Supplementary material

’Clickable lectins’: Bioorthogonal reactive handles facilitate the directed conjugation of lectins in a modular fashion

Felix Tobola[1,2], Elise Sylvander[1 †], Claudia Gafko[1,2 ‡] and Birgit Wiltschi[*1]

1 Austrian Centre of Industrial Biotechnology, Petersgasse 14, 8010 Graz, Austria

2 Institute of Molecular Biotechnology, Graz University of Technology, Petersgasse 14, 8010 Graz, Austria

* Author for correspondence: Birgit Wiltschi; e-mail: birgit.wiltschi@acib.at

† Present address: Institut National des Sciences Appliquées de Toulouse, 35 Avenue de Rangueil, 31400 Toulouse, Haute-Garonne (31), France

‡ Present address: Institute of Biochemistry, Department of Biology, ETH Zurich, Otto-Stern-Weg 3, 8093 Zürich, Switzerland

Supplementary results and discussion

Conjugation reactions with bis-propargyl-PEG linkers are shown in electronic supplementary material, figure S3. These linkers were less efficient for the conjugation of Gal-1[AzK] and Stx1B[AzK] than DBCO-PEG$_4$-DBCO (main text, figure 4*b,e*, Coomassie-stained gel images, white arrows *vs* electronic supplementary material, figure S3*b,e*) The presence of a hexahistidine-tag on both proteins could explain this observation. The tag chelates divalent cations, such as Ni^{2+} and Cu^{2+}, which is exploited for the purification of hexahistidine-tagged proteins by immobilized metal affinity chromatography [1]. The hexahistidine-tag on Gal-1[AzK] and Stx1B[AzK] could have chelated the copper ions that were used as a catalyst for the CuAAC-reaction [2]. Possibly, this resulted in the reduced conjugation-efficiency in the CuAAC reactions with the bis-propargyl-PEG linkers (that do require a copper catalyst) but not the SPAAC reactions with the DBCO-PEG$_4$-DBCO linker (which occurs without a copper catalyst). The other two azido-lectins did not contain a hexahistidine-tag and showed comparable homo-conjugation efficiencies with all PEG-linkers in this study (main text, figure 4*c,d*, Coomassie-stained gel images, white arrows *vs* electronic supplementary material, figure S3*c,d*). Noticeably, we observed protein loss due to precipitation in the CuAAC reactions, which was not the case with the SPAAC reactions. This confirms the previous observation [3] that copper and other additives in the CuAAC reaction may be detrimental to protein stability and/or function. The differences in PEG-spacer-size of the two bis-propargyl-linkers did not influence the conjugation efficiency (electronic supplementary material, figure S3).

References

1. Porath J. 1992 Immobilized metal ion affinity chromatography. *Protein Expr. Purif.* **3**, 263-281. (doi:10.1016/1046-5928(92)90001-D)
2. Presolski SI, Hong VP, Finn MG. 2011 Copper-catalyzed azide-alkyne click chemistry for bioconjugation. *Curr. Protoc. Chem. Biol.* **3**, 153-162. (doi:10.1002/9780470559277.ch110148)
3. Hong V, Presolski SI, Ma C, Finn MâG. 2009 Analysis and optimization of copper-catalyzed azide-alkyne cycloaddition for bioconjugation. *Angew. Chem. Int. Ed. Engl.* **48**, 9879-9883. (doi:10.1002/anie.200905087)

Table S1 Mass analysis of the lectin variants by HPLC ESI-MS.

Sample	calculated mass (Da)	measured mass (Da)	Δmass calculated / measured (Da)	Peak intensity (AU)	Peak intensity (%)	Note	Wildtype protein	Wildtype mass (Da)
RSL[Aha]	10062.71	10062.71	0.00	1771016	**100.0**	RSL[Aha]	RSL	10067.7
		10090.70	27.99	130301	7.4	Formylation of RSL[Aha]?		
		10079.71	17.00	108280	6.1			
		10100.65	37.94	106666	6.0	Pyroglutamic acid formation from Gln in RSL[Aha]?		
RSL[Hpg]	10045.71	10045.66	-0.05	167592	**100.0**	RSL[Hpg]	RSL	10067.7
		10072.66	26.95	54865	32.7			
		10065.65	19.94	15388	9.2			
		10084.62	38.91	14863	8.9			
		10110.61	64.90	13070	7.8			
Gal-1[Aha]	14860.50	14860.45	-0.05	4529037	**100.0**	Gal-1[Aha]	Gal-1	14865.48
		14877.47	16.97	594857	13.1			
		14842.46	-18.04	266000	5.9			
Gal-1[Hpg]	14843.49	14843.44	-0.05	2283595	**100.0**	Gal-1[Hpg]	Gal-1	14865.48
		14870.43	26.94	466896	20.4			
		14860.45	16.96	330488	14.5	Gal-1[Aha] ?		
		14882.40	38.91	181063	7.9			
Gal-1[AzK]	15703.81	15703.79	-0.02	3329524	**100.0**	Gal-1[AzK]	Gal-1 (His-tagged)	15576.74
		15676.79	-27.02	1235462	37.1			
		15719.79	15.98	387406	11.6	Oxidation of Met in Gal-1[AzK]?		
		15742.77	38.96	270089	8.1			
		15692.81	-11.00	164975	5.0			
Stx1B[AzK] w/o N-term Met	8893.28	8778.19	-115.09	5732636	100.0	Stx1B w/o N-term Met; S-S bridge	Stx1B w/o N-term Met	8780.26
		8891.21	-2.07	3415387	**59.6**	Stx1B[AzK] w/o N-term Met; S-S bridge		
		8865.21	-28.07	717589	12.5			

Table S2 DNA and amino acid (AA) sequences of genes, primers and proteins used in this study.

Name	Type	Sequence (DNA or AA)
rsl	gene	ATGGAAAGCAGCAGTGTTCAGACCGCAGCCACCAGCTGGGGTACCGTTCCGAGCA TTCGTGTGTATACAGCAAATAACGGTAAAATTACCGAACGTTGTTGGGATGGCAAA GGTTGGTATACCGGTGCATTTAATGAACCGGGCGATAACGTTAGCGTGACCTCTTG GCTGGTTGGTAGCGCAATTCATATCCGTGTGTATGCTAGCACCGGCACCACGACCA CGGAATGGTGTTGGGATGGTAATGGCTGGACCAAAGGTGCATATACCGCAACTAAT TAATGA
RSL	protein[a]	MESSSVQTAATSWGTVPSIRVYTANNGKITERCWDGKGWYTGAFNEPGDNVSVTSWL VGSAIHIRVYASTGTTTTEWCWDGNGWTKGAYTATN
gal-1	gene[b]	ATGCATCACCATCACCATCACGGATCGGCCAGCGGCCTTGTAGCCAGTAACTTGA ACTTAAAGCCGGGAGAGAGTTTGCGCGTACGTGGCGAAGTTGCTCCCGACGCCAA ATCCTTTGTGCTTAATCTGGGAAAAGACTCCAATAACCTGAGCTTGCACTTCAACCC TCGTTTCAATGCACACGGGGACGCAAACACTATTGTGTCAAATAGTAAAGATGGTG GTGCGTGGGGCACGGAGCAACGTGAGGCCGTGTTCCCTTTCCAACCGGGGAGTG TGGCTGAAGTAAGCATTACTTTCGACCAAGCCAACTTGACAGTTAAGTTGCCCGAT GGATATGAGTTCAAATTTCCGAATCGCTTAAACTTGGAAGCGATCAACTACATGGCC GCGGACGGTGACTTTAAGATTAAGTCTGTTGCCTTCGATTGATAA
Gal-1	protein[c]	MHHHHHHGSASGLVASNLNLKPGESLRVRGEVAPDAKSFVLNLGKDSNNLSLHFNPR FNAHGDANTIVSNSKDGGAWGTEQREAVFPFQPGSVAEVSITFDQANLTVKLPDGYEF KFPNRLNLEAINYMAADGDFKIKSVAFD
stx1B	gene[d]	ATGGCGACGCCTGATTGTGTAACTGGAAAGGTGGAGTATACAAAATATAATGATGA CGATACCTTTACAGTTAAAGTGGGTGATAAAGAATTATTTACCAACAGATGGAATCT TCAGTCTCTTCTTCTCAGTGCGCAAATTACGGGGATGACTGTAACCATTAAAACTAA TGCCTGTCATAATGGAGGGGGATTCAGCGAAGTTATTTTTCGTGGTTCTGGCCATC ACCATCACCATCACTAA
Stx1B	protein[e]	MATPDCVTGKVEYTKYNDDDTFTVKVGDKELFTNRWNLQSLLLSAQITGMTVTIKTNAC HNGGGFSEVIFRGSGHHHHHH
pBP_CSGal-1_fw	primer	ACAGAATTCATTAAAGAGGAGAAATTAACT
pBP_CSGal-1_rv	primer	ATACGCATCTGCAGTTATCAATC
pBP6H-CSGal1_bb_rev	primer	CGATCCGTGATGGTGATGGTGATGCATAGTTAATTTCTCCTCTTTAATGAATTCTGT
pBP6H-CSGal_bb_fw	primer	GCGCCACATAGCAGAACTT
pBP6H-CSGal-1_fw	primer	ATGCATCACCATCACCATCACGGATCGGCCAGCGGCC
pBPCSGal-1_rev2	primer	TTGGCTGCAGTTATCAATCG
pBP2093	primer	GTCCCCGTGTGCCTAGAAACGAGGGTT
pBP2092	primer	AACCCTCGTTTCTAGGCACACGGGGAC
pBP1097	primer	GCAACCGAGCGTTCTGAAC
pBP2035	primer	ACAGAATTCATTAAAGAGGAGAAATTAACTATGGAGGGATCGGCCAGCGG
pBP2036	primer	GTCCGCGGCCAGGTAGTTGATC
pBP2037	primer	GATCAACTACCTGGCCGCGGAC
pBP1647	primer	CATTAAAGAGGAGAAATTAACTATGGAAAGCAGCAGTGTTCAGACCGC
pBP1648	primer	GCGGTCTGAACACTGCTGCTTTCCATAGTTAATTTCTCCTCTTTAATG
pBP2130	primer	GCGGATAACAATTTCACACAGAATTC
pBP2131	primer	CTGGATCTATCAACAGGAGTCGA

[a] The methionine shown in **red** is the site of Aha or Hpg incorporation in the clickable variants RSL[Aha] and RSL[Hpg], respectively.

[b] The nucleotides shown in **blue** were exchanged against the triplet codon GAG in the variant for Aha incorporation. The triplet codon shown in **red** was exchanged against the amber stop codon (TAG) in the variant for AzK incorporation.

[c] The methionine shown in **green** is the site of Aha or Hpg incorporation in Gal-1[Aha] and Gal-1[Hpg], respectively. Additionally, the AAs shown in **blue** were exchanged against glutamic acid and the methionine shown in **purple** was exchanged against leucine in the same clickable variants. The arginine shown in **red** is the site of AzK incorporation in the variant Gal-1[AzK].

[d] The triplet codon shown in **red** was exchanged against the amber stop codon (TAG) for the site-specific incorporation of AzK.

[e] The lysine shown in **red** is the site of AzK incorporation in the clickable variant Stx1B[AzK].

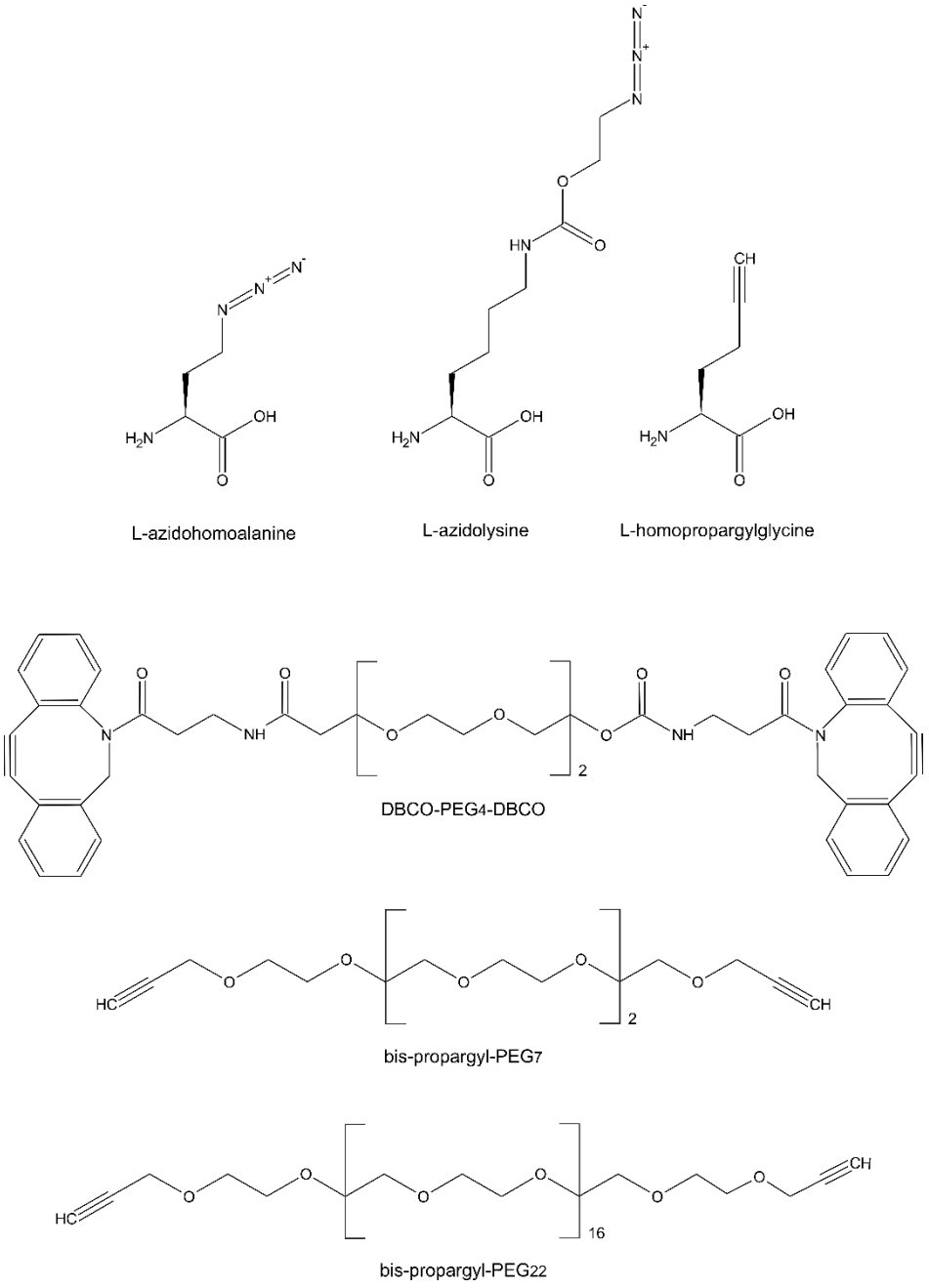

Figure S1 NcAA- and linker-structures. The chemical structures of the ncAAs, as well as the bi-functional PEG-linkers applied in this study are shown.

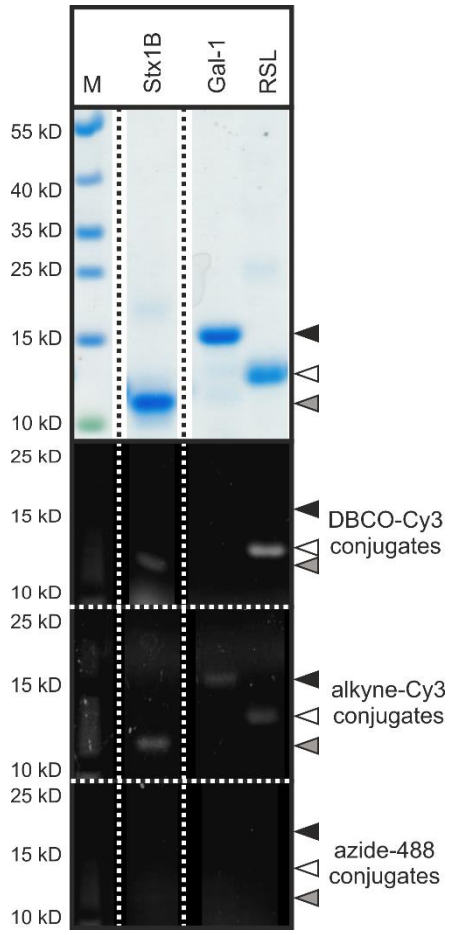

Figure S2 Unspecific interaction of the azido-and alkyne-fluorophores with the wildtype lectins. Wildtype lectins not containing reactive handles were mixed with three different fluorophores either carrying reactive azido- (azide-488) or alkyne-groups (DBCO-Cy3 or alkyne-Cy3). The samples were applied to SDS-PAGE and fluorophores interacting with lectins were visualised under UV light at 302 nm before the proteins were Coomassie stained. Black arrows, Gal-1; white arrows, RSL; grey arrows, Stx1B. Numbers on the left indicate the size of the molecular weight marker bands. The figure is composed of samples run on different gels, the individual images are separated by dotted lines.

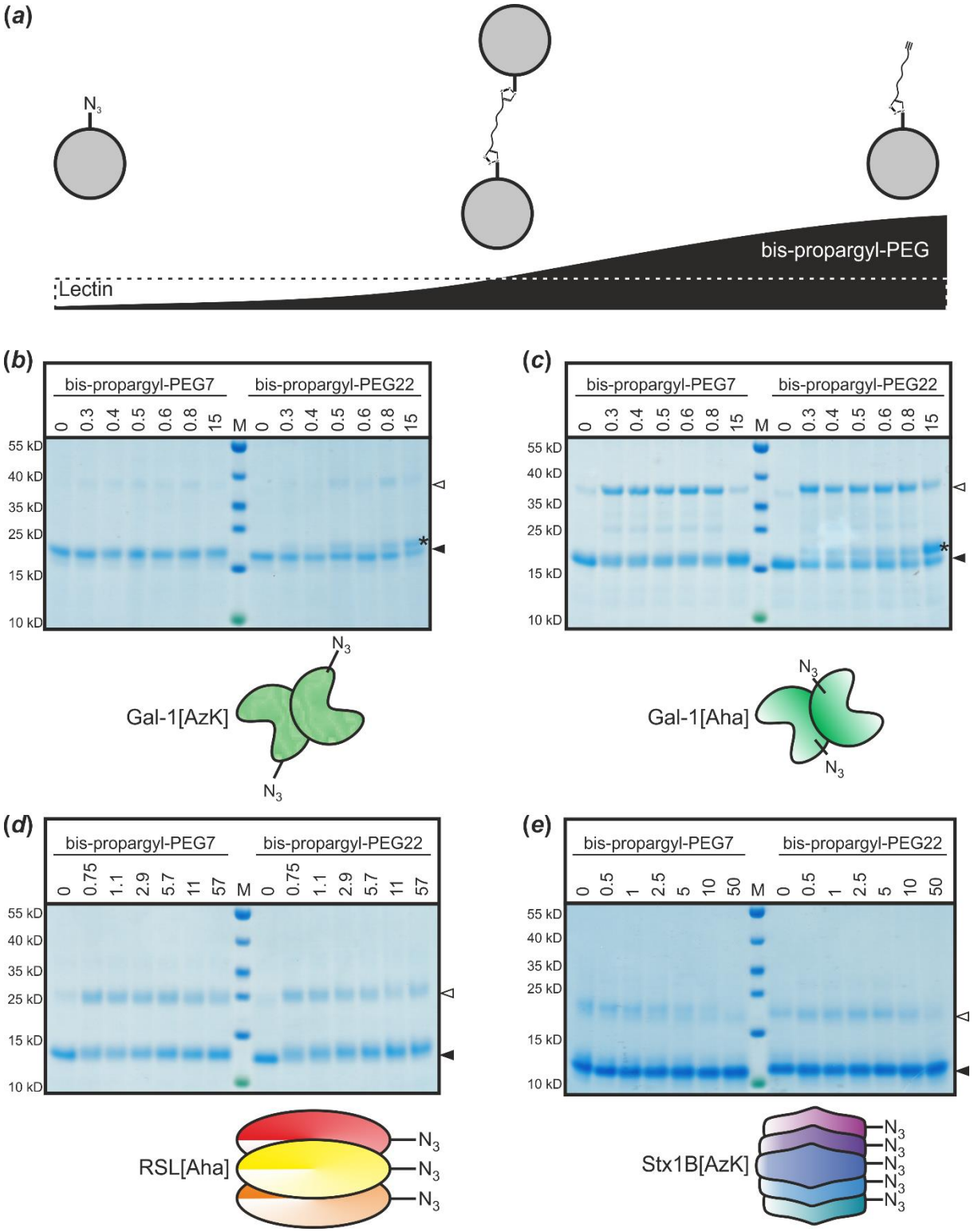

Figure S3 Linker-mediated homo-lectin conjugate formation with the bis-propargyl-PEG linkers.

(*a*) Hypothetical process of linker-mediated conjugation. The lectins contain reactive azido-groups in the absence of bis-propargyl-PEG-linker. Homo-lectin conjugates form when equal molar ratios of alkyne-groups on the linker and azido-groups on the lectins are present. Excess linker molecules provoke the alkyne-labelling of the lectins. (*b-e*) The molar ratio of the reactive groups on the linkers in relation to the reactive groups on the lectins is indicated on the top of the gels. 32 µM of Gal-1[AzK], 26 µM of

Gal-1[Aha], 15 µM of RSL[Aha] and 23.5 µM of Stx1B[AzK] were used in the reactions. The panels (*b*)-(*e*) show the Coomassie-stained SDS-gels for Gal-1[AzK], Gal-1[Aha], RSL[Aha] and Stx1B[AzK], respectively. Black triangles denote lectin monomers, while white triangles label the formed homo-conjugates. Proteins bands corresponding to alkyne-PEG-labelled Gal-1-variants are indicated by asterisks. The sizes of molecular weight marker (M) proteins are indicated on the left.

For the interpretation of the gels the reader is referred to the 'linker-mediated lectin:lectin conjugation' section and figure 4 in the main text. With an excess of PEG-linker over Gal-1-variants (panels *b*,*c*), we were able to label the azido-lectins with free alkyne groups. In these cases, one of the two reactive groups of the PEG-spacers conjugated with the azido-group on the protein, but not with another lectin, converting the azido-lectin to an alkyne-labelled lectin. This process was directly visible on the SDS-gels for reactions with the bis-propargyl-PEG_{22} (panels *b*,*c*, asterisks). The lectin:linker-conjugates ran with a higher apparent molecular weight compared to the unlabelled azido-lectins. We observed about one third (Gal-1[Aha]) to half (Gal-1[AzK]) of the protein to remain in the protein band with the lower molecular weight (panels *b*,*c*, black arrows). As discussed in the main text, this band contains proteins that did not undergo conjugation with the PEG-linker. In the case of Gal-1[AzK], this band might as well contain unidentified protein species of comparable size as observed by mass spectrometry (electronic supplementary material, table S1).