flemington Road, 3010 Parkville, Australia.
- ³¹ Chair of Bioanalytics, Institute of Biotechnology Technische Universität Berlin, 13355 Berlin, Germany.
- ³² Discipline of Chemistry, Faculty of Sciences, University of Adelaide, North Terrace, Adelaide, South Australia, 5005, Australia.
- ³³ Department of Pharmacy, University of Copenhagen, 2100 Copenhagen, Denmark.
- ³⁴ Wellcome Trust Centre for Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.
- ³⁵ University of Konstanz, Department of Biology, Universitätsstr. 10, 78457 Konstanz, Germany.
- ³⁶ PhD Program in Systems Biology, University of Zurich and ETH Zurich, 8092 Zurich, Switzerland.
- ³⁷ Dept. of Biochemistry & Molecular Biology, Robson DNA Science Centre, University of Calgary, 3330 Hospital Dr. NW, Calgary, Alberta T2N 4N1, Canada.
- ³⁸ UCSF Mass Spectrometry Facility, Genentech Hall, 600 16th Street, San Francisco, CA 94158, USA.
- ³⁹ Thermo Fisher Scientific, 355 River Oaks Parkway, San Jose, CA 95134, USA.

⁴⁰ Department of Cell Biology, University of Pittsburgh, School of Medicine, Pittsburgh, PA 15213, USA.

⁴¹ Department of Biochemistry, University of Washington, Seattle, WA 98195, USA.

⁴² Department of Chemistry, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium.

⁴³ The Astbury Centre for Structural Molecular Biology, and ∇ School of Molecular and Cellular Biology, University of Leeds, LS2 9JT Leeds, UK.

⁴⁴ Laboratory of Toxinology, Oswaldo Cruz Institute, Fiocruz, Avenida Brasil 4.365, Manguinhos, Rio de Janeiro - RJ, CEP 21040-900 (Moorish Castle), Brazil.

Corresponding Authors

* Andrea Sinz, Department of Pharmaceutical Chemistry & Bioanalytics, Institute of Pharmacy, Charles Tanford Protein Center, Martin Luther University Halle-Wittenberg, Kurt-Mothes-Str. 3a, D-06120 Halle (Saale), Germany.

Email address: andrea.sinz@pharmazie.uni-halle.de

* Alexander Leitner, Institute of Molecular Systems Biology, Otto-Stern-Weg 3, ETH Zurich, CH-8093 Zurich, Switzerland.

Email address: leitner@imsb.biol.ethz.ch

Both authors contributed equally.

Table of Contents

Supplementary Scheme

Page S-5

Scheme S1: General workflow of cross-linking mass spectrometry. Adapted from Götze *et al.*, 2019, bioRxiv preprint, <https://doi.org/10.1101/524314>.

Supplementary Figures

Pages S-6 – S-10

Figure S1: Details on (A) enrichment of cross-linked species, (B) considered charge states, (C) fragmentation methods, and (D) MS³ resolution are presented; normalized collision energies (NCE) are given in % for HCD and CID. For (B) and (C); the number of experiments are given as y-axes. As cross-linked peptides carry higher charge states, usually charge states >2 are considered for analysis. SEC: Size exclusion chromatography; SCX: strong cation exchange.

Figure S2: Influence of cross-linking sites considered in data analysis.

Figure S3: Cross-links with homobifunctional amine-reactive reagents, found in the monomer band of BSA using *in-gel* digestion (in total 10 datasets), were mapped into the published 3D structure of BSA (pdb entry 4F5S). Only cross-links were considered that were identified in at least two independent experiments.

Figure S4: Cross-links with homobifunctional, amine-reactive reagents, identified in the monomer band of BSA using *in-gel* digestion (in total 10 datasets), were mapped into the published 3D-structure of BSA (pdb entry 4F5S); a) distribution of C α -C α distances of the identified cross-links as a function of their reproducibility across the datasets. 30 Å was set as maximum C α -C α distance; b) percentage of overlength cross-links (> 30 Å) as a function of cross-link reproducibility.

Supplementary Tables

Pages S-11 – S-19

Table S1: List of cross-linking reagents used in this study.

Table S2: List of cross-linking software used in this study.

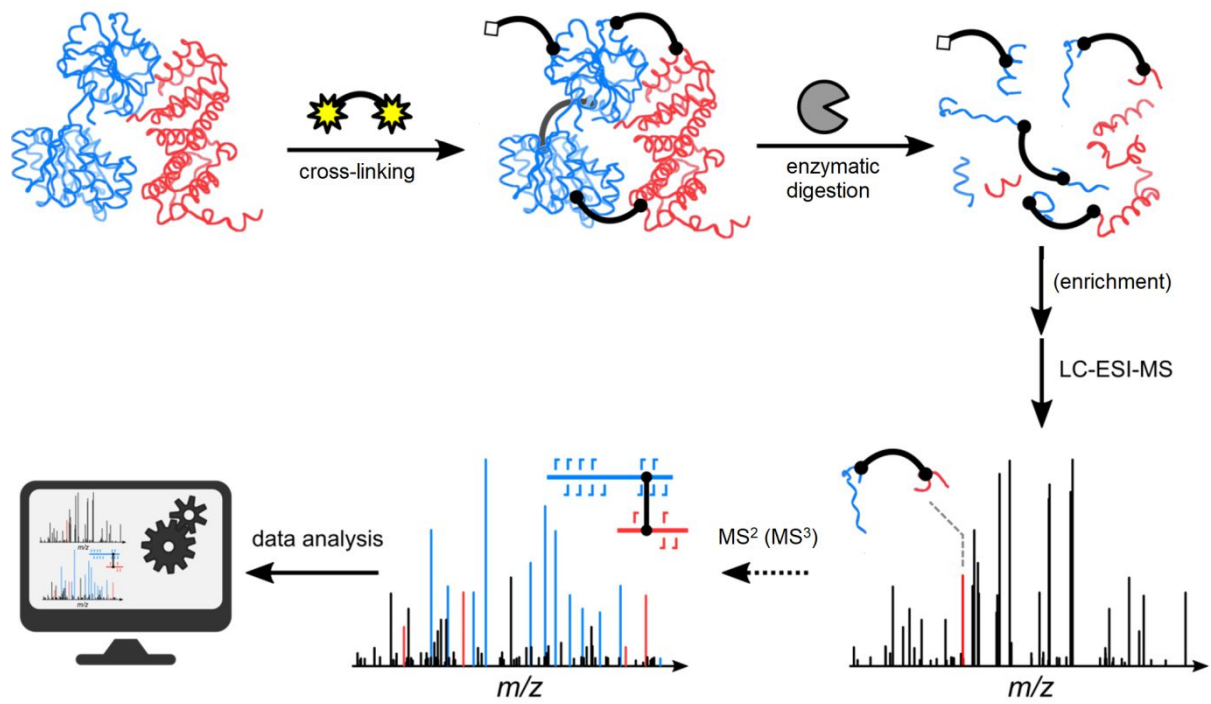
Table S3: List of unique cross-links identified with homobifunctional, amine-reactive reagents after *in-gel* digestion of the BSA monomer band. Cross-links are sorted by the number of times they were identified across the 10 relevant datasets; C α -C α distances are reported in Å.

Table S4: Unique cross-links identified at different BSA concentrations. The cross-linking reactions were conducted in duplicates (500, 250, 50, 25 μ M DSBU, 25 mM HEPES, pH 7.2, incubation time 1 hour, 21 °C).

Supplementary Text

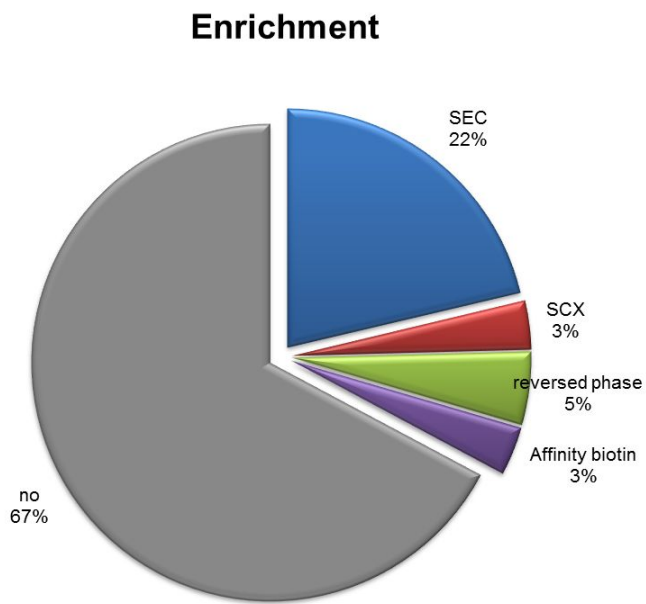
Pages S-19 – S-24

Supplementary Figures

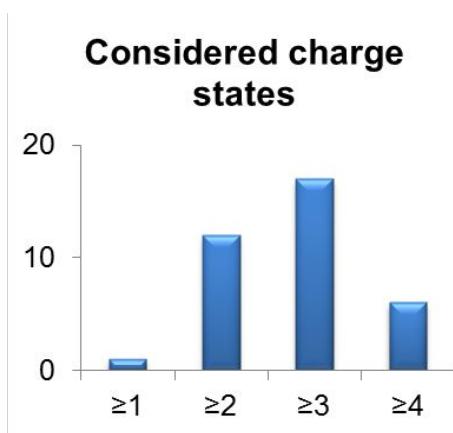


Scheme S1: General workflow of cross-linking mass spectrometry. Adapted from Götze *et al.*, 2019, bioRxiv preprint, <https://doi.org/10.1101/524314>.

A



B



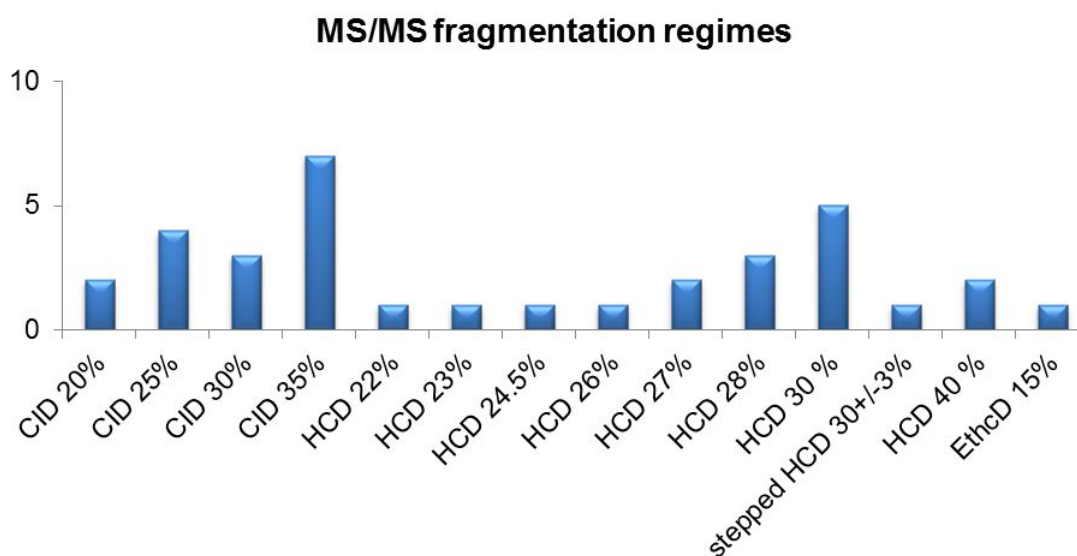
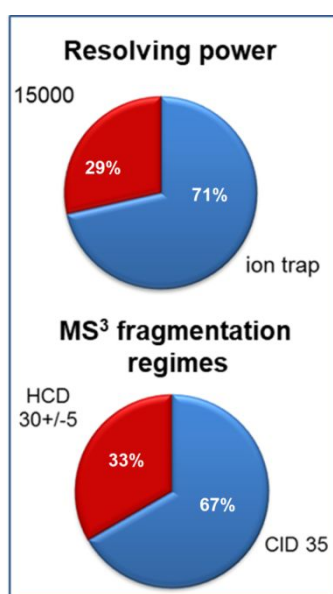
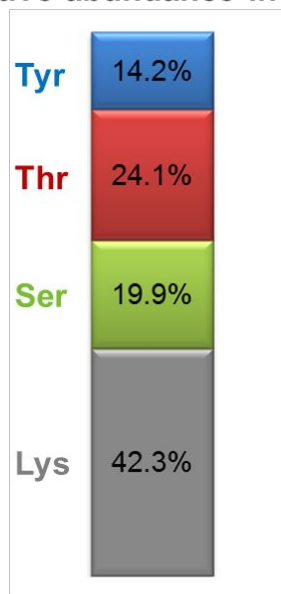
C**D**

Figure S1: Details on (A) enrichment of cross-linked species, (B) considered charge states, (C) fragmentation methods, and (D) MS³ resolution are presented; normalized collision energies (NCE) are given in % for HCD and CID. For (B) and (C); the number of experiments are given as y-axes. As cross-linked peptides carry higher charge states, usually charge states >2 are considered for analysis. SEC: Size exclusion chromatography; SCX: strong cation exchange.

**Lys, Ser, Thr, and Tyr
relative abundance in BSA**



**Relative abundance of targeted
residues in cross-links**

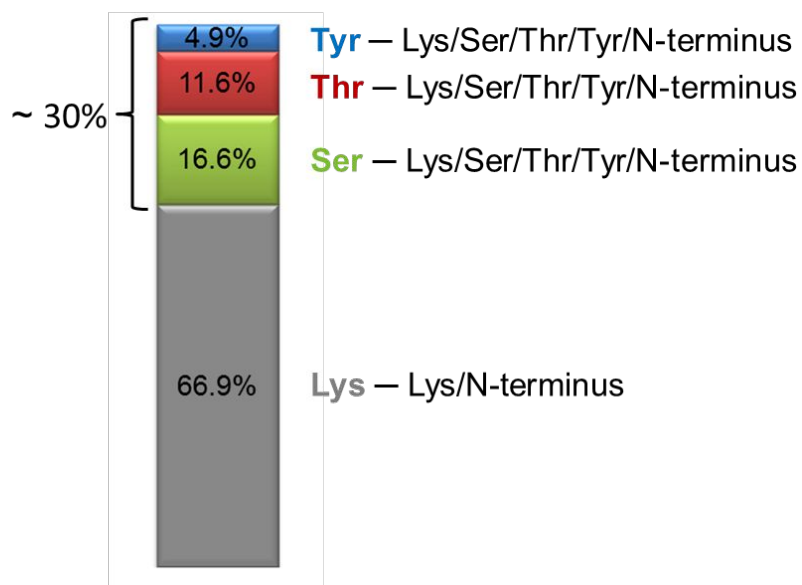


Figure S2: Influence of cross-linking sites considered in data analysis.

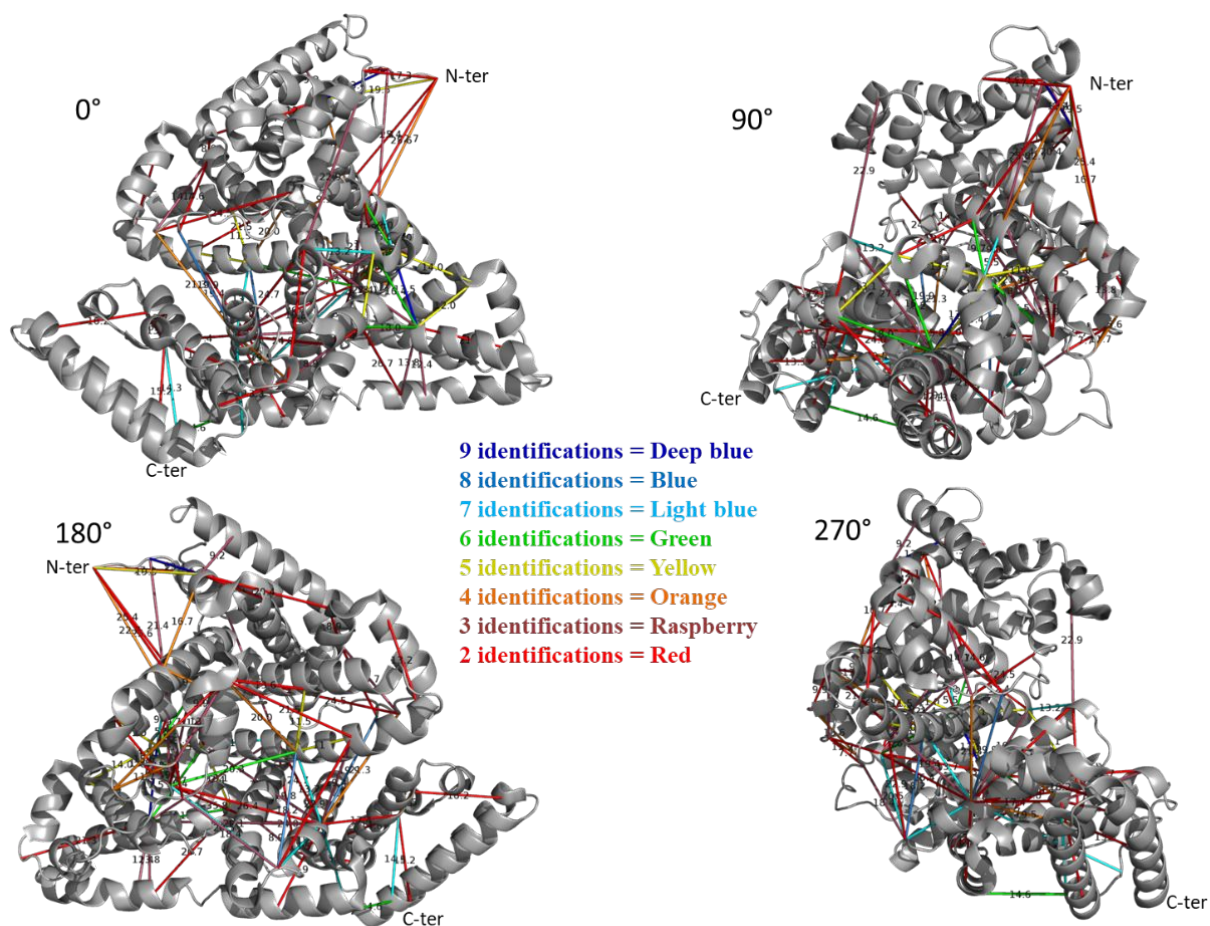


Figure S3: Cross-links with homobifunctional amine-reactive reagents, found in the monomer band of BSA using *in-gel* digestion (in total 10 datasets), were mapped into the published 3D structure of BSA (pdb entry 4F5S). Only cross-links were considered that were identified in at least two independent experiments.

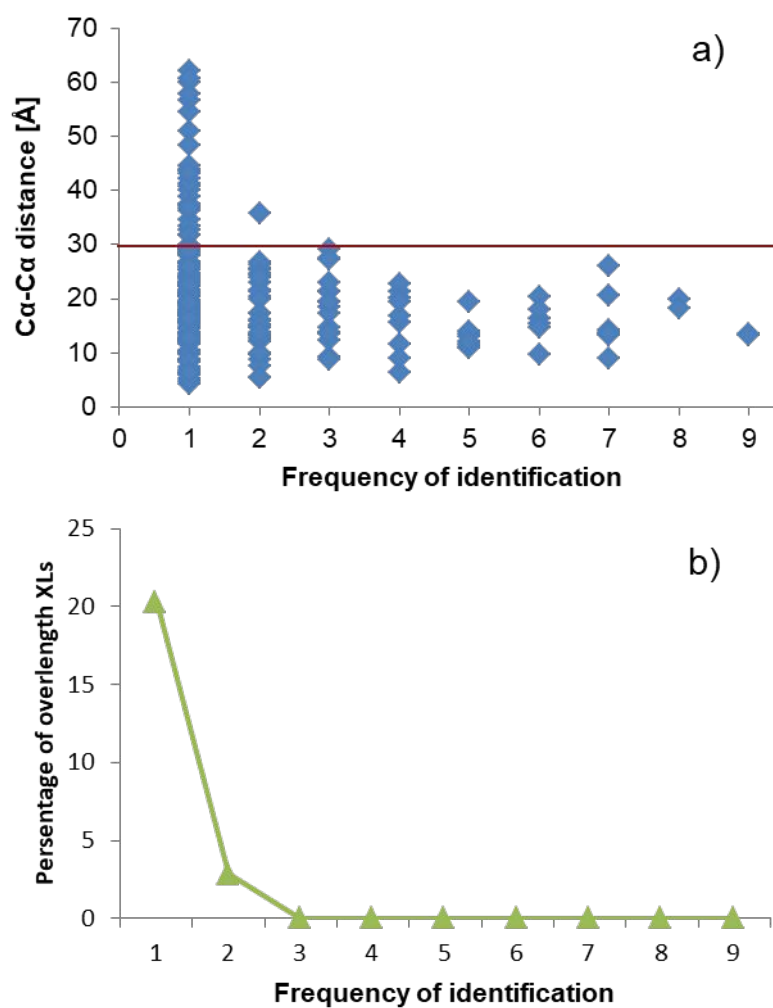
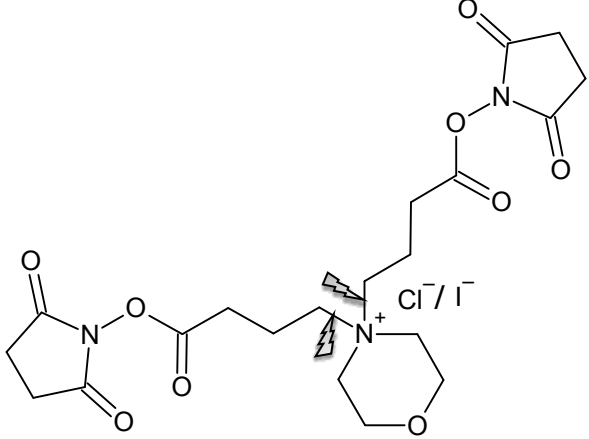
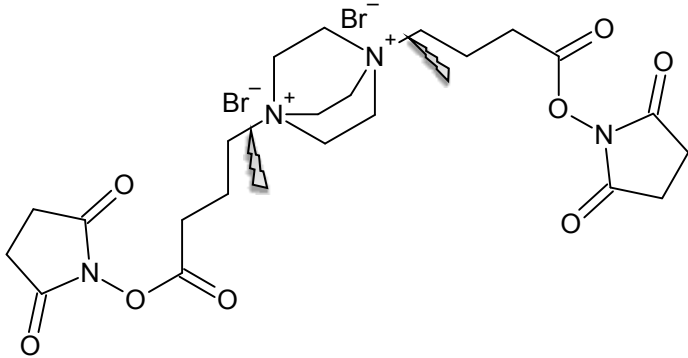
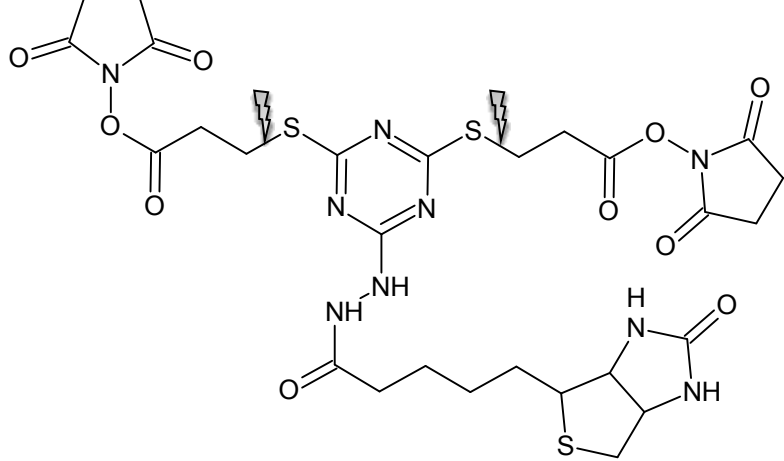
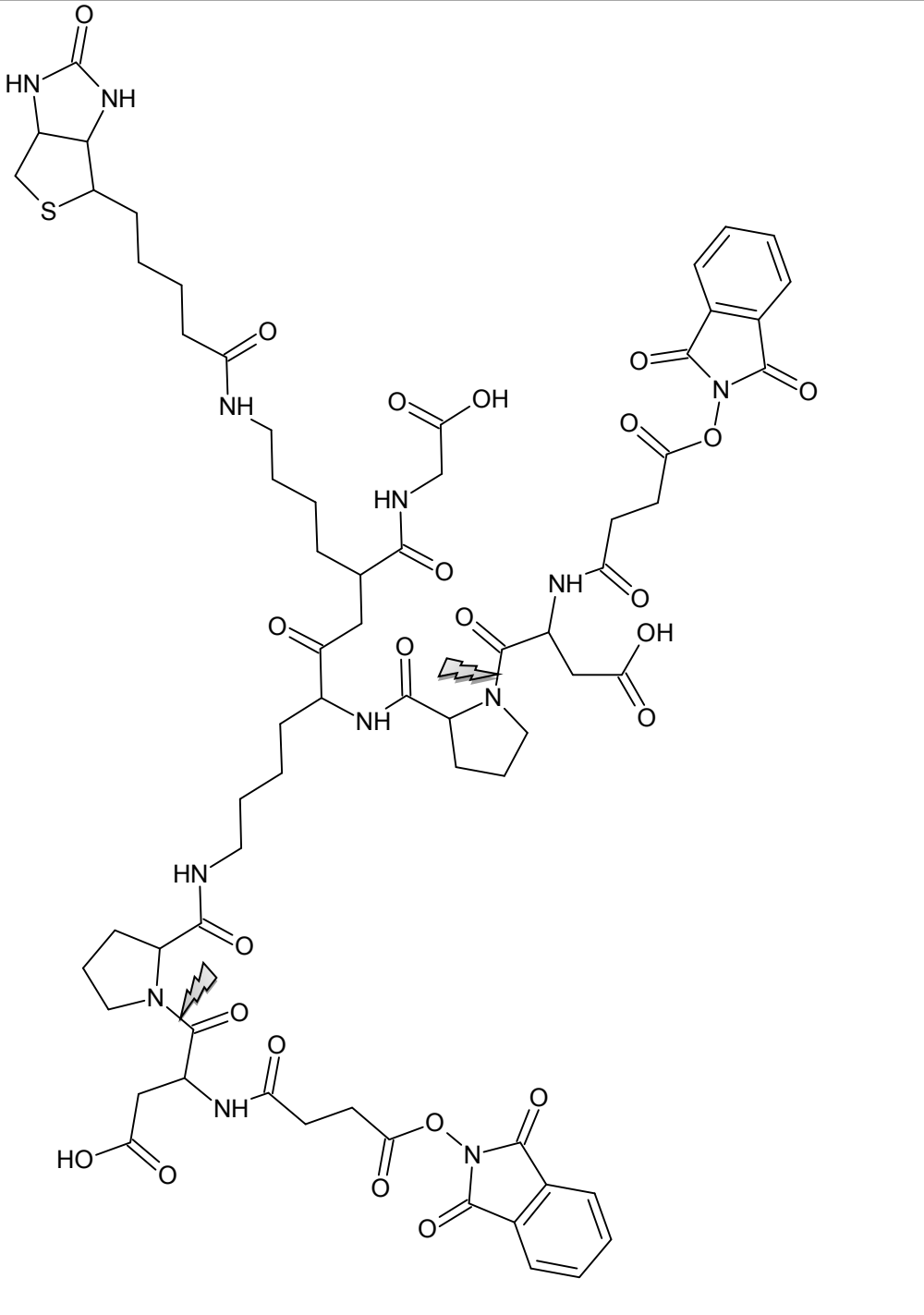
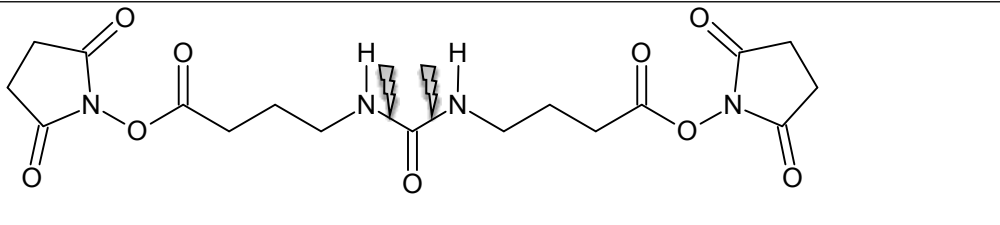
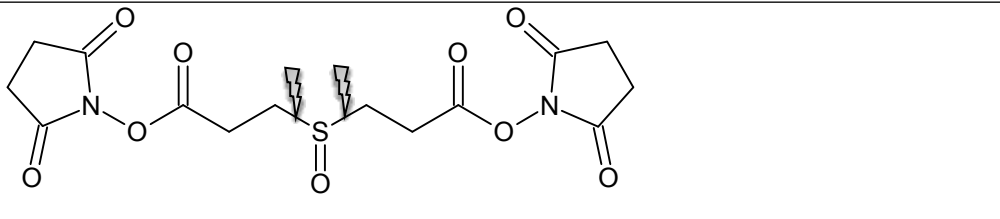
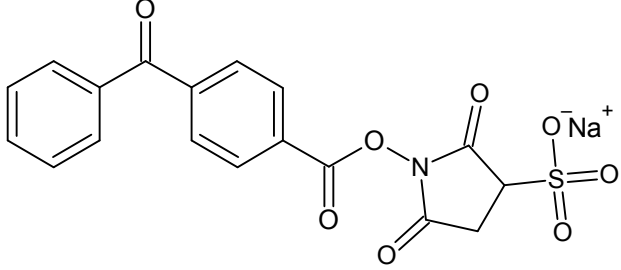
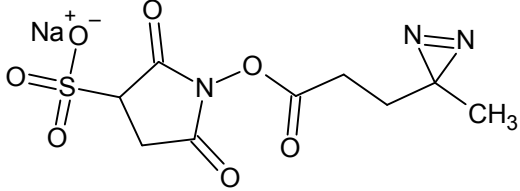
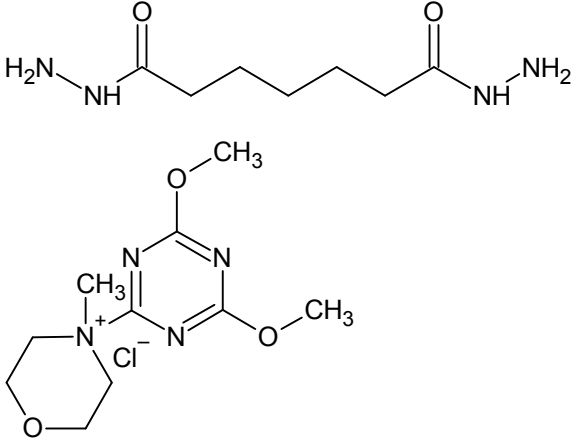
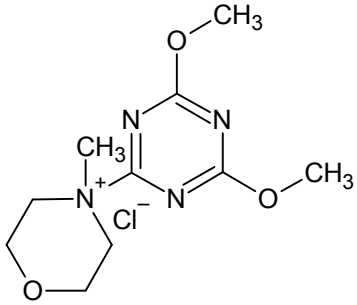
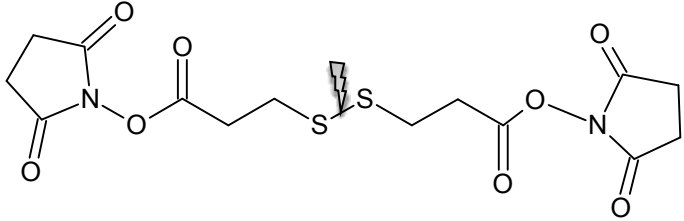
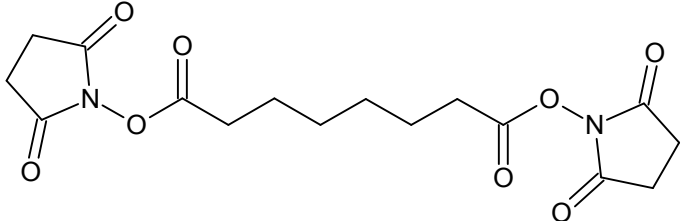


Figure S4: Cross-links with homobifunctional, amine-reactive reagents, identified in the monomer band of BSA using in-gel digestion (in total 10 datasets), were mapped into the published 3D-structure of BSA (pdb entry 4F5S); a) distribution of C α -C α distances of the identified cross-links as a function of their reproducibility across the datasets. 30 Å was set as maximum C α -C α distance; b) percentage of overlength cross-links (> 30 Å) as a function of cross-link reproducibility.

Table S1: List of cross-linking reagents used in this study.

Cross-linker	Structure
MC4	 <p>The structure of MC4 consists of a central nitrogen atom with a positive charge, coordinated to a chloride or iodide ion (Cl⁻ / I⁻). This central nitrogen is also bonded to a morpholine ring and two propyl chains. Each propyl chain is terminated with a succinimidyl ester group, which is used for cross-linking.</p>
DC4	 <p>The structure of DC4 features a central nitrogen atom with a positive charge, coordinated to two bromide ions (Br⁻). The central nitrogen is also bonded to two propyl chains, each ending in a succinimidyl ester group. Additionally, the central nitrogen is bonded to a bicyclic diamine core.</p>
CBDPS	 <p>The structure of CBDPS is a complex molecule. It features a central benzothiazine core. Two propyl chains are attached to the benzothiazine core via sulfur atoms, each terminating in a succinimidyl ester group. The benzothiazine core also has a hydrazide group (-NH-NH-) and a thiazolidine ring system.</p>

BDP-NHP	 <p>The structure of BDP-NHP is a complex molecule. It features a central chain of amino acid residues: Lysine, Aspartic acid, Glutamic acid, and another Lysine. The first Lysine is modified with a BDP (benzyl-DL-glutamate) group. The second Lysine is modified with a BOP (benzotriazol-1-yloxytris(dimethylamino)phosphonium hexafluorophosphate) group. The Aspartic acid and Glutamic acid residues are also modified with BOP groups. The third Lysine is modified with a BOP group. The fourth Lysine is modified with a BOP group. The structure is highly branched and contains multiple amide and ester linkages.</p>
DSBU	 <p>The structure of DSBU (diisobutylsuccinimide) is a succinimide ring (a five-membered ring with one nitrogen and two carbonyl groups) attached to a butyl chain. The butyl chain is further attached to a nitrogen atom, which is part of a diisobutylsuccinimide group. The structure is shown as a dimeric molecule with two such units linked together.</p>
DSSO	 <p>The structure of DSSO (diisobutylsuccinimide sulfonamide) is a succinimide ring attached to a butyl chain. The butyl chain is further attached to a sulfur atom, which is part of a diisobutylsuccinimide sulfonamide group. The structure is shown as a dimeric molecule with two such units linked together.</p>

CBSS	
Sulfo-SDA	
PDH/ DMTMM	
DMTMM	
DSP	
DSS	

BS3

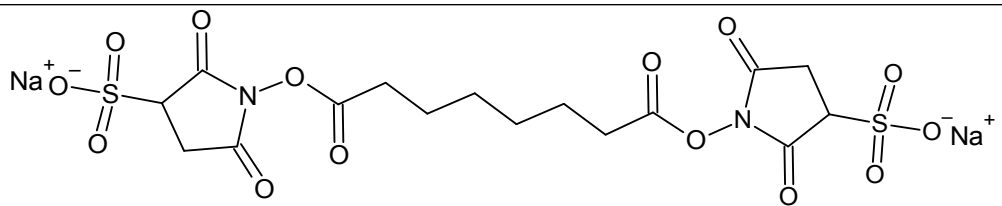


Table S2: List of cross-linking software used in this study.

Software	Reference	Website
StavroX	Götze M, Pettelkau J, Schaks S, Bosse K, Ihling CH, Krauth F, Fritzsche R, Kühn U, Sinz A (2012b) StavroX-A software for analyzing crosslinked products in protein interaction studies. <i>J Am Soc Mass Spectrom</i> 23, 76–87	www.stavrox.com
pLink	Yang B, Wu YJ, Zhu M, Fan SB, Lin J, Zhang K, Li S, Chi H, Li YX, Chen HF, Luo SK, Ding YH, Wang LH, Hao Z, Xiu LY, Chen S, Ye K, He SM, Dong MQ (2012) Identification of cross-linked peptides from complex samples. <i>Nat Methods</i> 9, 904–906	http://pfind.ict.ac.cn/software/pLink1/index.html
MeroX	Götze M, Pettelkau J, Fritzsche R, Ihling CH, Schafer M, Sinz A (2015) Automated assignment of MS/MS cleavable cross-links in protein 3D-structure analysis. <i>J Am Soc Mass Spectrom</i> 26, 83–97	www.stavrox.com
Proteome Discoverer	Liu F, Rijkers DT, Post H, Heck AJ (2015) Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. <i>Nat Methods</i> 12, 1179–1184	www.thermofisher.com
xQuest xProphet	Rinner O, Seebacher J, Walzthoeni T, Mueller LN, Beck M, Schmidt A, Mueller M, Aebersold R (2008) Identification of cross-linked peptides from large sequence databases. <i>Nat Methods</i> 5, 315–318	http://proteomics.ethz.ch/cgi-bin/xquest2.cgi/index.cgi
Kojak	Hoopmann MR, Zelter A, Johnson RS, Riffle M, MacCoss MJ, Davis TN, Moritz RL (2015) Kojak: efficient analysis of chemically cross-linked protein complexes. <i>J Proteome Res</i> 14, 2190–2198	www.kojak-ms.org
Protein Prospector	Chu F, Baker PR, Burlingame AL, Chalkley RJ (2010) Finding chimeras: a bioinformatics strategy for identification of cross-linked peptides. <i>Mol Cell Proteomics</i> 9, 25–31	http://prospector.ucsf.edu
Comet	Eng JK, Jahan TA, Hoopmann MR (2013) Comet: an open-source MS/MS sequence database search tool. <i>Proteomics</i> 13, 22–24	http://comet-ms.sourceforge.net
XlinX	Liu F, Rijkers DT, Post H, Heck AJ (2015) Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. <i>Nat Methods</i> 12, 1179–1184	www.hecklab.com/software/xlinkx
MaxQuant	Cox J, Mann M (2018) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. <i>Nat Biotechnol</i> 26, 1367–1372	www.maxquant.org
XiSearch	Fischer L, Chen ZA, Rappsilber J (2013) Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. <i>J Proteomics</i> 88, 120–128	https://github.com/Rappsilber-Laboratory/XiSearch
Find_XL/pXL	Slavin M, Kalisman N (2018) Structural Analysis of Protein Complexes by Cross-Linking and Mass Spectrometry. In: Marsh J. (eds) <i>Protein Complex Assembly. Methods in Molecular Biology</i> , 1764, Humana Press, New York, NY	http://biolchem.huji.ac.il/nirka/software.html
DXMSMS Match	Petrotchenko EV, Makepeace KA, Borchers CH (2014) DXMSMS match program for automated analysis of LC-MS/MS data obtained using isotopically coded CID-cleavable cross-linking reagents. <i>Curr Prot Bioinform</i> 48:8.18, 11–19	http://www.creativemolecules.com/cm_software.htm
SIM-XL	Lima DB, de Lima TB, Balbuena TS, Neves-Ferreira AGC, Barbosa VC, Gozzo FC, Carvalho PC (2015) SIM-XL: a powerful and user-friendly tool for peptide cross-linking analysis. <i>J Proteomics</i> 129, 51–55	http://patternlabforproteomics.org/sim-xl/
MassSpec Studio	Sarpe V, Rafiei A, Hepburn M, Ostan N, Schryvers AB, Schriemer DC (2016) High sensitivity crosslink detection coupled with integrative structure modeling in the mass spec studio. <i>Mol Cell Proteomics</i> 15, 3071–3080	www.msstudio.ca
MassAI	Rasmussen MI, Refsgaard JC, Peng L, Houen G, Hojrup P (2011) CrossWork: software-assisted identification of cross-linked peptides. <i>J Proteomics</i> 74, 1871–1883	www.massai.dk

Table S3: List of unique cross-links identified with homobifunctional, amine-reactive reagents after in-gel digestion of the BSA monomer band. Cross-links are sorted by the number of times they were identified across the 10 relevant datasets; C α -C α distances are reported in Å.

Unique cross-links identified by in-gel digestion of BSA monomer band				Unique cross-links identified by in-gel digestion of BSA monomer band			
Site 1	Site 2	n° of identifications	C α -C α distance	Site 1	Site 2	n° of identifications	C α -C α distance
211	350	9	13.537	350	431	3	29.145
4	12	9	13.378	350	465	3	27.385
114	431	8	19.903	431	524	3	17.365
187	439	8	18.216	504	524	3	13.513
524	544	7	14.335	93	465	3	22.889
187	431	7	13.69	180	431	3	19.399
204	465	7	13.237	204	431	3	27.154
211	239	7	8.958	413	471	3	8.862
221	439	7	20.539	1	64	2	17.328
431	537	7	26.11	1	242	2	25.613
431	439	7	13.513	4	64	2	9.618
187	221	6	20.304	12	20	2	12.094
204	350	6	16.426	12	131	2	20.422
211	221	6	15.451	106	116	2	24.484
211	242	6	9.735	114	136	2	14.647
350	474	6	17.992	127	173	2	13.156
396	544	6	14.584	131	136	2	8.856
1	12	5	19.543	159	280	2	13.552
159	187	5	11.533	180	280	2	21.469
180	187	5	10.97	180	439	2	22.889
204	211	5	13.472	204	242	2	12.437
204	474	5	13.142	221	232	2	14.856
211	232	5	11.021	221	273	2	9.71
211	322	5	13.993	221	471	2	35.831
322	350	5	11.985	273	294	2	7.675
116	431	4	21.274	275	280	2	9.923
224	273	4	11.658	312	350	2	21.318
224	275	4	15.587	377	474	2	26.673
520	524	4	6.418	388	431	2	15.871
1	239	4	22.716	413	465	2	19.797
12	261	4	16.738	413	535	2	13.277
187	280	4	20.037	431	524	2	17.365
261	285	4	9.073	471	474	2	5.409
413	431	4	19.491	504	523	2	12.689
4	239	3	21.371	520	563	2	16.155
12	51	3	9.158	524	533	2	14.583
116	136	3	14.676	524	545	2	15.211
221	239	3	19.349	1	261	2	25.43
221	242	3	21.401	211	238	2	5.484
273	280	3	13.825	221	231	2	13.758
294	439	3	18.376	221	431	2	26.429
350	375	3	12.436	431	465	2	24.671
350	377	3	13.801	431	471	2	23.977

Unique cross-links identified by in-gel digestion of BSA monomer band				Unique cross-links identified by in-gel digestion of BSA monomer band			
Site 1	Site 2	n° of identifications	Cα-Cα distance	Site 1	Site 2	n° of identifications	Cα-Cα distance
1	5	1	12.708	187	275	1	23.981
1	20	1	29.054	187	281	1	20.361
1	136	1	40.897	187	440	1	21.515
1	187	1	43.959	187	544	1	34.577
1	233	1	26.59	191	431	1	15.607
1	262	1	24.726	204	221	1	26.441
1	431	1	54.473	211	224	1	19.284
1	523	1	62.098	211	233	1	10.206
2	12	1	16.994	211	235	1	7.429
12	136	1	24.258	211	377	1	24.138
12	431	1	43.596	211	413	1	32.671
20	132	1	11.506	211	431	1	29.182
20	136	1	12.539	221	132	1	37.438
28	36	1	8.358	221	233	1	12.134
51	76	1	21.044	221	276	1	15.391
64	76	1	16.887	221	285	1	15.953
64	132	1	29.478	221	294	1	4.292
64	525	1	50.947	221	350	1	20.74
93	106	1	17.595	221	440	1	20.504
106	431	1	26.843	224	211	1	19.284
106	563	1	42.12	228	235	1	10.539
114	116	1	6.097	228	374	1	28.457
114	127	1	23.235	228	489	1	38.915
114	131	1	22.781	228	498	1	56.739
114	322	1	48.479	232	242	1	15.608
116	132	1	18.945	232	261	1	13.849
116	173	1	15.604	232	377	1	29.406
116	520	1	15.011	235	256	1	12.078
116	523	1	18	235	263	1	13.27
127	132	1	10.275	235	266	1	16.068
127	525	1	40.139	235	346	1	17.584
131	137	1	10.059	235	374	1	31.692
132	136	1	6.039	242	261	1	19.343
132	221	1	37.438	242	285	1	21.036
136	281	1	22.735	242	37	1	29.772
138	455	1	23.802	242	431	1	29.847
159	173	1	16.787	242	465	1	21.108
159	180	1	10.069	245	463	1	19.265
159	181	1	9.652	273	285	1	12.421
180	188	1	12.408	275	285	1	10.23
180	190	1	15.708	280	377	1	41.27
183	187	1	6.793	294	316	1	33.526
187	273	1	22.329	294	317	1	29.96

Unique cross-links identified by in-gel digestion of BSA monomer band				Unique cross-links identified by in-gel digestion of BSA monomer band			
Site 1	Site 2	n° of identifications	Cα-Cα distance	Site 1	Site 2	n° of identifications	Cα-Cα distance
312	316	1	8.686	520	523	1	5.239
312	377	1	21.679	520	525	1	8.642
316	524	1	59.931	523	534	1	16.917
316	535	1	57.78	524	520	1	6.418
322	351	1	13.859	524	536	1	17.398
322	377	1	22.37	524	537	1	19.71
331	377	1	17.138	524	539	1	18.674
333	377	1	11.918	524	563	1	17.617
346	374	1	16.773	526	544	1	16.441
350	396	1	36.378	544	548	1	6.441
350	439	1	28.354	548	568	1	19.254
362	377	1	23.012	556	573	1	12.899
374	377	1	4.768	1	1	1	
374	498	1	44.673				
377	396	1	31.799				
377	431	1	29.581				
377	465	1	37.277				
377	471	1	28.985				
377	563	1	60.734				
388	413	1	20.643				
388	439	1	12.318				
388	524	1	28.025				
396	413	1	22.332				
396	524	1	20.784				
396	545	1	17.941				
413	350	1	26.807				
413	439	1	28.228				
413	474	1	13.568				
419	535	1	14.523				
431	440	1	17.055				
431	474	1	25.423				
431	537	1	26.11				
437	561	1	43.12				
439	535	1	36.735				
455	463	1	12.27				
465	474	1	14.54				
465	524	1	25.43				
465	537	1	28.668				
471	524	1	25.895				
471	525	1	22.403				
499	524	1	18.38				
499	536	1	8.868				
504	520	1	17.023				

Table S4: Unique cross-links identified at different BSA concentrations. The cross-linking reactions were conducted in duplicates (500, 250, 50, 25 μ M DSBU, 25 mM HEPES, pH 7.2, incubation time 1 hour, 21 $^{\circ}$ C).

BSA conc.	Unique XLs	>30	%
10	80	2	2.5
5	71	2	2.8
1	71	2	2.8
0.5	41	1	2.4
all	127	3	2.4

SUPPLEMENTARY TEXT

Comparison of data acquisition and analysis strategies from one participating lab

Comparison of different instrument platforms with the same sample. As a first example, we analyzed the very same samples (representing three experimental replicates) of cross-linked BSA on different instrument platforms within the same laboratory. BSA was either cross-linked with the non-cleavable, amine-reactive DSS, or a combination of the dihydrazide PDH and the coupling reagent DMTMM, resulting in a combination of carboxyl-carboxyl and amine-carboxyl cross-links. Samples were either directly injected without fractionation or after fractionation with size exclusion chromatography (resulting in three fractions per sample), and analyzed by LC/MS/MS on three different instruments equipped with an orbitrap analyzer: (i) LTQ Orbitrap XL, (ii) Orbitrap Elite, and (iii) Orbitrap Fusion Lumos (only DSS sample). These three instruments differ in several essential performance characteristics, such as sensitivity and sequencing speed. Considering the relatively low sample complexity, it was not clear whether there would be significant differences between the three orbitrap instruments.

As shown in Figure S3, not unexpectedly, the positive influence of sample fractionation was most noticeable for the oldest generation instrument (LTQ Orbitrap XL), and was especially visible for the PDH links, which are known to be less abundant due to the low yield of this cross-linking chemistry. Nevertheless, an increase in the cross-links identified was observed for all three instrument platforms, at the expense of a three-fold increase in analysis time. It has to be pointed out that different fragmentation methods (CID in the ion trap on the Orbitrap Elite and HCD in the ion routing multipole on the Orbitrap Fusion Lumos, respectively) may explain part of the differences in performance. Nevertheless, the results show that reasonable cross-link coverage for individual proteins can easily be obtained on more than 10-year old instrumentation.

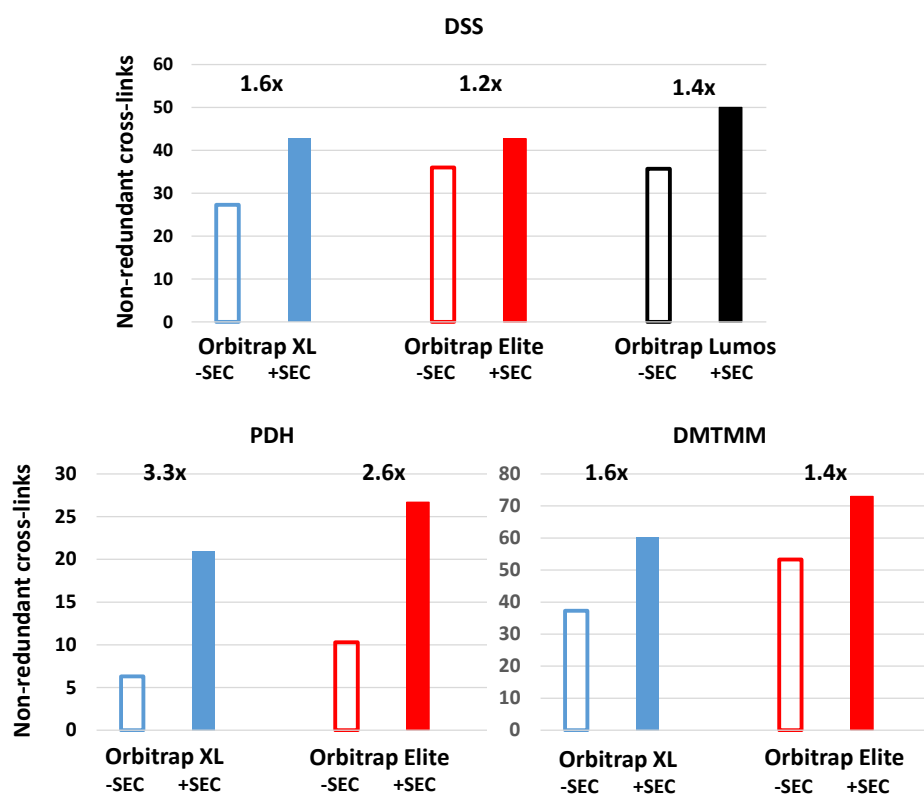


Figure S5: Comparison of non-redundant BSA cross-links identified on three different orbitrap MS platforms. Data were analyzed by the xQuest software. The cross-linkers (DSS, PDH, DMTMM) and the relative fold increase from unfractionated (-SEC) to fractionated (+SEC) samples are indicated.

Effect of different data analysis strategies on one single data set. To study the effect of search parameters and validation strategies, we re-searched DSS data sets (unfractionated samples) from the Orbitrap Elite and Orbitrap Fusion Lumos platforms using xQuest as search engine. Specifically, starting from a “minimal” parameter set (maximum of two missed cleavage sites, no variable modification and allowing only Lys as cross-linking site), we expanded the search space either slightly (including Met oxidation as variable modification) or severely (allowing up to four missed cleavages or considering Lys, Ser, Thr, Tyr and *N*-terminus as potential cross-linking sites). The latter setting, in particular, conforms to the relaxed specificity settings for *N*-hydroxysuccinimide ester-based reagents discussed in the main text. In addition, although the experiments were carried out with stable-isotope coded DSS (a 1:1 mixture of DSS- D_0 and $-D_{12}$), we also analyzed data by only

considering spectra from the “light” (D_0) form of the reagent, reflecting a scenario where a non-labeled cross-linker is used. Altogether, the different parameter sets reflect highly different search spaces that were expected to have different effects on data MS/MS data acquired in low or high resolution.

As shown in Figure S6 A, the different search parameter settings resulted in an up to 4.4-fold expansion of potential peptide pairs (from 8,128 to 36,046) considered during the search for four missed cleavages. However, despite the increased search space, neither considering more missed cleavages increased the number of identified cross-links noticeably, nor did the inclusion of Met-oxidized peptides (Figure S6 B-C). In fact, applying the same 5% FDR threshold on all data sets even led to a decrease of confident identifications after allowing for more missed cleavages, due to the increase in random matches (condition 2 vs. condition 1). Therefore, if efficient enzymatic digestion can be assumed (in this case, a two-step digestion with endoprotease Lys-C, followed by trypsin was used), it may be counterproductive to relax the stringency of this proteolysis-related setting.

In contrast, relaxing the specificity of the DSS reaction by including the side chains containing hydroxy groups in Ser, Thr, and Tyr as possible targets for cross-linking led to a higher number of accepted cross-links at the same FDR. This effect appeared to be particularly noticeable on the Orbitrap Fusion Lumos platform (for which MS/MS spectra were acquired in high-resolution mode). However, a closer evaluation revealed that many new low-quality assignments were accepted at the 5% FDR threshold due to the overall low frequency of decoy hits that made the standard target/decoy model less suitable in such cases. After additional validation/filtering criteria were applied, the increase in confident cross-links was comparable for the Orbitrap Elite and Orbitrap Fusion Lumos platforms (Figure S6 D-E). Moreover, as could be expected, if more residues are considered as potential cross-linking sites, the number of identifications with unequivocal site localization decreases (Figure S6 D-E). If only Lys is considered as reactive site, this will obviously result in clear cross-link assignments.

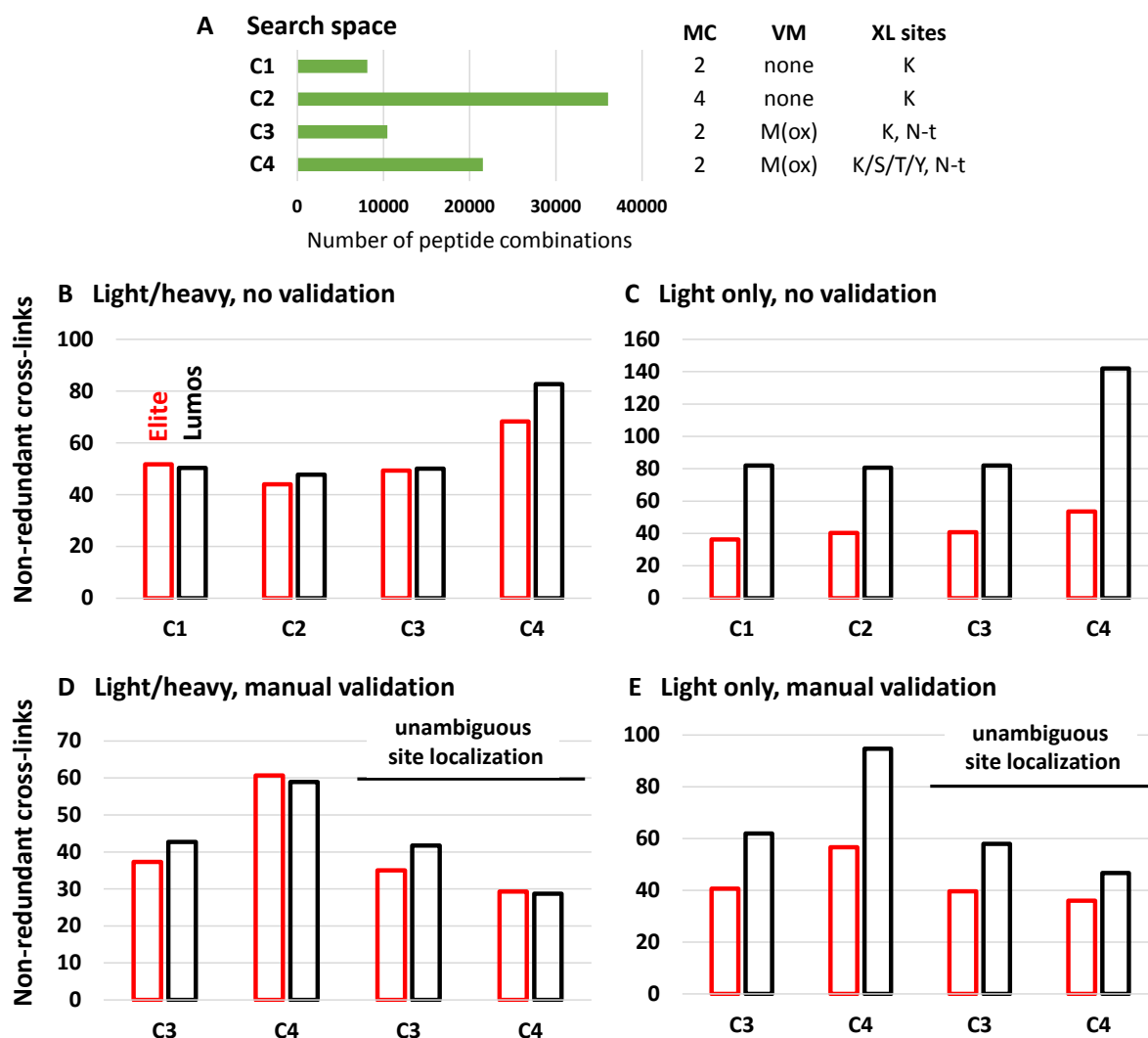


Figure S6: Comparison of different data analysis strategies for DSS (“light” D_0 and “heavy” D_{12}) data sets acquired on Orbitrap Elite or Orbitrap Fusion Lumos instruments. (A) Effect of search parameters on the search space, defined as possible peptide pair combinations without applying any filters. (B-E) Non-redundant cross-links identified for different parameter sets: (B) “light”/“heavy” linker, no manual validation, (C) “light” linker, no-manual validation, (D) “light”/“heavy” linker, with manual validation, (E) “light” linker, with manual validation. In (D) and (E), results are shown either with the highest scoring cross-linking site localization or after manual assessment of site localization, excluding identifications with ambiguous site assignment. Condition 1 (C1): two missed cleavages (MC), no variable modification (VM), Lys as cross-linking site (XL site); C2: four missed cleavages; C3: two missed cleavages, Met oxidation (M(ox)) as variable modification, Lys and *N*-terminus (N-t) as

cross-linking sites; C4: two missed cleavages, Met oxidation as variable modification, and Lys, Ser, Thr, Tyr, and *N*-terminus as cross-linking sites.