Cross-linking/mass spectrometry (CLMS) is a widespread method for investigating protein structure and interactions. Selective cross-linking chemistry is one factor that limits the current resolution of this approach. In CLMS, new covalent bonds are introduced between atoms near in space but not necessarily close in the sequence of the protein. These linkages are then detected by mass spectrometry and function as distance constraints when modeling the structure. Currently, data density of identified distance constraints is curtailed by the restricted specificity of homobifunctional N-hydroxysuccinimide (NHS) ester cross-linkers predominantly used in CLMS. An earlier study aimed at comparing exogenous bifunctional cross-linkers containing different photoactivatable groups, including diazirine and benzophenone. However, resulting data from using carboxybenzophenone-succinimide was sparse, with only 3 links observed and a total lack of identified links using a diazirine derivative cross-linker. An earlier study, having synthesized a benzophenone derivative cross-linker and its deuterated equivalent, also identified 3 linked residue pairs. In contrast, when using the diazirine derivative sulfoSDA in a previous study, we identified 500 cross-linked residue pairs in purified HSA at 5% FDR. We aimed to establish whether the use of a benzophenone based cross-linker could result in a similarly high number of cross-links and whether cross-linking site selection would be influenced by the different chemical nature of the two photoactivatable groups.

## METHODS

### Chemicals and Materials

The cross-linking reagent sulfoSDA was purchased from Thermo Fisher Scientific (Rockford, IL). HSA (A8763) was purchased from Sigma-Aldrich (St. Louis, MO).

#### Synthesis of 4-(N-Sulfo succinimidylcarboxy) benzophenone

4-Benzoylbenzoic acid (37 mg, 0.16 mmol) used to capture protein interactions in the context of amino acid analogue incorporation, but this has typically involved targeting single incorporation sites, limiting the extent of identified links. A previous study aimed at comparing exogenous bifunctional cross-linkers containing different photoactivatable groups, including diazirine and benzophenone. However, resulting data from using carboxybenzophenone-succinimide was sparse, with only 3 links observed and a total lack of identified links using a diazirine derivative cross-linker. An earlier study, having synthesized a benzophenone derivative cross-linker and its deuterated equivalent, also identified 3 linked residue pairs. In contrast, when using the diazirine derivative sulfoSDA in a previous study, we identified 500 cross-linked residue pairs in purified HSA at 5% FDR. We aimed to establish whether the use of a benzophenone based cross-linker could result in a similarly high number of cross-links and whether cross-linking site selection would be influenced by the different chemical nature of the two photoactivatable groups.

#### Synthesis of 4-(N-Sulfo succinimidylcarboxy) benzophenone

4-Benzoylbenzoic acid (37 mg, 0.16 mmol)
and N-hydroxysulfosuccinimide (50 mg, 0.23 mmol) were stirred in 1 mL of DMF and to this was added EDC-HCl (40 mg, 0.21 mmol) and the mixture stirred overnight. Solvents were removed under reduced pressure and the residue purified using semipreparative RP-HPLC (Agilent 1100 prep-HPLC system equipped with a preparative autosampler (G2260A), preparative scale pumps (G1361A), a fraction collector (G1364B-prep), and a multil wavelength UV detector (G13658 MWD with preparative flow cell)). Material was separated at a flow rate of 20 mL/min on an Agilent RP-C-18 column (21.2 mm × 150 mm, 10 μm particle size), using the following method: 0% B for 5 min, 0% to 80% B in 40 min, 80% B for 5 min, 80% to 0% B in 5 min, 0% B for 5 min, with mobile phase A consisting of water and 0.1% trifluoroacetic acid and mobile phase B consisting of acetonitrile and 0.1% trifluoroacetic acid) to yield the title product as a white solid (20 mg, 0.05 mmol, 31%). 1H-NMR (500 MHz, D2O) to 8.39 (2H, t, J = 8.7 Hz, CH2), 8.26 (1H, m, CH), 7.90–7.66 (5H, m, ArH), 7.61–7.48 (2H, m, ArH), 7.36 (1H, m, CH2). 13C-NMR (125 MHz, D2O) δ 199.2 (C), 170.0 (C), 166.0 (C), 161.9 (C), 142.6 (C), 135.8 (C), 134.1 (CH), 130.5 (CH), 130.4 (CH), 130.3 (CH), 128.7 (CH), 127.6 (C), 56.6 (CH), 30.0 (CH2). MS (ESI) Exact mass calcld for C18H14NO8S [M + H]+, 404.44; found, 404.15.

Cross-Linking HSA. Cross-linking was carried out in triplicate. Aliquots of HSA (15 μg, 0.75 mg/mL) in cross-linking buffer (20 mM HEPES—OH, 20 mM NaCl, 5 mM MgCl2, pH 7.8) were cross-linked with either sulfo-SDA or sulfo-SBP, using eight cross-linker-to-protein ratios (w/w) and UV irradiation times of either 25 or 50 min. In the case of sulfo-SBP, the cross-linker-to-protein ratios (w/w) used were 0.187:1, 0.25:1, 0.375:1, 0.5:1, 0.75:1, 1:1, 1.5:1, and 2:1. The equivalent molar ratios of sulfo-SDA were used, which meant that the cross-linker-to-protein ratios (w/w) used were 0.152:1, 0.203:1, 0.303:1, 0.406:1, 0.606:1, 0.811:1, 1:1:1, and 1:6.2:1. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group. The cross-linking reaction was a two-step process. First the cross-linker and protein mixture was allowed to incubate to form the diazirine group.

Sample Preparation for Mass Spectrometric Analysis. Bands corresponding to monomeric HSA were excised from the gel and the proteins reduced with 20 mM DTT, alkylated using 55 mM IAA and digested with trypsin following standard protocols. The resulting digests were desalted using self-made C18 StageTips prior to mass spectrometric analysis.

Mass Spectrometry and Data Analysis. Peptides were loaded directly onto a spray emitter analytical column (75 μm inner diameter, 8 μm opening, 250 mm length; New Objectives) packed with C18 material (RepriSol-Pur C18-AQ 3 μm; Dr Maisch GmbH, Ammerbuch-Entringen, Germany) using an air pressure pump (Proxeon Biosystems, Odense, Denmark). Mobile phase A consisted of water and 0.1% formic acid. Mobile phase B consisted of acetonitrile and 0.1% formic acid. Peptides were loaded onto the column with 1% B at 700 nL/min flow rate and eluted at 300 nL/min flow rate with a gradient: 1 min linear increase from 1% B to 9% B; linear increase to 35% B in 169 min; 5 min increase to 85% B. Eluted peptides were sprayed directly into a hybrid linear ion-trap-Orbitrap mass spectrometer (LTQ-Orbitrap Velos, Thermo Fisher Scientific). Peptides were analyzed using a “high/high” acquisition strategy, detecting at high resolution in the Orbitrap and analyzing the subsequent fragments also in the Orbitrap. Survey scan (MS) spectra were recorded in the Orbitrap at 100 000 resolution. The eight most intense signals in the survey scan for each acquisition cycle were isolated with an m/z window of 2 Th and fragmented with collision-induced dissociation (CID) in the ion trap. 1+ and 2+ ions were excluded from fragmentation. Fragmentation (MS2) spectra were acquired in the Orbitrap at 7500 resolution. Dynamic exclusion was enabled with 90 s exclusion time and repeat count equal to 1. Mass spectrometric raw files were processed into peak lists using MaxQuant version 1.3.0.530 using default parameters except the setting for “Top MS/MS peaks per 100 Da” being set to 100.

Peak lists were subsequently searched against an HSA sequence database (UniProt P02768) using X! (https://github.com/Rappsilber-Laboratory/X!Search) for identification of cross-linked peptides. Search parameters for sulfo-SDA search were MS accuracy, 6 ppm; MS/MS accuracy, 20 ppm; enzyme, trypsin; specificity, fully tryptic; allowed number of missed cleavages, four; cross-linker, SDA; fixed modifications, none; variable modifications, carbamidomethylation on cysteine, oxidation on methionine, SDA-loop (SDA cross-link within a peptide that is also cross-linked to a separate peptide). Search parameters for sulfo-SBP were the same except: cross-linker, BENZO (C14H8O2); fixed modifications, none; variable modifications, carbamidomethylation on cysteine and oxidation on methionine. The linkage specificity for sulfo-SDA and sulfo-SBP was assumed to be at lysine, serine, threonine, tyrosine, and protein N-termini at one end, with the other end having specificity for any amino acid residue. False discovery rates (FDR) were estimated following a modified target-decoy search strategy. Cross-link results from the FDR analysis with scores at the peptide-spectrum match level and confidence values at cross-link level are also available in Tables S1 and S2. Cross-links were visualized in the crystal structure of HSA (PDB1AO6) using PyMOL (http://www.pymol.org).

To compare the fragmentation behavior of peptides linked with either sulfo-SDA or sulfo-SBP, we extracted 49 common PSMs that were identified with both sulfo-SBP and sulfo-SDA. A common PSM is defined through the sequences of the two peptides in the cross-link, the cross-link sites and the charge state of the precursor. Next, we performed a two-sided Mann–Whitney–U tests to check for differences in the number of b- and y-type fragments (ignoring loss of H2O and NH3). The tests were highly significant (n = 49, p-value ≤0.01 for the b-ions, and ≤0.001 for the y-ions, respectively). The mean difference between the number of b- and y-fragments between sulfo-SDA and sulfo-SBP cross-linked peptides is 4. Next, the fragmentation preferences in terms of fragment ion intensities were compared with an adjusted cosine similarity. Instead of computing the cosine similarity of the complete spectrum only the relative intensities of shared fragment ions were considered. The two distributions are clearly distinct indicating conservation of the preferred bond cleavage during fragmentation.

RESULTS AND DISCUSSION

Herein we report the synthesis and use of a photoactivatable benzophenone derivative cross-linker, 4-
(sulfosuccinimidylcarboxy)benzophenone (sulfo-SBP), and compare its behavior in the reactivity with proteins to that of sulfo-SDA (Figure 1). While fewer links were detected when using sulfo-SBP compared to sulfo-SDA, complementary photochemistry allowed access to additional parts of protein structure, previously inaccessible using the diazirine-based cross-linker only.

We aimed to benchmark sulfo-SBP against the performance of sulfo-SDA. Sulfo-SBP was first synthesized in a one-step coupling procedure from the corresponding acid (Figure 1B). Purified HSA was cross-linked in triplicate using either sulfo-SDA or sulfo-SBP. (D) Purified HSA, with either no (−) cross-linker, reacted with (+) sulfo-SBP or with (+) sulfo-SDA. (E) Unique cross-linked residue pairs identified using only either sulfo-SBP or sulfo-SDA and those common to both.

Using sulfo-SDA, 792 unique residue pairs were identified (supported by 1874 MS spectra) at 5% FDR (using a score cutoff of 5). In contrast, using sulfo-SBP a total of 173 unique residues pairs were identified (supported by 356 MS spectra) at 5% FDR (score cutoff = 5) (Figure 1E), of which 66 were also found using sulfo-SDA and 107 were unique to sulfo-SBP. For comparison, previous work cross-linking acid residues and primary amines using 5 different cross-linkers, was based on an average of 52 links per protein, detected in separate studies.3

Unambiguous linkage site assignment is dependent on adequate flanking fragmentation events (Figure 2). Manual validation of peptide spectra matches revealed this to be the case for 313/792 (40%) sulfo-SBP cross-linked residue pairs

Figure 1. Comparison of sulfo-SDA and sulfo-SBP reaction with HSA. (A) Chemical structures of sulfo-SDA and sulfo-SBP. (B) Synthesis of sulfo-SBP. (C) Experimental design. HSA was cross-linked in triplicate using either sulfo-SDA or sulfo-SBP. (D) Purified HSA, with either no (−) cross-linker, reacted with (+) sulfo-SBP or with (+) sulfo-SDA. (E) Unique cross-linked residue pairs identified using only either sulfo-SBP or sulfo-SDA and those common to both.

Figure 2. Fragmentation analysis. (A and B) LC/ESI-MS/MS spectra of the same peptide pair cross-linked with either (A) sulfo-SBP or (B) sulfo-SDA. (C and D) Comparison of b- and y-ion fragmentation of common sulfo-SDA and sulfo-SBP cross-linked peptide pairs. (C) A Mann–Whitney U test indicates highly significant changes in the number of fragments ($n = 49$, **p-value ≤ 0.01 for b-ions, and *** ≤ 0.001 for y-ions, respectively). (D) Fragmentation spectra similarity computed as a simplified cosine similarity and compared with a random reference distribution.
and 55/173 (32%) sulfo-SBP cross-linked residue pairs. Looking further at cross-linked peptide pairs common to both sulfo-SDA and sulfo-SBP revealed differences in the numbers of fragments between cross-linkers (Figure 2C); however, cosine similarity was high, suggesting the impact was minimal (Figure 2D).

Cross-link coverage is even across the structure of HSA for both cross-linkers (Figure 3). Very few links have Ca-Ca distances in excess of 25 Å, 5.6% for sulfo-SDA (44/776 residue pairs) and 2.9% for sulfo-SBP (5/172 residue pairs). Although sulfo-SDA has a shorter calculated backbone distance compared to sulfo-SBP (3.9 Å compared to 5.7 Å as predicted for sulfo-SBP), the data suggests that sulfo-SBP has a shorter upper limit distance constraint than sulfo-SDA. This is also demonstrated by 94% of sulfo-SBP linked residue pairs having Ca-Ca distance less than 20 Å, compared to 86% of sulfo-SDA residue pairs. The number of identified linked residue pairs in excess of these proposed upper limits are in agreement with the 5% FDR applied to identifications. Sulfo-SBP data fits well to the crystal structure of HSA (PDB1AO6), corroborating with a previous study that showed benzophenone incorporation into proteins is well tolerated.36 While HSA does have extensive disulfide bridges, it is not incapable of changing its structure in response to changes in its surface charge. Changing the pH does result in conformational changes as does protein oxidation.38,39 The close agreement of cross-link data with the crystal structure suggests that the reduction of charged residues during the initial anchoring step of the reagent does not noticeably disrupt the native fold of the protein.

Both cross-linkers are heterobifunctional, with cross-linking occurring as a two-step reaction: as an anchoring step, the sulfo-NHS ester component of each cross-linker reacts, primarily with the amine groups of lysine residues and protein N-termini, but also with the hydroxyl groups of serine, threonine, and tyrosine. Second, following UV activation, the photoactivatable groups react by insertion into available N−H/C−H bonds at any residue (photo-cross-linking step). Reactivity can therefore be characterized in terms of both the anchoring and photoreacted sites. Distinct anchoring residue hotspots within HSA account for up to 46 unique residue pairs for a single hotspot. In total, the top eight anchoring hotspots (K161, Y162, K183, K186, K214, K223, K456, and K549) account for 26% of all identified residue pairs (Figure S1).

Fewer anchoring lysine residues are observed with sulfo-SBP than sulfo-SDA at 23/59 and 45/59 lysine residues, respectively (Figure S1). However, linear peptides with sulfo-SBP and sulfo-SDA modifications (reacted at the NHS-ester) were observed for 55 and 57 (out of 59) lysine residues, respectively, suggesting that the observation of cross-links is not limited by the anchoring reaction but must depend on the photoreaction. Both cross-linkers show distinct regional preferences in the photoreaction step (Figure 3D,E). This is marked in the case of anchoring residues K223, K236, K565 and residues K161, K161, and Y162 (Figure 3D). Although all these residues have reacted with both cross-linkers, the reaction has proceeded in opposing directions. Lysine K565, in particular, was linked to 12 different residues, 5 of which by sulfo-SDA and 7 by sulfo-SBP. Sulfo-SDA linked K565 to residues in a single loop. In contrast, sulfo-SBP linked K565 to residues in 4 separate

Figure 3. Sterical analysis of cross-links. (A) Observed cross-links plotted in PDB1AO6. (B) Ca–Ca distogram of observed cross-links, compared with a random distance distribution (PDB1AO6). (C) Matrix plot of residue pairs across the sequence of HSA. (D and E) Zoomed in regions of HSA in PDB1AO6 with cross-linked residue pairs involving NHS-ester reactive residue hot-spots. (D) Photoreaction exhibits a cross-linker dependent regional preference. Top left, K223. Top right, K565. Bottom left, K236. Bottom right, K160, K161, Y162; (E) Top panel, K375. Bottom panel, K183, Y185 and K186. Sulfo-SDA (red), sulfo-SBP (blue) and common to both cross-linkers (yellow). Figures were generated using PyMOL (http://www.pymol.org).
neighboring alpha helices. In the case of anchoring residue K375 and residues K183, Y185, and K186 (Figure 3E), opposite cross-linking orientation is also evident. The difference here being that there is also considerable overlap, with some degree of linker specificity.

There is an observed cross-linking bias for particular residue types, which are different for sulfo-SDA and sulfo-SBP (Figure S2). Both cross-linkers disfavor polar residues (sulfo-SBP more strongly than sulfo-SDA); however, sulfo-SDA has a preference for acidic residues (Asp, Glu), which are strongly disfavored by sulfo-SBP. Notably, we identify a bias between the benzophenone of sulfo-SBP toward methionine residues, which corroborates a previous study demonstrating a “Magnet Effect” of methionine residues.40 The benzophenone of sulfo-SBP has a stronger preference for linking hydrophobic residues than sulfo-SDA, in particular glycine and phenylalanine residues.

Photo-cross-linking has emerged as a means to greatly increase the density of cross-linking/mass spectrometry data, which allows interrogation of protein structures in very complex environments.15 Alternative photo-cross-linking chemistry may become routinely tuned for physicochemical characteristics. Ultimately, cross-linking/mass spectrometry chemistry may become routinely tuned for probing a likely structure and to determine protein folds.

**ACKNOWLEDGMENTS**

The Wellcome Trust generously funded this work through a Senior Research Fellowship to J.R. (Grant 103139), a Centre Core Grant (Grant 203149), and an Instrument Grant (Grant 108504).

**REFERENCES**


**ASSOCIATED CONTENT**

1 Supporting Information
The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.6b04938.

H and 13C NMR spectra for 4-[(N-sulfosuccinimidylcarboxy) benzophenone, gel images, replica analysis and site-selection differences following photoactivation of sulfo-SDA and sulfo-SBP (PDF)

List of peptide spectrum matches for Sulfo-SBP and Sulfo-SDA (5% FDR) (ZIP)

**AUTHOR INFORMATION**

Corresponding Author
*E-mail: juri.rappisilber@ed.ac.uk. Phone: +44-(0)131-651-7056. Fax: +44-(0)131-650-5379.

ORCID
Adam Belsom: 0000-0002-8442-4964
Sven Giese: 0000-0002-9886-2447
Manfred Auer: 0000-0001-8920-3522
Jurí Rappisilber: 0000-0001-5999-1310

**Notes**
The authors declare no competing financial interest. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium35 (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the data set identifier PXD004920.