Supporting Information

CLMSVault : a software suite for protein cross-linking mass spectrometry data analysis and visualization

Mathieu Courcelles[1]*, Jasmin Coulombe-Huntington*[1]*, Émilie Cossette*[1]*, Anne-Claude Gingras*[2,3]*, Pierre Thibault*[1,4]*, Mike Tyers*[1,5]***

[1] Institute for Research in Immunology and Cancer, Université de Montréal, Montréal, Québec H3C 3J7, Canada

[2]Lunenfeld-Tanenbaum Research Institute at Sinai Health Service, Toronto, Ontario, Canada

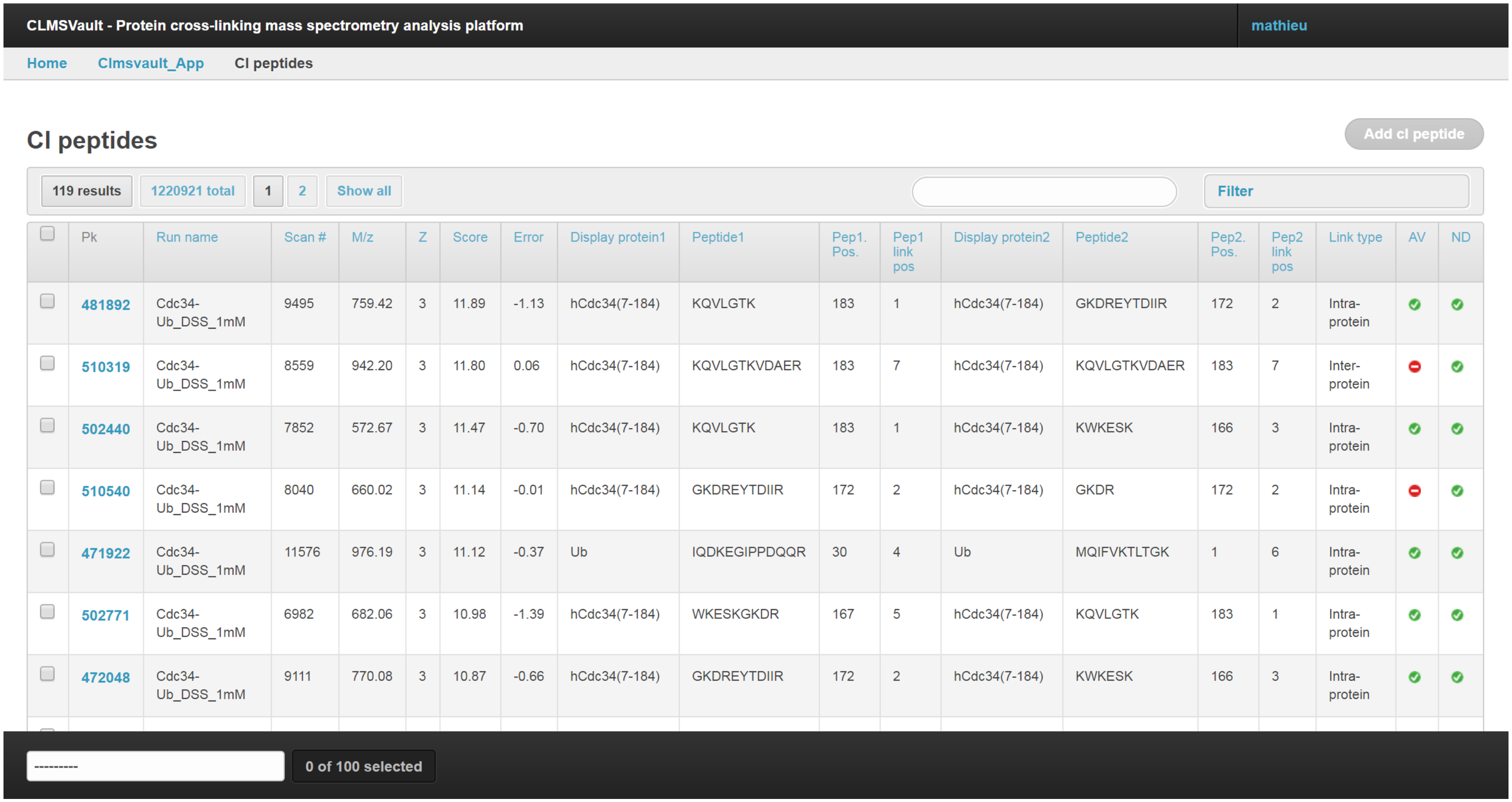
[3]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

[4]Department of Chemistry, Université de Montréal, Montréal, Québec H3C 3J7, Canada

[5]Department of Medicine, Université de Montréal, Montréal, Québec H3C 3J7, Canada

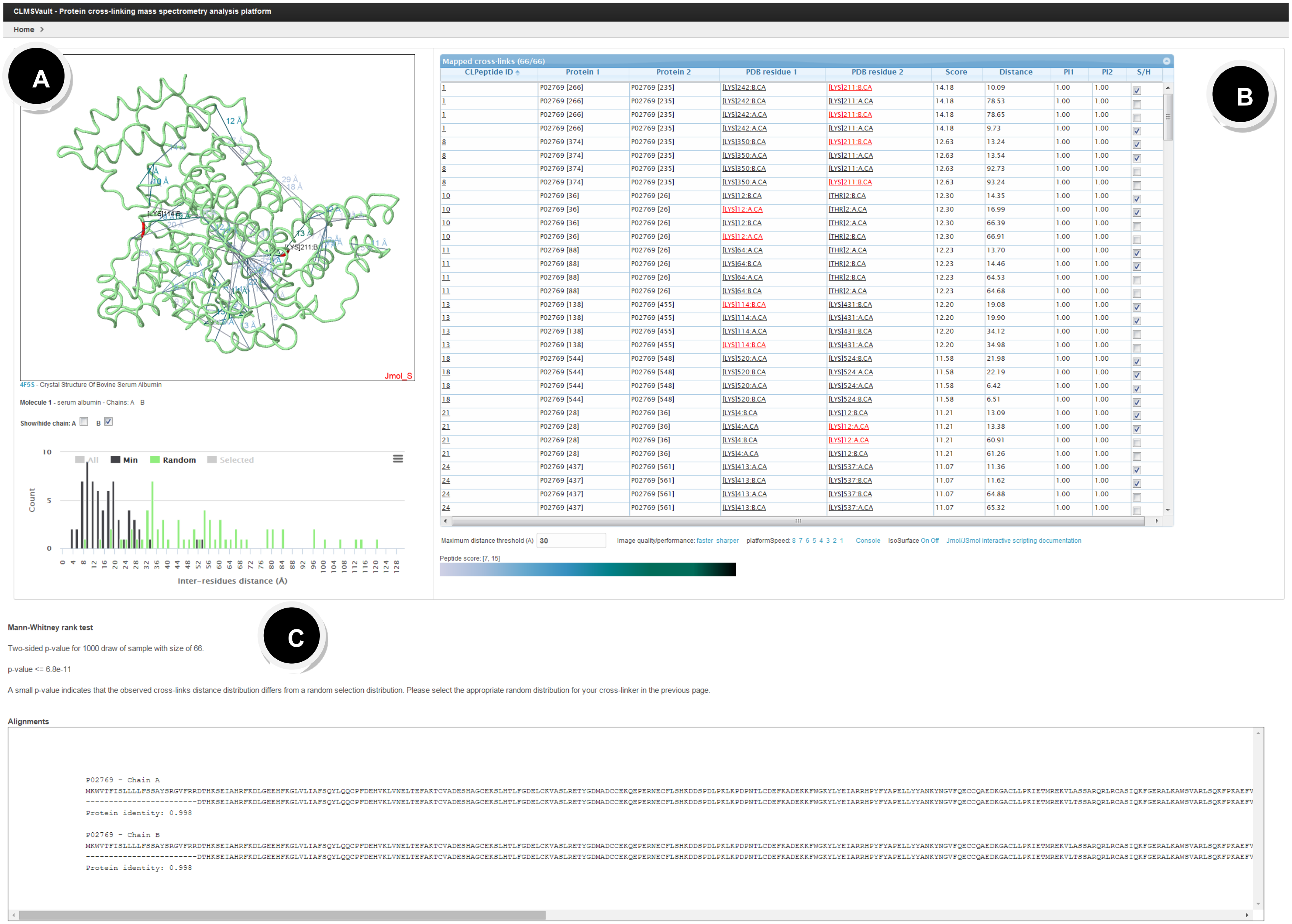
* Correspondence: md.tyers@umontreal.ca; tel 514-343-6668

TABLE OF CONTENT:



☐	Pk	Run name	Scan #	M/z	Z	Score	Error	Display protein1	Peptide1	Pep1. Pos.	Pep1 link pos	Display protein2	Peptide2	Pep2. Pos.	Pep2 link pos	Link type	AV	ND
☐	481892	Cdc34-Ub_DSS_1mM	9495	759.42	3	11.89	-1.13	hCdc34(7-184)	KQVLGTK	183	1	hCdc34(7-184)	GKDREYTDIIR	172	2	Intra-protein	✓	✓
☐	510319	Cdc34-Ub_DSS_1mM	8559	942.20	3	11.80	0.06	hCdc34(7-184)	KQVLGTKVDAER	183	7	hCdc34(7-184)	KQVLGTKVDAER	183	7	Inter-protein	⊖	✓
☐	502440	Cdc34-Ub_DSS_1mM	7852	572.67	3	11.47	-0.70	hCdc34(7-184)	KQVLGTK	183	1	hCdc34(7-184)	KWKESK	166	3	Intra-protein	✓	✓
☐	510540	Cdc34-Ub_DSS_1mM	8040	660.02	3	11.14	-0.01	hCdc34(7-184)	GKDREYTDIIR	172	2	hCdc34(7-184)	GKDR	172	2	Intra-protein	⊖	✓
☐	471922	Cdc34-Ub_DSS_1mM	11576	976.19	3	11.12	-0.37	Ub	IQDKEGIPPDQQR	30	4	Ub	MQIFVKTLTGK	1	6	Intra-protein	✓	✓
☐	502771	Cdc34-Ub_DSS_1mM	6982	682.06	3	10.98	-1.39	hCdc34(7-184)	WKESKGKDR	167	5	hCdc34(7-184)	KQVLGTK	183	1	Intra-protein	✓	✓
☐	472048	Cdc34-Ub_DSS_1mM	9111	770.08	3	10.87	-0.66	hCdc34(7-184)	GKDREYTDIIR	172	2	hCdc34(7-184)	KWKESK	166	3	Intra-protein	✓	✓

Figure S-1. User interface screenshots of a cross-linked peptides dataset.



CLMSVault - Protein cross-linking mass spectrometry analysis platform

Home >

4F5S - Crystal Structure Of Bovine Serum Albumin

Molecule 1 - serum albumin - Chains: A B

Show/hide chain: A ☐ **B** ☑

Mapped cross-links (66/66)

CLPeptide ID	Protein 1	Protein 2	PDB residue 1	PDB residue 2	Score	Distance	PI1	PI2	S/H
1	P02769 [266]	P02769 [235]	[LYS]242:B.CA	[LYS]211:B.CA	14.18	10.09	1.00	1.00	☑
1	P02769 [266]	P02769 [235]	[LYS]242:B.CA	[LYS]211:A.CA	14.18	78.53	1.00	1.00	☐
1	P02769 [266]	P02769 [235]	[LYS]242:A.CA	[LYS]211:B.CA	14.18	78.65	1.00	1.00	☐
1	P02769 [266]	P02769 [235]	[LYS]242:A.CA	[LYS]211:A.CA	14.18	9.73	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:B.CA	[LYS]211:B.CA	12.63	13.24	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:A.CA	[LYS]211:A.CA	12.63	13.54	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:B.CA	[LYS]211:A.CA	12.63	92.73	1.00	1.00	☐
8	P02769 [374]	P02769 [235]	[LYS]350:A.CA	[LYS]211:B.CA	12.63	93.24	1.00	1.00	☐
10	P02769 [36]	P02769 [26]	[LYS]12:B.CA	[THR]2:B.CA	12.30	14.35	1.00	1.00	☑
10	P02769 [36]	P02769 [26]	[LYS]12:A.CA	[THR]2:A.CA	12.30	16.99	1.00	1.00	☑
10	P02769 [36]	P02769 [26]	[LYS]12:B.CA	[THR]2:A.CA	12.30	66.39	1.00	1.00	☐
10	P02769 [36]	P02769 [26]	[LYS]12:A.CA	[THR]2:B.CA	12.30	66.91	1.00	1.00	☐
11	P02769 [88]	P02769 [26]	[LYS]64:A.CA	[THR]2:A.CA	12.23	13.70	1.00	1.00	☑
11	P02769 [88]	P02769 [26]	[LYS]64:B.CA	[THR]2:B.CA	12.23	14.46	1.00	1.00	☑
11	P02769 [88]	P02769 [26]	[LYS]64:A.CA	[THR]2:B.CA	12.23	64.53	1.00	1.00	☐
11	P02769 [88]	P02769 [26]	[LYS]64:B.CA	[THR]2:A.CA	12.23	64.68	1.00	1.00	☐
13	P02769 [138]	P02769 [455]	[LYS]114:B.CA	[LYS]431:B.CA	12.20	19.08	1.00	1.00	☑
13	P02769 [138]	P02769 [455]	[LYS]114:A.CA	[LYS]431:A.CA	12.20	19.90	1.00	1.00	☑
13	P02769 [138]	P02769 [455]	[LYS]114:A.CA	[LYS]431:B.CA	12.20	34.12	1.00	1.00	☐
13	P02769 [138]	P02769 [455]	[LYS]114:B.CA	[LYS]431:A.CA	12.20	34.98	1.00	1.00	☐
18	P02769 [544]	P02769 [548]	[LYS]520:A.CA	[LYS]524:B.CA	11.58	21.98	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:B.CA	[LYS]524:A.CA	11.58	22.19	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:A.CA	[LYS]524:A.CA	11.58	6.42	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:B.CA	[LYS]524:B.CA	11.58	6.51	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:B.CA	[LYS]12:B.CA	11.21	13.09	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:A.CA	[LYS]12:A.CA	11.21	13.38	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:B.CA	[LYS]12:A.CA	11.21	60.91	1.00	1.00	☐
21	P02769 [28]	P02769 [36]	[LYS]4:A.CA	[LYS]12:B.CA	11.21	61.26	1.00	1.00	☐
24	P02769 [437]	P02769 [561]	[LYS]413:A.CA	[LYS]537:A.CA	11.07	11.36	1.00	1.00	☑
24	P02769 [437]	P02769 [561]	[LYS]413:B.CA	[LYS]537:B.CA	11.07	11.62	1.00	1.00	☑
24	P02769 [437]	P02769 [561]	[LYS]413:A.CA	[LYS]537:B.CA	11.07	64.88	1.00	1.00	☐
24	P02769 [437]	P02769 [561]	[LYS]413:B.CA	[LYS]537:A.CA	11.07	65.32	1.00	1.00	☐

Maximum distance threshold (A) 30

Image quality/performance: faster sharper platformSpeed: 8 7 6 5 4 3 2 1 Console IsoSurface On Off Jmol/JSmol interactive scripting documentation

Peptide score: [7, 15]

Mann-Whitney rank test

Two-sided p-value for 1000 draw of sample with size of 66.

p-value <= 6.8e-11

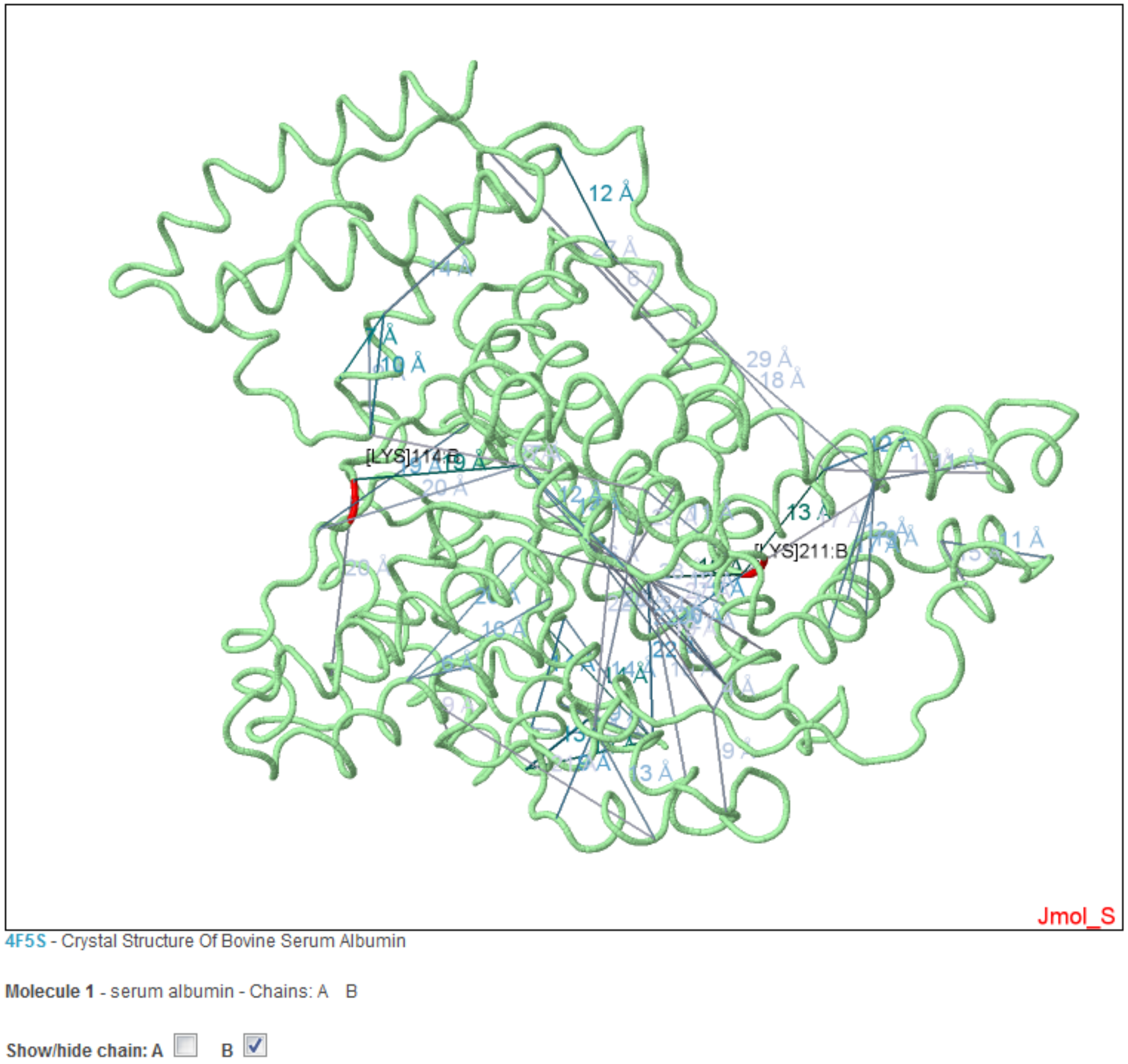
A small p-value indicates that the observed cross-links distance distribution differs from a random selection distribution. Please select the appropriate random distribution for your cross-linker in the previous page.

Alignments

```
P02769 - Chain A
MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFV
------------------------DTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLTSSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFV
Protein identity: 0.998

P02769 - Chain B
MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFV
------------------------DTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFAKTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTLCDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLTSSARQRLRCASIQKFGERALKAWSVARLSQKFPKAEFV
Protein identity: 0.998
```

Figure S-2. User interface screenshot of the 3D cross-links viewer

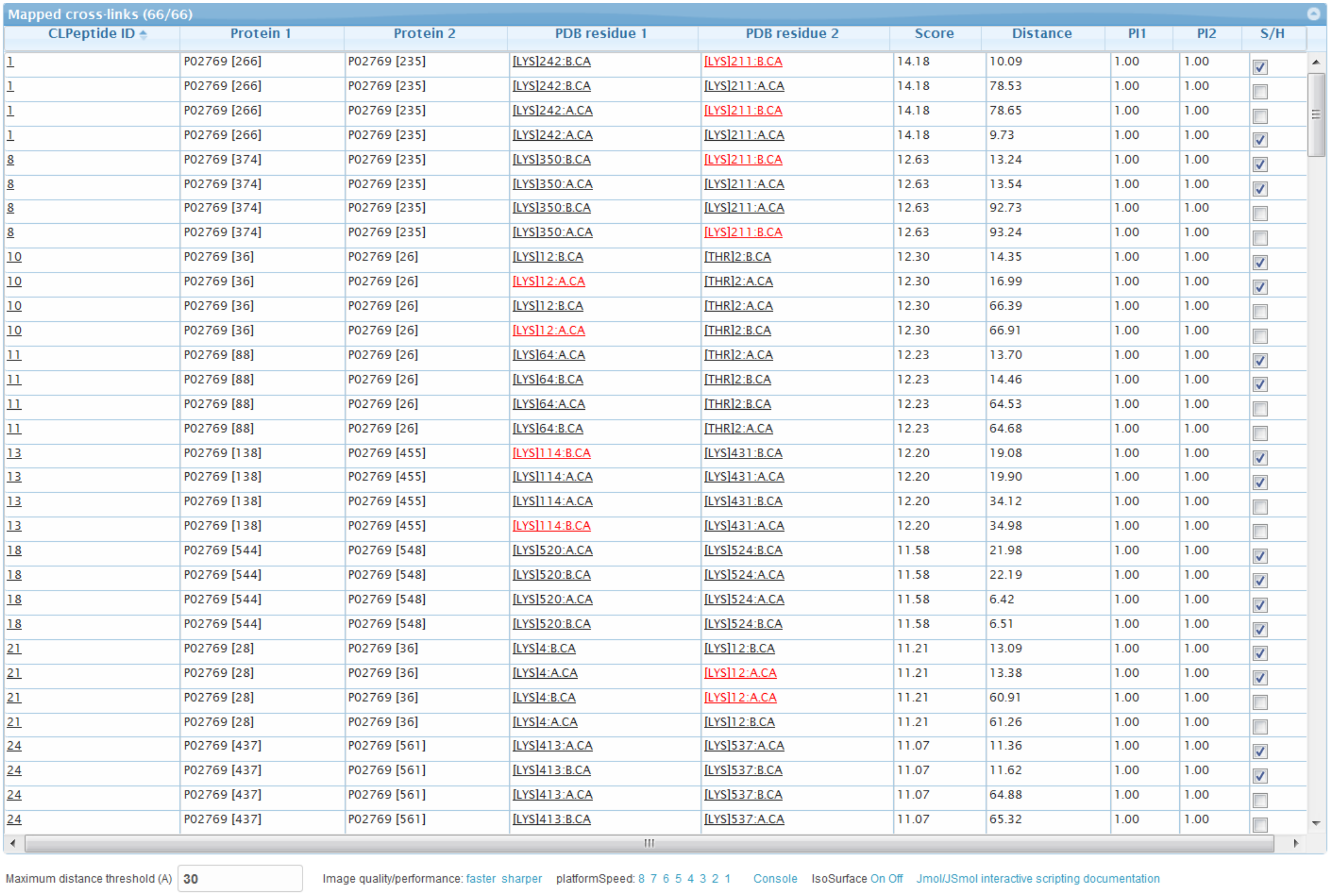


4F5S - Crystal Structure Of Bovine Serum Albumin

Molecule 1 - serum albumin - Chains: A B

Show/hide chain: A ☐ **B** ☑

Figure S-2A. Zoomed user interface screenshot of the 3D cross-links viewer

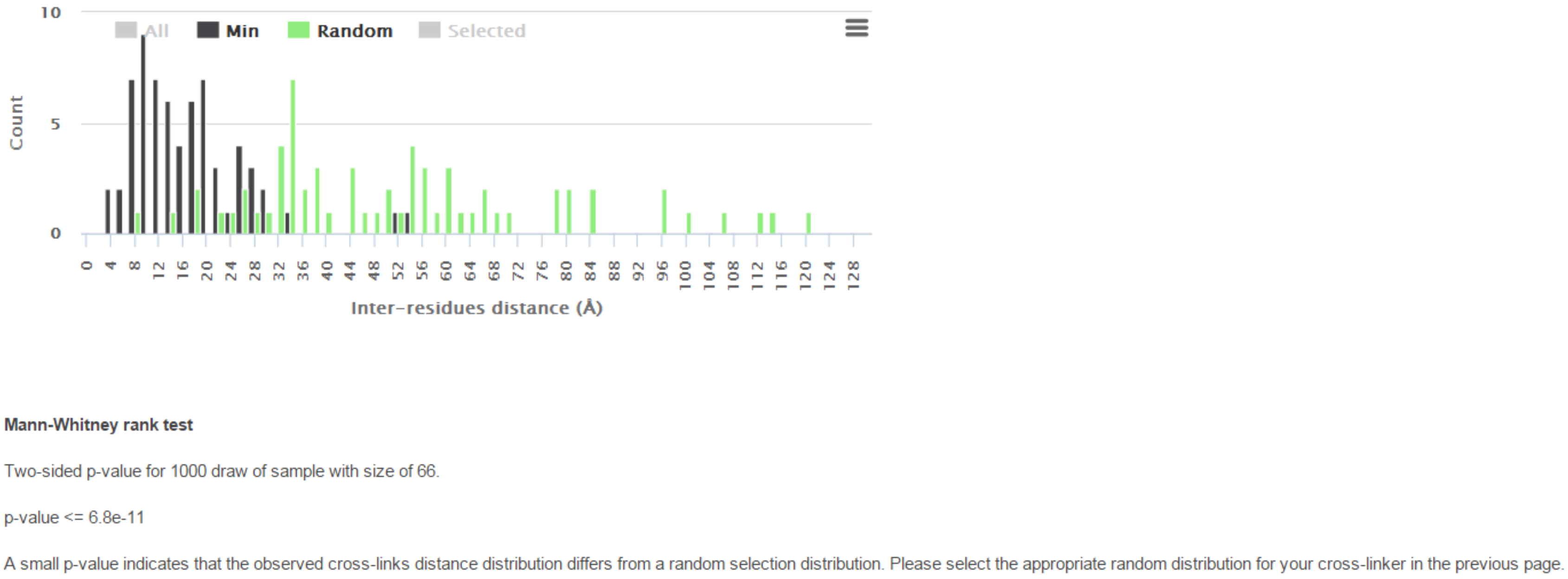


Mapped cross-links (66/66)

CLPeptide ID	Protein 1	Protein 2	PDB residue 1	PDB residue 2	Score	Distance	PI1	PI2	S/H
1	P02769 [266]	P02769 [235]	[LYS]242:B.CA	[LYS]211:B.CA	14.18	10.09	1.00	1.00	☑
1	P02769 [266]	P02769 [235]	[LYS]242:B.CA	[LYS]211:A.CA	14.18	78.53	1.00	1.00	☐
1	P02769 [266]	P02769 [235]	[LYS]242:A.CA	[LYS]211:B.CA	14.18	78.65	1.00	1.00	☐
1	P02769 [266]	P02769 [235]	[LYS]242:A.CA	[LYS]211:A.CA	14.18	9.73	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:B.CA	[LYS]211:B.CA	12.63	13.24	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:A.CA	[LYS]211:A.CA	12.63	13.54	1.00	1.00	☑
8	P02769 [374]	P02769 [235]	[LYS]350:B.CA	[LYS]211:A.CA	12.63	92.73	1.00	1.00	☐
8	P02769 [374]	P02769 [235]	[LYS]350:A.CA	[LYS]211:B.CA	12.63	93.24	1.00	1.00	☐
10	P02769 [36]	P02769 [26]	[LYS]12:B.CA	[THR]2:B.CA	12.30	14.35	1.00	1.00	☑
10	P02769 [36]	P02769 [26]	[LYS]12:A.CA	[THR]2:A.CA	12.30	16.99	1.00	1.00	☑
10	P02769 [36]	P02769 [26]	[LYS]12:B.CA	[THR]2:A.CA	12.30	66.39	1.00	1.00	☐
10	P02769 [36]	P02769 [26]	[LYS]12:A.CA	[THR]2:B.CA	12.30	66.91	1.00	1.00	☐
11	P02769 [88]	P02769 [26]	[LYS]64:A.CA	[THR]2:A.CA	12.23	13.70	1.00	1.00	☑
11	P02769 [88]	P02769 [26]	[LYS]64:B.CA	[THR]2:B.CA	12.23	14.46	1.00	1.00	☑
11	P02769 [88]	P02769 [26]	[LYS]64:A.CA	[THR]2:B.CA	12.23	64.53	1.00	1.00	☐
11	P02769 [88]	P02769 [26]	[LYS]64:B.CA	[THR]2:A.CA	12.23	64.68	1.00	1.00	☐
13	P02769 [138]	P02769 [455]	[LYS]114:B.CA	[LYS]431:B.CA	12.20	19.08	1.00	1.00	☑
13	P02769 [138]	P02769 [455]	[LYS]114:A.CA	[LYS]431:A.CA	12.20	19.90	1.00	1.00	☑
13	P02769 [138]	P02769 [455]	[LYS]114:A.CA	[LYS]431:B.CA	12.20	34.12	1.00	1.00	☐
13	P02769 [138]	P02769 [455]	[LYS]114:B.CA	[LYS]431:A.CA	12.20	34.98	1.00	1.00	☐
18	P02769 [544]	P02769 [548]	[LYS]520:A.CA	[LYS]524:B.CA	11.58	21.98	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:B.CA	[LYS]524:A.CA	11.58	22.19	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:A.CA	[LYS]524:A.CA	11.58	6.42	1.00	1.00	☑
18	P02769 [544]	P02769 [548]	[LYS]520:B.CA	[LYS]524:B.CA	11.58	6.51	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:B.CA	[LYS]12:B.CA	11.21	13.09	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:A.CA	[LYS]12:A.CA	11.21	13.38	1.00	1.00	☑
21	P02769 [28]	P02769 [36]	[LYS]4:B.CA	[LYS]12:A.CA	11.21	60.91	1.00	1.00	☐
21	P02769 [28]	P02769 [36]	[LYS]4:A.CA	[LYS]12:B.CA	11.21	61.26	1.00	1.00	☐
24	P02769 [437]	P02769 [561]	[LYS]413:A.CA	[LYS]537:A.CA	11.07	11.36	1.00	1.00	☑
24	P02769 [437]	P02769 [561]	[LYS]413:B.CA	[LYS]537:B.CA	11.07	11.62	1.00	1.00	☑
24	P02769 [437]	P02769 [561]	[LYS]413:A.CA	[LYS]537:B.CA	11.07	64.88	1.00	1.00	☐
24	P02769 [437]	P02769 [561]	[LYS]413:B.CA	[LYS]537:A.CA	11.07	65.32	1.00	1.00	☐

Maximum distance threshold (A) 30

Figure S-2B. Zoomed user interface screenshot of the 3D cross-links viewer



Mann-Whitney rank test

Two-sided p-value for 1000 draw of sample with size of 66.

p-value <= 6.8e-11

A small p-value indicates that the observed cross-links distance distribution differs from a random selection distribution. Please select the appropriate random distribution for your cross-linker in the previous page.

Figure S-2C. Zoomed user interface screenshot of the 3D cross-links viewer

Table S-1. CLMSVault features comparison against other processing pipelines

Feature	XLinkDB	ProXL DB	CLMSVault
Dataset import	Tab delimited text file	ProXL XML, converter for Kojak, Crux, pLink, StavroX, xQuest	Results files from pLink, Xi, xQuest, 14N15N DXMSMS, Mascot (dead-end)
Dataset export	Tab delimited text file	ProXL XML, Tab delimited text file	Comma-separated text file, PSI-MI TAB and XML 2.5
Dataset filtering	-	Yes	Yes
Spectrum viewer	-	Lorikeet	xiSPEC
Protein network viewer	Cytoscape web, xiNET	xVis	xiNET
Protein interactions reference database	IntAct, EciD	-	-
Cross-links overlay on PDB structure	Jmol/JSmol viewer	Pv viewer	Jmol/JSmol viewer
Structure Modeling	Modeller, PatchDock	-	-
Quantitative analysis	-	-	Cross-links color scheme based on abundance fold change for protein network and structure viewers. Filter.
Protein database	Uniprot (E. coli, H. sapiens, S. cerevisae, A. thaliana)	No restriction	No restriction
Source code	No	Yes	Yes