THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes

Citation for published version:
Chen, Z, Fischer, L, Tahir, S, Bukowski-Wills, J-C, Barlow, PN & Rappsilber, J 2016, 'Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes: QCLMS reveals protein conformational changes' Wellcome Open Research . DOI: 10.12688/wellcomeopenres.9896.1

Digital Object Identifier (DOI):
10.12688/wellcomeopenres.9896.1

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Wellcome Open Research

Publisher Rights Statement:
The data referenced by this article are under copyright with the following copyright statement: Copyright: © 2016 Chen Z et al.
KT

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. Apr. 2019

Quantitative cross-linking/mass spectrometry reveals subtle protein conformational changes

Zhuo A. Chen[1], Lutz Fischer[1], Salman Tahir[1], Jimi-Carlo Bukowski-Wills[1], Paul N. Barlow[2], Juri Rappsilber[1,3*]

[1] Wellcome Trust Centre for Cell Biology, Institute of Cell Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3BF, UK

[2] Schools of Chemistry and Biological Sciences, University of Edinburgh, Edinburgh EH9 3JJ, UK

[3] Department of Bioanalytics, Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany

Correspondence should be addressed to Juri Rappsilber.

Email: juri.rappsilber@ed.ac.uk

Phone: +44-(0)131-651-7056, FAX: +44-(0)131-650-5379

Running title: QCLMS reveals protein conformational changes

Abbreviations

BS^3 - Bis[sulfosuccinimidyl] suberate

CLMS - Cross-linking/mass spectrometry

FDR - False discovery rate

HCD - Higher energy collision induced dissociation

LC-MS/MS - Liquid chromatography tandem mass spectrometry

LTQ - Linear trap quadrupole

MS2 - Tandem mass spectrometry

QCLMS - Quantitative cross-linking/mass spectrometry

SCX - Strong cation exchange

Summary

We have developed quantitative cross-linking/mass spectrometry (QCLMS) to interrogate conformational rearrangements of proteins in solution. Our workflow was tested using a structurally well-described reference system, the human complement protein C3 and its activated cleavage product C3b. We found that small local conformational changes affect the yields of cross-linking residues that are near in space while larger conformational changes affect the detectability of cross-links. Distinguishing between minor and major changes required robust analysis based on replica analysis and a label-swapping procedure. By providing workflow, code of practice and a framework for semi-automated data processing, we lay the foundation for QCLMS as a tool to monitor the domain choreography that drives binary switching in many protein-protein interaction networks.

Introduction

Domain rearrangements of individual proteins act as molecular switches, govern the assembly of complexes and regulate the activity of networks. A typical example of a predominantly conformational change-driven finely tuned protein-protein interaction network is the mammalian complement system of ~40 plasma and cell-surface proteins. This is responsible for clearance of immune complexes from body fluids along with other hazards to health including bacteria and viruses. The study of such networks is hampered by the lack of tools to follow dynamic aspects of protein structures.

Cross-linking combined with mass spectrometry and database searching is a tool that can reveal the topology and other structural details of proteins and their complexes (1-3). It is currently unclear if dynamic information could also be obtained from such straight applications of cross-linking/mass spectrometry (CLMS) analysis. However, dynamic information could come from a comparison of the cross-links obtained from different protein states. We are developing a strategy to perform such a comparison in a rigorous manner by quantitative CLMS (QCLMS) (4) using isotope-labeled cross-linkers (5-8). In our strategy, distinct isotopomers of a cross-linker are used, to cross-link different stable conformers of the same, or very similar, protein molecules. The mass spectrometric signals of cross-linked peptides derived from different conformations can then be distinguished by their masses and thus quantified and related to conformational differences (Fig. 1A).

Despite being straightforward in principle, quantifying cross-linked peptides is technically challenging. Currently available quantitative proteomics software is linked to protein identification workflows meaning that cross-link analysis is not routinely possible. Previously performed proof-of-principle work relied on an elementary computational tool, XiQ (4). A subsequent application (9) relied on manual data analysis. These works neglected to include

replicated analysis and label-swaps; which are important standard procedures of quantitative proteomics. Recently, a workflow using replicated analysis and software mMass has been demonstrated using a calmodulin (17kDa) in presence and absence of Ca2+ (10). It remains to be seen if the approach scales to larger protein system. Additional experimental challenges derive from cross-linking itself. Cross-linking is known for low reproducibility. In any case, the signal intensity of an individual cross-linked peptide does not just depend on cross-linking efficiency. For example, additional variable chemical modifications of the peptides, such as methionine oxidation and reaction with the hydrolyzed cross-linker, may also depend on protein conformation. Furthermore, the network of cross-links that form within the molecule may influence the efficiency of protease digestion (10). Finally, peak interference or stochastic processes during data acquisition may prevent accurate quantitation. This is a general problem for quantification but it may especially affect cross-linking; cross-links are generally sub-stoichiometric and thus of low signal intensity in mass spectra.

To explore solutions to the challenges of QCLMS we have investigated here key proteins in the complement system, a central player in human innate immune defenses. As the pivotal activation step of the complement system, C3 convertases excise the small anaphylatoxin domain (ANA) from the complement component C3 (184 kDa) leaving its activated form, C3b (175 kDa) (Fig. 1). Both C3 and C3b are stable, and comparisons of the crystal structures of both (C3 (11) and C3b (12)) (Fig. 1) revealed details of the structural rearrangements during this conversion. Based on this comparison of C3 and C3b, we have been able to develop a workflow and tested the reliability and usefulness of QCLMS. Our workflow involves replicated analysis, label-swap and some level of automation of the quantitation process, here utilizing the quantitative proteomics software tool Pinpoint (Thermo Scientific). In this way it proved possible to infer the conformational changes that accompany

cleavage of C3 to form C3b and then compare these to the difference between the two crystal structures. This both allowed cross-validation of the existing structures and revealed details of the relationship between cross-linking yields and conformational changes. Based on our experiences, we suggest a code of practice for the use of QCLMS in the study of protein dynamics.

Experimental Procedures

Protein preparation for cross-linking

Plasma-derived human C3 and C3b were purchased from Complement Technology, Inc. USA (and stored at -80 °C). Native C3 was depleted of low amounts of contaminating iC3 using cation-exchange chromatography (13). Chromatography was performed using a Mini S PC 3.2/3 column (GE Healthcare) at a flow-rate of 500 µl/min at 4 °C. Immediately after purification, C3 and C3b samples were exchanged, using 30-kDa molecular weight cutoff (MWCO) filters (Millipore), into cross-linking buffer (20 mM HEPES-KOH, pH 7.8, 20 mM NaCl and 5 mM $MgCl_2$) with a final concentration of 2 µM.

Protein cross-linking

50 µg C3 and C3b were each, separately, cross-linked with either bis[sulfosuccinimidyl] suberate (BS^3) (Thermo Scientific) or its deuterated analogue bis[sulfosuccinimidyl] 2,2,7,7-suberate-d4 (BS^3-d4) (Fig. 1C), at 1:3 protein to cross-linker mass ratios, giving rise to four different protein-cross-linker combinations: C3+BS^3, C^3+BS3-d4, C3b+BS^3 and C3b+BS^3-d4. After incubation (two hours) on ice, reactions were quenched with 5 µl 2.5 M ammonium bicarbonate for 45 minutes on ice. For the monitoring of cross-linked products, an aliquot containing 5 pmol of cross-linked protein from each of the above four reaction mixtures were subjected to SDS-PAGE using a NuPAGE 4-12% Bis-Tris gel and MOPS running buffer (Thermo Fisher Scientific) according to the manufacturer's instructions. The protein bands were visualized using the Colloidal Blue Staining Kit (Thermo Fisher Scientific) according to the manufacturer's instructions.

Sample preparation for mass spectrometric analysis

Native C3 and C3b molecules each contain two, disulfide-linked, polypeptide chains. The monomeric (two-polypeptide chain) products of cross-linked C3 and C3b were resolved as distinct bands using SDS-PAGE. The products were in-gel reduced and alkylated, then digested using trypsin following a standard protocol (14). For quantitation, equimolar quantities of the tryptic products from the four cross-linked protein samples were mixed pair-wise, yielding two combinations: C3+BS^3 and C3b+BS^3-d4; C3+BS^3-d4 and C3b+BS^3.

From each of these two samples, a 20-μg (40 μL) aliquot was taken and fractionated using SCX-Stage-Tips (15) with a small variation of the protocol previously described for linear peptides (16). In short, peptide mixtures were first loaded on a SCX-Stage-Tip in loading buffer (0.5% v/v acetic acid, 20% v/v acetonitrile, 50 mM ammonium acetate). The bound peptides were eluted in two steps, with buffers containing 100 mM ammonium acetate and 500 mM ammonium acetate, producing two fractions. These peptide fractions were desalted using C18-StageTips (17) prior to mass spectrometric analysis. In parallel, a 4-μg (10 μL) aliquot of each sample was desalted using C18-Stage-Tips for mass spectrometric analysis without pre-fractionation.

Mass spectrometric analysis

Analysis by LC-MS/MS was performed using two instruments, both applying a “high-high” acquisition strategy. SCX-Stage-Tip fractions were analyzed using a hybrid linear ion trap-Orbitrap mass spectrometer (LTQ-Orbitrap Velos, Thermo Fisher Scientific). Peptides were separated on an analytical column that was packed with C18 material (ReproSil-Pur C18-AQ 3 μm; Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) in a spray emitter (75-μm inner

diameter, 8-µm opening, 250-mm length; New Objectives, Woburn, MA, USA) (18). Mobile phase A consisted of water and 0.5% v/v acetic acid. Mobile phase B consisted of acetonitrile and 0.5% v/v acetic acid. Peptides were loaded at a flow-rate of 0.7 µl/min and eluted at 0.3 µl/min using a linear gradient going from 3% mobile phase B to 35% mobile phase B over 130 minutes, followed by a linear increase from 35% to 80% mobile phase B in five minutes. The eluted peptides were directly introduced into the mass spectrometer. MS data were acquired in the data-dependent mode. For each acquisition cycle, the mass spectrum was recorded in the Orbitrap with a resolution of 100,000. The eight most intense ions with a precursor charge state 3+ or greater were fragmented in the linear ion trap by collision-induced disassociation (CID). The fragmentation spectra were then recorded in the Orbitrap at a resolution of 7,500. Dynamic exclusion was enabled with single repeat count and 60-second exclusion duration.

Non-fractionated peptide samples were analyzed using a hybrid quadrupole-Orbitrap mass spectrometer (Q Exactive, Thermo Fisher Scientific). Peptides were separated on a reversed-phase analytical column of the same type as described above. Mobile phase A consisted of water and 0.1% v/v formic acid. Mobile phase B consisted of 80% v/v acetonitrile and 0.1% v/v formic acid. Peptides were loaded at a flow rate of 0.5 µl/min and eluted at 0.2 µl/min. The separation gradient consisted of a linear increase from 2% mobile phase B to 40% mobile phase B in 169 minutes and a subsequent linear increase to 95% B over 11 minutes. Eluted peptides were directly sprayed into the Q Exactive mass spectrometer. MS data were acquired in the data-dependent mode. For each acquisition cycle, the MS spectrum was recorded in the Orbitrap at 70,000 resolution. The ten most intense ions in the MS spectrum, with a precursor change state 3+ or greater, were fragmented by Higher Energy Collision Induced Dissociation (HCD). The fragmentation spectra were thus recorded in the Orbitrap at 35,000 resolution. Dynamic exclusion was enabled, with single-repeat count and a 60 second exclusion duration.

Identification of cross-linked peptides

The raw mass spectrometric data files were processed into peak lists using MaxQuant version 1.2.2.5 (19) with default parameters, except that “Top MS/MS Peaks per 100 Da” was set to 20. The peak lists were searched against C3 and decoy C3 sequences using Xi software (ERI, Edinburgh) for identification of cross-linked peptides. Search parameters were as follows: MS accuracy, 6 ppm; MS2 accuracy, 20 ppm; enzyme, trypsin; specificity, fully tryptic; allowed number of missed cleavages, four; cross-linker, BS^3/BS^3-d4; fixed modifications, carbamidomethylation on cysteine; variable modifications, oxidation on methionine, modifications by BS^3/BS^3-d4 that are hydrolyzed or amidated on the other end. The reaction specificity for BS^3 was assumed to be for lysine, serine, threonine, tyrosine and protein N-termini. Identified candidates for cross-linked peptides were validated manually in Xi, after applying an estimated FDR of 3% for cross-linked peptides (Fischer and Rappsilber, submitted MCP/2015/049502). We used 3%FDR as this was the best FDR that returned reasonable number of decoys to provide a meaningful FDR. All identified cross-linked peptides are listed in Supplemental Table S1 and for each identified cross-linked peptide, the annotated best-matched MS2 spectrum is provided in Supplemental File S1.

Quantitation of cross-link data using Pinpoint software

Identified cross-linked peptides were quantified based on their MS signals. The quantitative proteomics software tool Pinpoint (Thermo Fisher Scientific) was used to retrieve intensities of both light and heavy signals for each cross-linked peptides in an automated manner. An input library for cross-linked peptides was constructed according to “General Spectral Library Format for Pinpoint Comma Separated Values” (Thermo Fisher Scientific) (Supplemental File S1). In the input library, the sequence of every cross-linked peptide was converted into a linear version

with identical mass (Fig. 2A), following an idea from Maiolica et al. (13). Six modifications that were not listed in Pinpoint modification list were defined in “customized modifications”:

[BS3-OH]; mass, 156.07860; Site modified, KSTY

[BS3-NH2]; mass, 155.09460; Site modified, KSTY

[BS3d4-OH]; mass, 160.10374; Site modified, KSTY

[BS3d4-NH2]; mass, 159.11972; Site modified, KSTY

[Xlink]; mass, 27.98368; Site modified, K

[Xlinkd4]; mass, 32.00878; Site modified, K

The three most abundant of the first four signals in the isotope envelope were used for quantitation. The error tolerance for precursor m/z was set to 6 ppm. Signals are only accepted within a window of retention time (defied in spectral library) ±10 minutes. Manual inspection was carried out to ensure the correct isolation of elution peaks. For each cross-linked peptide, elution peak areas of the light and the heavy signals were measured and reported as log2 (C3b/C3). “Match between runs” (20) was carried out for all cross-linked peptides in Pinpoint interface manually, based on high mass accuracy and reproducible LC retention time, hence quantitation was conducted for each identified cross-linked peptides in both forward-labeled and reverse-labeled samples even when they were only identified in one of them. Within each quantitation sample, signal fold-changes of all quantified cross-linked peptides were normalized against their median, in order to correct systematic error introduced by minor shift on mixing ratio during sample preparation.

The quantitation data was subsequently summarized at the level of unique residue pairs (cross-links). A cross-link was defined as a unique cross-link in either C3 or C3b only if all its supporting cross-linked peptides were observed as the corresponding singlet signals. Otherwise, cross-links were regarded as having being observed in both conformations. For a cross-link

shared by C3 and C3b, the signal fold-change was defined as the median of all its supporting cross-linked peptides. Only those cross-links that were consistently quantified in both paired replica analyses (with label-swap) were accepted for subsequent structural analysis. Singlet cross-links were further confirmed by a mass shift of 4 Da, resulting from the label-swap. For cross-links observed as a doublet, the average of signal fold-changes from replicated analyses was reported (Fig 3C). All quantified cross-links are listed in Supplemental Table S2. Cross-links that were significantly enriched in C3 or C3b were determined using “Significance A” test from the standard proteomics data analysis tool Perseus (version 1.4.1.2) (19) based on log2(C3b/C3) values. The following parameters were used in Perseus for the test: “Side”: both; “Use for truncation”: P value; “Threshold value”: 0.05.

Comparison with crystal structures

To compare cross-linking data with X-ray crystallographic data, cross-links were displayed using PyMol (version 1.2b5) (21) in the crystal structures of C3 (PDB|2A73) and C3b (PDB|2I07). Cross-links were represented as strokes between the C-α atoms of linked residues. In the case of a residue missing from the crystal structures, the nearest residue in the sequence was used for display purposes. The distance of a cross-linked residue pair in the crystal structures was measured between the C-α atoms. Measured distances of linked residue pairs in crystal structures were compared to a theoretical cross-linking limit, which was calculated as side-chain length of cross-linked residues plus the spacer length of the cross-linker. An additional 2 Å was added for each residue as allowance for residue displacement in crystal structures. The following side-chain lengths were used for the calculation: 6.0 Å for lysine, 2.4 Å for threonine, 2.4 Å for serine and 6.5 Å for tyrosine. For example, for a lysine-lysine cross-link, the theoretical cross-linking limit is 27.4 Å.

Accession codes

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (22) (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository with the dataset identifier PXD001675.

Reviewer account details:

Username: reviewer31036@ebi.ac.uk

Password: CWaMcjvi

Results

Quantitative cross-linking/mass spectrometry (QCLMS) of C3 and C3b in solution

To assess the abilities of QCLMS to reflect conformational changes, we studied the structurally well-characterized differences between C3 (PDB|2A73) and its activated cleavage product C3b (PDB|2I07). We cross-linked both purified proteins in solution using bis[sulfosuccinimidyl] suberate (BS^3) and its deuterated analogue BS^3-d4 (5-7) (Fig 1C) in four distinct protein-cross-linker combinations, generating C3+BS^3, C3+BS^3-d4, C3b+BS^3 and C3b+BS^3-d4. BS^3 is a homo-bifunctional cross-linker containing an amine-reactive *N*-hydroxysulfosuccinimide (NHS) ester at each end. It reacts primarily with the ε-amino groups of lysine residues and the amino-terminus of a protein, however it can also react with hydroxyl groups of serine, tyrosine and threonine residues (23). Native C3 and C3b are each composed of two polypeptides chains (β and α for C3, β and α' for C3b) connected by a disulfide bond. In all four cross-linking reactions, the two polypeptide chains of C3/C3b were efficiently cross-linked by BS^3 or BS^3-d4, and the resultant two-chain products of cross-linking could be isolated by SDS-PAGE (Fig. 1D) and subjected to in-gel trypsin digestion (14). A 1:1 mixture of C3+BS^3 and C3b+BS^3-d4 digests along with a "label-swapped" replica (C3+BS^3-d4 and C3b+BS^3 digests) were analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (24). Cross-linked peptides were identified by database searching and then quantified based on their MS signals (Fig. 1E).

Wherever C3 and C3b have regions of identical structure, cross-linked peptides should be seen in mass spectra as 1:1 doublet signals (separated by 4 Da). In contrast, where C3 and C3b structures differ, unique cross-links may occur and these will lead to singlets in mass spectra (Fig. 1A). Signal interferences, extended isotope envelopes or experimental variations

may affect quantification. Experimental variations are addressed by replica analysis. Swapping the use of labels in the replica addresses problems arising from extended isotope envelopes and signal interference. Importantly, the label-swap prevents singlet signals of BS^3 cross-linked (light) peptides from being mistakenly quantified as doublets (Fig. 3A). Swapping the use of labels also assists in the identification of cross-links. Singlet signals move by 4 Da, while doublet signals remain as doublets. This behavior of signals corresponding to cross-linked peptides is also clearly distinct from those of peptides contain no cross-linked amino acid residues (Fig. 3A).

We conducted strong cation exchange (SCX)-based fractionation to increase the number of identifiable cross-linked peptides (Fig. 4A). Only 19% of unique cross-linked peptide pairs (50/257) eluted in more than a single fraction. For those 50 cross-linked peptides that were identified and quantified in two sequential SCX fractions, we observed that the C3b/C3 signal ratios were highly reproducible (*R2=0.997*, Fig. 4B). We conclude that SCX appears to be compatible with QCLMS and a worthwhile step in the procedure while cross-linked peptides can be quantified with high technical reproducibility.

In an ideal case, the ratio of the component signals of a doublet would simply be a function of the relative likelihood of linking a particular residue pair in each of the two different protein conformations. However, we observed that different cross-linked peptide pairs that contained the same pair of cross-linked residues, did not share the same component signal ratios. For example, for the seven different cross-linked peptides that covered the residue pair Lys1 and Lys646 (Fig. 3B), the (C3b/C3) signal ratio varied between 0.1 and 2.5. This by far exceeds the variation seen in technical replicates. It may arise, for example, from the impact of conformation on proteolysis cleavage efficiencies. To translate data on multiple peptide pairs containing the same cross-linked residues into a single data point for that residue pair, we took

the median ratio [i.e. log2(C3b/C3)] of all supporting cross-linked peptide pairs for each residue pair. This is equivalent to the procedure used for the inference of protein ratios from peptide ratios in standard quantitative proteomics.

Semi-automated quantitation for cross-linked peptides using Pinpoint

A practical challenge for quantifying cross-link data is how to capture efficiently the signal intensities of cross-linked peptides from a large quantity of mass spectrometric raw data. Available quantitative proteomics software does not accommodate cross-linked peptides (4). A previous protocol for QCLMS resorted to manual quantification (25), a time-consuming and error-prone process. To alleviate this we developed a semi-automated workflow using the quantitative proteomics software package, Pinpoint. The resultant workflow utilizes the well-established functionality of Pinpoint to retrieve intensities of both light and heavy version of every cross-linked peptide in an automatic manner. Furthermore, the user interface of Pinpoint provides, when necessary, a platform for visualizing and validating the quantitation results of any chosen cross-linked peptides. Hence, improvements to the accuracy of quantitation by introducing knowledge-based expertise can be achieved easily and rapidly.

Pinpoint is normally restricted to work with single peptides. Therefore it was necessary to convert our cross-linked peptide pairs into a “linearized” single-peptide format, following a previous approach developed for database searches (14) (Fig. 2A). This allowed generation of a tailored input library of cross-linked peptides (Fig. 2B) based on “General Spectral Library Format for Pinpoint” (Supplemental File S1). In general, Pinpoint requires a very basic set of information to quantify a cross-linked peptide: amino acid sequence, precursor charge state, retention time, and chemical modifications. A consequence of this minimal requirement of

information is that quantification can be performed in Pinpoint independently of the method used for identifying the cross-linked peptides in the first place.

Pinpoint calculates the theoretical m/z of each converted (“linearized”) cross-linked peptide and identifies its signal (within 6 ppm error tolerance in our case) in the raw MS data. Although isotope-related shifts in retention time often occur, in most cases, the retention time of the BS^3 and BS^3-d4 cross-linked versions of a cross-linked peptide in LC-MS/MS overlap to some extent. In order to accurately define singlet and doublet signals of cross-linked peptides, we programmed Pinpoint to retrieve intensity information of both light and heavy signals for each cross-linked peptide by including both a BS^3 cross-linked and a BS^3-d4 cross-linked version of them in the input library.

As discussed above, accurate quantitation of a cross-linked peptide relies on consistent read-out from both replicates of the label-swap analysis. This is especially important for cross-linked peptides that are unique in one of the two conformations. However, cross-linked peptides were not always fragmented and identified in each replica. Such “under sampling” of signals is common in shotgun proteomics and likely exasperated due to the generally low abundance of cross-linked peptides. The Pinpoint interface also provides a platform to conduct “Match between runs” (20). The high mass accuracy achieved in the analysis of the high-resolution Orbitrap data and reproducible LC retention time facilitated this transfer of peptide identities. Thus, all identified cross-linked peptide were quantified in both replica experiments even though they were not necessarily identified in both replicas.

The aggregated intensity of each cross-linked peptide was calculated from the summed area of the three most intense peaks among the first four peaks in the isotope envelope. In cases where a cross-linked peptide was identified with more than one charge state, intensities

derived from the different charge states were automatically combined by Pinpoint. Pinpoint did not select correctly the start or end of an elution peak in every instance. Manual curation was therefore still necessary, albeit largely supported by the Pinpoint interface. In addition, cross-linked peptides were discarded for quantitation if they did not have proper signals for the first three peaks of their isotope envelopes.

Consequently, for each cross-linked peptide, Pinpoint provided intensities for both light (BS^3) and heavy (BS^3-d4) signals. The results were exported into a .csv file. The subsequent processes, for summarizing the results of quantitation for each residue pair, and then combining the outcomes of quantitation for the label-swapped duplicates, were carried out using a Microsoft Excel spreadsheet.

Cross-linking confirms in solution, C3 to C3b conformational transition

In our comparative analysis of C3 and C3b, we quantified signals corresponding to 104 linked residue pairs (cross-links) (Fig. 5A). The 104 cross-links could be divided into three groups: 31 C3-unique cross-links, 24 C3b-unique cross-links and 49 cross-links observed in both proteins. When the “Significance A” test from the standard proteomics data analysis tool Perseus (version 1.4.1.2) (19) was applied to the 49 cross-links observed in both C3 and C3b, based on their log2(C3b/C3) values, three subgroups appeared: 37 showed no significant change (highlighting structural features that are unaffected by cleavage of C3 to C3b), four linked residue pairs were significantly enriched ($p<0.05$) in C3b and eight linked residue pairs were significantly enriched in C3 ($p<0.05$); these two subgroups potentially reflect minor conformational changes. Importantly, the cross-links in these quantified groups and subgroups were not randomly distributed in the 3D structures of C3 (PDB|2A73) (11) and C3b (PDB|2I07) (12) (Figs. 5B, C, D, E).

The linkage pattern revealed structural differences and similarities between C3 and C3b that are in agreement with the crystal structures of the two proteins. Conversion from C3 to C3b is triggered by proteolytic cleavage of the 7-kDa anaphylatoxin (ANA) domain from the N-terminus of the alpha chain. It is therefore not surprising that 20 out of the 31 residue pairs that were found to be cross-linked only in C3 involved ANA. Six of these 20 residue pairs included ANA residues as one partner and residues of MG3, MG8 and TED as the other. Thus these C3-exclusive ANA-specific cross-links clearly define a spatial location for ANA in the C3 molecule that is wholly consistent with the crystal structure (Fig. 5C). The remaining 84 cross-links (11 unique to C3, 24 unique to C3b, and 49 mutual ones) (Figs. 5B, C, D E), report on the extent and nature of rearrangements of the 12 non-ANA domains, namely C345C, TED, CUB, LNK and eight MG domains.

The data confirmed that the spatial arrangement of the domains of the β-chain is conserved following cleavage of the alpha chain: only two cross-links unique to either C3 or C3b involved residues in this chain, compared to 33 C3-unique or C3b-unique cross-links for α-chain domains. The cross-links in the β-chain that are conserved between C3 and C3b occur within (13 cross-links) or between (10 cross-links) its seven intact domains (six MG domains and the LNK domain). In contrast, the remaining domains of the α-chain appears to rearrange extensively following ANA excision: A total of 28 of the 33 C3-unique or C3b-unique cross-links in the α-chain are between domains and only five are within domains; this compares to the identification of nine out of the 12 preserved (not unique to C3 or C3b) α-chain cross-links within domains. In essence, all but one of the domains remain largely unaltered in structure following activation of C3, despite their arrangement changing. An exception is a set of four C3b-unique cross-links identified within MG8. This suggests that conformational changes occur within MG8, to a greater extent than in the other MG domains, following C3 cleavage

Inspection of pairs of amino acid residues that were cross-linked exclusively either in C3 or in C3b illuminates three major rearrangements in the α-chain that accompany the conversion of C3 to C3b. These observations, made in solution, strongly agree with the conformational changes inferred from comparing the crystal structures of C3 (PDB 2A73) and C3b (PDB 2I07) (Fig. 5C, D; Supplemental Table S2). First, residues within the α'- N-terminal (NT) segment (the region of the C3b α-chain that becomes the new terminus after ANA cleavage) are involved in four C3-unique cross-links and six C3b-unique cross-links. These cross-links captured, in C3, the proximity of the α-NT segment to ANA and MG8, In C3b the cross-links confirmed re-location of α'-NT from the MG8/MG3 side to the opposite, C345C/MG7, side of the structure (12). Second, migration of the CUB and TED domains is reflected by two subsets of cross-links. Five cross-links from TED to MG8, MG7, ANA and MG2 support the location of TED at "shoulder-height" as observed in the traditional view of the crystal structure of C3 (Fig. 5C) (11). In C3b, these cross-links were no longer detectable and had been effectively replaced by five TED-MG1 cross-links and a CUB-MG1 cross-link, locating the TED at the "bottom" of the β-chain key ring (12) structure (Fig. 5D), which is again consistent with the crystal structure. Third, distinct sets of cross-links for C3 versus C3b amongst the MG7, MG8, C345C and anchor domains suggest the rearrangements of domains in this region. Such a domain rearrangement is entirely consistent with a comparison of the crystal structures..

Interestingly, two pairs of residues, with one partner in TED and the other in MG1, were cross-linked in C3b, in solution, yet were far apart in the crystal structure (38.3 and 40.4 Å compared to a theoretical cross-linking limit of 27.4 Å for the cross-linker used here (Fig. 6A). A change in the juxtaposition of the TED and MG1 domain could explain this but, on the other hand, three further TED-MG1 cross-links support the arrangement seen in the C3b crystal structure. This apparent conflict can be resolved assuming the TED domain to be mobile with

respect to MG1 in solution (Fig. 6B). This is consistent with several other indications that TED is mobile in C3b (26-28).

QCLMS may reveal subtle protein conformational changes

Based on a comparison of QCLMS data and crystal structures some other observations may be more generally relevant to the challenging task of translating differences in yields of cross-linked products into inferred changes in protein conformation. Due to the absence of some residues in the crystal structures of C3 and C3b, not all cross-links can be evaluated against crystallographic evidence. Of the 104 cross-links observed, 70 bridged residues that are present in both the crystal structures of C3 and C3b. In the case of cross-links for which similar yields were obtained in both C3 and C3b (n=36), the corresponding inter-residue distances are conserved between the two crystal structures. Some C3-specific or C3b-specific cross-links corresponded to pairs of residues that are within cross-linkable distances in one structure and not in the other (C3-unique n=5; C3b-unique n=8) (Fig. 7A).

So far, these observations are entirely consistent with expectations. On the other hand, 11 other conformation-specific cross-links (C3-unique n=4; C3b-unique n=7) are feasible, distance-wise, in both proteins. To account for their absence from one or other protein we must invoke steric effects that would prevent formation of a bridge (Fig. 7B), or changes in accessibility of cross-linkable residues (Fig. 7C). Importantly, ten signals that are relatively enriched (rather than being unique) in either C3 compared to C3b, or *vice versa*, do not correspond to differences in residue proximities between the two crystal structures (Fig. 7A). These ten cross-links are located near to the C3b-unique and C3-unique cross-links; they lie within regions that are rearranged during the C3-to-C3b conversion (Fig. 5E). This suggests that differences in the yields of cross-link formation may arise from subtle local rearrangements

affecting the chemical environment and through this the relative cross-linking yields (Fig. 7D). In summary, cross-linking appears to be very sensitive to conformational changes and the careful quantitative analysis of cross-link strengths could be a useful tool in structural biology.

Discussion

We have established a workflow for QCLMS that we believe is in accordance with good practice in quantitative proteomics. This includes replication and label-swapping. In addition, we place the cross-linked pairs of amino acid residues at the focus of the analysis by gathering and summarizing all the relevant peptide quantitation data. We have lowered the barrier to entry for researchers wishing to apply this technique by establishing a semi-automated approach; this should also facilitate application of QCLMS to more complex systems involving, for example, multiple proteins. To achieve this, we enabled the standard proteomic quantitation software Pinpoint to work with data for cross-linked peptides by "linearizing" their sequences (Fig. 2A). Importantly, this mode of quantitation is independent of the specific algorithm used for identifying cross-linked peptides. As has recently become standard for all publications in *Molecular and Cellular Proteomics*, we propose that good practice in QCLMS would include open access to both the raw data and the lists of cross-links with associated quantitation. Our data are available via ProteomeXchange (22) with identifier PXD001675 and in the supplementary material.

Our study demonstrated that QCLMS is able to explore, in solution, the differences and similarities between the arrangements of domains in C3 and C3b. The excellent agreement of QCLMS-derived data with the structural transitions suggested by the crystal structures of C3 and C3b provide strong support for our approach. It also suggests some rules that determine how changes in protein conformation influence the yields of cross-linked peptides. Clearly, residue proximity can in many cases act as a simple binary switch for cross-link formation. But even in instances where a bridgeable distance between two cross-linkable residues is largely preserved during a conformational change (here as close as 0.1 Å), other factors than distance may impact cross-link formation; these include changes in surface accessibilities or the

positions of the two partners relative to other structural features. While these factors could completely prohibit formation of a cross-link between two residues that are within range, complete negation seems to be rare. More commonly, these non-distance factors cause variation in yields of cross-linked products. This is manifested in a linkage that is enriched in one conformer relative to the other. In general, it seems that complete loss or gain of a cross-link reflects large conformational changes while depletion or enrichment of the cross-link correlates with smaller or more local conformational changes. It is therefore essential to distinguish experimentally between the two scenarios. This is possible through the robust quantitation, using replicated analysis with isotopic label-swapping, outlined herein

In conclusion, QCLMS is emerging as a tool for studying dynamic protein architectures. We have presented a carefully designed and thoroughly evaluated workflow for QCLMS analysis using isotope-labeled cross-linkers. A limitation of the approach is that each conformer to be analyzed must be stable on a time-scale during which it can be resolved from other conformers and then cross-linked. On the other hand, the automation in quantitation achieved in our protocols means that it should be feasible to extend a QCLMS study to the conformational changes that occur in multiple-protein assemblies (29). Thus it should be possible, for example, to follow the conformational preferences of a protein subunit in a series of assemblies as proteins are sequentially incorporated. The quantitation module in our workflow can also be adapted for SILAC-based quantitation or label-free quantitation (30) for cross-linked peptides. We envision that QCLMS will greatly facilitate the investigation of conformational dynamics in solution and help to animate the current largely crystallography-derived “snapshots” of biological processes.

Acknowledgments

We thank Christoph Schmidt, Carla Clark and Paul McLaughlin for helpful discussions and Morten Rasmussen for initial software developments. We acknowledge the PRIDE team for the deposition of our data to the ProteomeXchange Consortium. This work was supported by a European Community Marie Curie Excellence Grant (ZAC, JR: MS-MODIB), the Wellcome Trust (PNB: 081179, JR: 084229, 103139, 092076, 091020) and BBSRC (BB/I007946/1). JR is a Senior Research Fellow of the Wellcome Trust.

References

1. Leitner, A., Walzthoeni, T., Kahraman, A., Herzog, F., Rinner, O., Beck, M., and Aebersold, R. (2010) Probing native protein structures by chemical cross-linking, mass spectrometry, and bioinformatics. *Mol Cell Proteomics* 9, 1634-1649
2. Rappsilber, J. (2011) The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *J Struct Biol* 173, 530-540
3. Sinz, A. (2006) Chemical cross-linking and mass spectrometry to map three-dimensional protein structures and protein-protein interactions. *Mass Spectrom Rev* 25, 663-682
4. Fischer, L., Chen, Z. A., and Rappsilber, J. (2013) Quantitative cross-linking/mass spectrometry using isotope-labelled cross-linkers. *Journal of proteomics* 88, 120-128
5. Kalkhof, S., Ihling, C., Mechtler, K., and Sinz, A. (2005) Chemical cross-linking and high-performance Fourier transform ion cyclotron resonance mass spectrometry for protein interaction analysis: application to a calmodulin/target peptide complex. *Anal Chem* 77, 495-503
6. Mueller, D. R., Schindler, P., Towbin, H., Wirth, U., Voshol, H., Hoving, S., and Steinmetz, M. O. (2001) Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. *Anal Chem* 73, 1927-1934
7. Pearson, K. M., Pannell, L. K., and Fales, H. M. (2002) Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. *Rapid Commun Mass Spectrom* 16, 149-159
8. Rinner, O., Seebacher, J., Walzthoeni, T., Mueller, L. N., Beck, M., Schmidt, A., Mueller, M., and Aebersold, R. (2008) Identification of cross-linked peptides from large sequence databases. *Nat Methods* 5, 315-318

9. Schmidt, C., Zhou, M., Marriott, H., Morgner, N., Politis, A., and Robinson, C. V. (2013) Comparative cross-linking and mass spectrometry of an intact F-type ATPase suggest a role for phosphorylation. *Nature communications* 4, 1985
10. Kukacka, Z., Rosulek, M., Strohalm, M., Kavan, D., and Novak, P. (2015) Mapping protein structural changes by quantitative cross-linking. *Methods*
11. Janssen, B. J., Huizinga, E. G., Raaijmakers, H. C., Roos, A., Daha, M. R., Nilsson-Ekdahl, K., Nilsson, B., and Gros, P. (2005) Structures of complement component C3 provide insights into the function and evolution of immunity. *Nature* 437, 505-511
12. Janssen, B. J., Christodoulidou, A., McCarthy, A., Lambris, J. D., and Gros, P. (2006) Structure of C3b reveals conformational changes that underlie complement activity. *Nature* 444, 213-216
13. Sanchez-Corral, P., Anton, L. C., Alcolea, J. M., Marques, G., Sanchez, A., and Vivanco, F. (1989) Separation of active and inactive forms of the third component of human complement, C3, by fast protein liquid chromatography (FPLC). *J Immunol Methods* 122, 105-113
14. Maiolica, A., Cittaro, D., Borsotti, D., Sennels, L., Ciferri, C., Tarricone, C., Musacchio, A., and Rappsilber, J. (2007) Structural analysis of multiprotein complexes by cross-linking, mass spectrometry, and database searching. *Mol Cell Proteomics* 6, 2200-2211
15. Ishihama, Y., Rappsilber, J., and Mann, M. (2006) Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *J Proteome Res* 5, 988-994
16. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc* 2, 1896-1906

17. Rappsilber, J., Ishihama, Y., and Mann, M. (2003) Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal Chem* 75, 663-670
18. Ishihama, Y., Rappsilber, J., Andersen, J. S., and Mann, M. (2002) Microcolumns with self-assembled particle frits for proteomics. *J Chromatogr A* 979, 233-239
19. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367-1372
20. Thakur, S. S., Geiger, T., Chatterjee, B., Bandilla, P., Frohlich, F., Cox, J., and Mann, M. (2011) Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol Cell Proteomics* 10, M110 003699
21. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA
22. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolome, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* 32, 223-226
23. Kalkhof, S., and Sinz, A. (2008) Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. *Anal Bioanal Chem* 392, 305-312
24. Chen, Z. A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J. C., Nilges, M., Cramer, P., and Rappsilber, J. (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J* 29, 717-726

25. Schmidt, C., and Robinson, C. V. (2014) A comparative cross-linking strategy to probe conformational changes in protein complexes. *Nat Protoc* 9, 2224-2236
26. Alcorlo, M., Tortajada, A., Rodriguez de Cordoba, S., and Llorca, O. (2013) Structural basis for the stabilization of the complement alternative pathway C3 convertase by properdin. *Proc Natl Acad Sci U S A* 110, 13504-13509
27. Nishida, N., Walz, T., and Springer, T. A. (2006) Structural transitions of complement component C3 and its activation products. *Proc Natl Acad Sci U S A* 103, 19737-19742
28. Pechtl, I. C., Neely, R. K., Dryden, D. T., Jones, A. C., and Barlow, P. N. (2011) Use of time-resolved FRET to validate crystal structure of complement regulatory complex between C3b and factor H (N terminus). *Protein Sci* 20, 2102-2112
29. Tomko, R. J., Jr., Taylor, D. W., Chen, Z. A., Wang, H. W., Rappsilber, J., and Hochstrasser, M. (2015) A Single alpha Helix Drives Extensive Remodeling of the Proteasome Lid and Completion of Regulatory Particle Assembly. *Cell* 163, 432-444
30. Herbert, A. P., Makou, E., Chen, Z. A., Kerr, H., Richards, A., Rappsilber, J., and Barlow, P. N. (2015) Complement Evasion Mediated by Enhancement of Captured Factor H: Implications for Protection of Self-Surfaces from Complement. *J Immunol*

Figure Legends

Figure 1: Quantitative CLMS analysis of C3, C3b and iC3 in solution.

(**A**) The strategy of QCLMS using isotopologues of a cross-linker to compare two protein conformations. (**B**) The crystal structures of C3 (PDB|2A73) and C3b (PDB|2I07) with the α-chain (in C3)/α'-chain (in C3b) (blue) and the β-chains (grey) highlighted. (**C**) Chemical structure of cross-linkers BS^3 and BS^3-d4. (**D**) SDS-PAGE shows that BS^3 (light cross-linker) and BS^3-d4 (heavy cross-linker) cross-link C3 and C3b with roughly equivalent overall efficiencies, and that broadly similar sets of cross-linked protein products were obtained. (**E**) High-resolution fragmentation spectrum of BS^3-cross-linked peptides ISLPESLK(cl)R-K(cl)VLLDGVQNLR that reveals a cross-link between Lys 267 and Lys 283. The mass spectrum of the precursor ion is shown (blue) in the inset; also present (red) is the signal of the precursor ion corresponding to the equivalent BS^3-d4 cross-linked peptides.

Figure 2: Construction of Pinpoint peptide library for cross-linked peptides.

(**A**) The principle of converting a cross-linked peptide pair into a "linearized" format that has an identical mass. (**B**) An example Pinpoint-input peptide library for cross-linked peptides with its key contents annotated.

Figure 3: Quantitation of cross-linked peptides.

(**A**) MS signals from paired quantitation analyses with label-swapping. These allow a clear distinction to be made between cross-linked peptides with a singlet signal, cross-linked peptides with a doublet signal, and a non-cross-linked peptide; these are all potentially similar and could therefore be confused in a non-paired quantitative analysis (**B**) Spread of quantitation data at the level of cross-linked peptides. For each of 104 quantified cross-links, the ratio of C3b to C3-

derived signal is shown in red, while the ratio of C3b to C3-derived cross-link signal for its supporting, cross-linked peptides is plotted in blue (**C**) Ratios of C3b to C3-derived signals for 48 cross-links that were observed in both C3 and C3b. The error bar indicated the maximum variation range of each cross-link in replicated analyses (n= max 4, if a cross-link was quantified in both replica and on both mass spectrometers used here).

Figure 4: Strong cation exchange based enrichment for cross-linked peptides is compatible with QCLMS analysis.

(**A**) Strong cation exchange (SCX)-based enrichment using SCX-Stage Tips applied in QCLMS analysis of C3 and C3b. (**B**) Reproducibility of C3b/C3 signal ratios of50 cross-linked peptides in two SCX fractions in which they were detected.

Figure 5 Quantitative CLMS analysis reveals the domain rearrangements, in solution, which accompany the activation of C3.

(**A**) 104 unique cross-links that were quantified in C3 and C3b samples were allocated to three groups: 31 are unique to C3 (20 involve the ANA domain); 24 are unique to C3b; 49 were observed in both C3 and C3b (of these, four are statistically enriched in C3b, eight are enriched in C3 and 37 (denoted “mutual”) are equally represented in both C3 and C3b. (**B**) The 37 mutual cross-links (black) are displayed on the crystal structures of C3 (PDB|2A73) and C3b (PDB|2I07). Of these, 19 are intra-domain, consistent with the conserved structures of these domains in C3 *versus* C3b. The remaining (18) links pairs of domains whose relative positions are similar in C3 and C3b. (**C**) As displayed in the C3 structure, C3-unique cross-links (blue) lie mainly within the C3-specific ANA domain (light blue) and between domains in the α-chain, which are significantly rearranged upon formation of C3b. (**D**) The C3b-unique cross-links (red)

occur primarily between domains within the rearranged α-chain and at the new interface formed between the TED and MG1 domains (displayed on PDB|2I07).

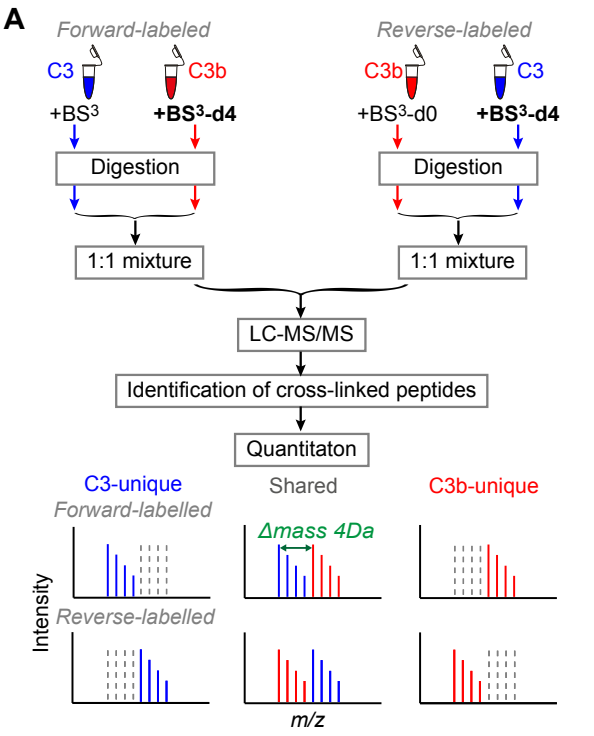
Figure 6: Flexibility of the TED domain in C3b revealed by cross-linking data in solution.

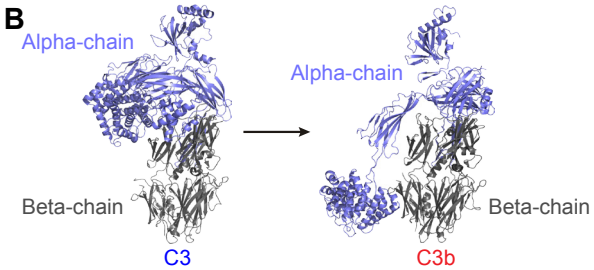
(**A**) Two of five observed C3b-unique cross-links between the MG1 domain and the TED involve pairs of residues separated, in the crystal structure, by more than the theoretical maximum of 27.4 Å (Cα-Cα). (**B**) Putative flexibility of the TED domain in solution (indicated by the arrow) would explain the observation of these two over-length cross-links.

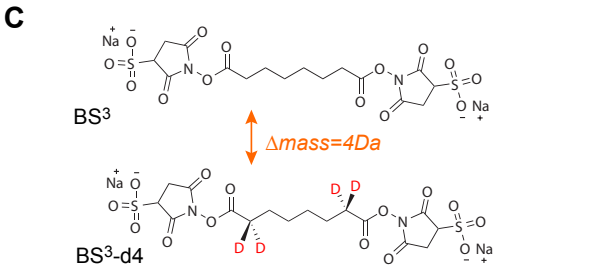
Figure 7: Impacts of conformational changes on yields of cross-links.

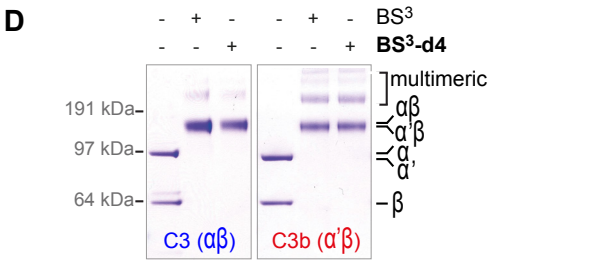
(**A**) Measured distances (Cα-Cα) between cross-linked residues (in the C3 structure) minus maximal cross-linkable distance are plotted against the equivalents distances in the C3b structure (cross-links are only plotted where both cross-linked residues are present in the crystal structures). (**B**) Distance-wise, cross-link K1479-K1573 is within cross-link limit (27.4 Å) both in C3 (12.3 Å) and C3b (23.3 Å) structures. However these cross-links were only observed in C3 because, in C3b, the steric access of cross-linkers to both residues is blocked. (**C**) A dramatic decrease in the surface accessibility of K578 coincides with the absence of a C3 unique cross-link, K578-K588 (blue line), in C3b. Even though the proximities between K578 and K588 are nearly identical in the crystal structures of C3 (12.9 Å) and C3b (12.7 Å). (**D**) Local rearrangement resulted in significant enrichment of cross-links $K857^{MG7}$-$S1501^{C345C}$ and $K857^{MG7}$-$K1504^{C345C}$ in C3 and significant enrichment of cross-links $K857^{MG7}$-$K1513^{C345C}$ in C3b. Residue K1501 is not present in the C3b crystal structure (PDB|2I07), residue 1500 is used instead for display purposes.

Figure 1


A
Forward-labeled
Reverse-labeled
C3
C3b
+BS3
+BS3-d4
+BS3-d0
Digestion
1:1 mixture
LC-MS/MS
Identification of cross-linked peptides
Quantitaton
C3-unique
Shared
C3b-unique
Forward-labelled
Δmass 4Da
Intensity
Reverse-labelled
m/z


B
Alpha-chain
Beta-chain
C3
C3b


C
BS3
Δmass=4Da
BS3-d4


D
BS3
BS3-d4
multimeric
αβ
α'β
α'
β
191 kDa
97 kDa
64 kDa
C3 (αβ)
C3b (α'β)

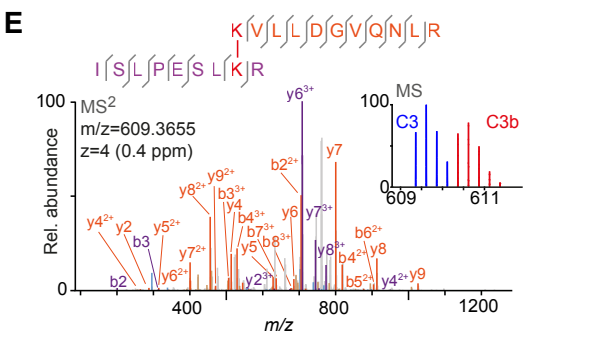

E
K V L L D G V Q N L R
I S L P E S L K R
MS2
m/z=609.3655
z=4 (0.4 ppm)
Rel. abundance
m/z
MS
C3
C3b

Figure 2

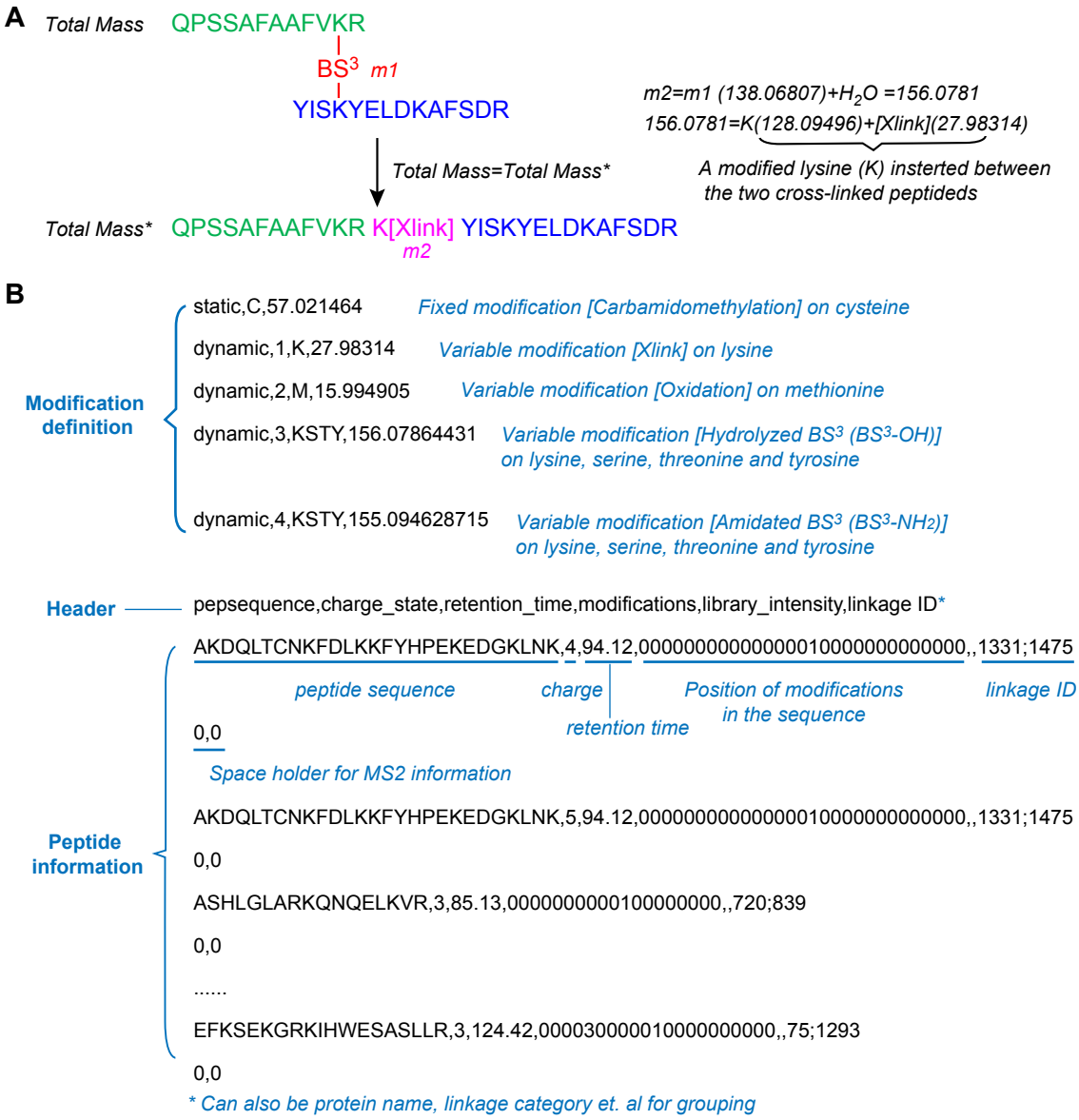

A
Total Mass
QPSSAFAAFVKR
BS3 m1
YISKYELDKAFSDR
Total Mass=Total Mass*
Total Mass*
QPSSAFAAFVKR K[Xlink] YISKYELDKAFSDR
m2
m2=m1 (138.06807)+H2O =156.0781
156.0781=K(128.09496)+[Xlink](27.98314)
A modified lysine (K) insterted between the two cross-linked peptideds
B
Modification definition
static,C,57.021464
Fixed modification [Carbamidomethylation] on cysteine
dynamic,1,K,27.98314
Variable modification [Xlink] on lysine
dynamic,2,M,15.994905
Variable modification [Oxidation] on methionine
dynamic,3,KSTY,156.07864431
Variable modification [Hydrolyzed BS3 (BS3-OH)] on lysine, serine, threonine and tyrosine
dynamic,4,KSTY,155.094628715
Variable modification [Amidated BS3 (BS3-NH2)] on lysine, serine, threonine and tyrosine
Header
pepsequence,charge_state,retention_time,modifications,library_intensity,linkage ID*
Peptide information
AKDQLTCNKFDLKKFYHPEKEDGKLNK,4,94.12,00000000000000010000000000000,,1331;1475
peptide sequence
charge
retention time
Position of modifications in the sequence
linkage ID
0,0
Space holder for MS2 information
AKDQLTCNKFDLKKFYHPEKEDGKLNK,5,94.12,00000000000000010000000000000,,1331;1475
0,0
ASHLGLARKQNQELKVR,3,85.13,0000000000100000000,,720;839
0,0
.....
EFKSEKGRKIHWESASLLR,3,124.42,000030000010000000000,,75;1293
0,0
* Can also be protein name, linkage category et. al for grouping

Figure 3

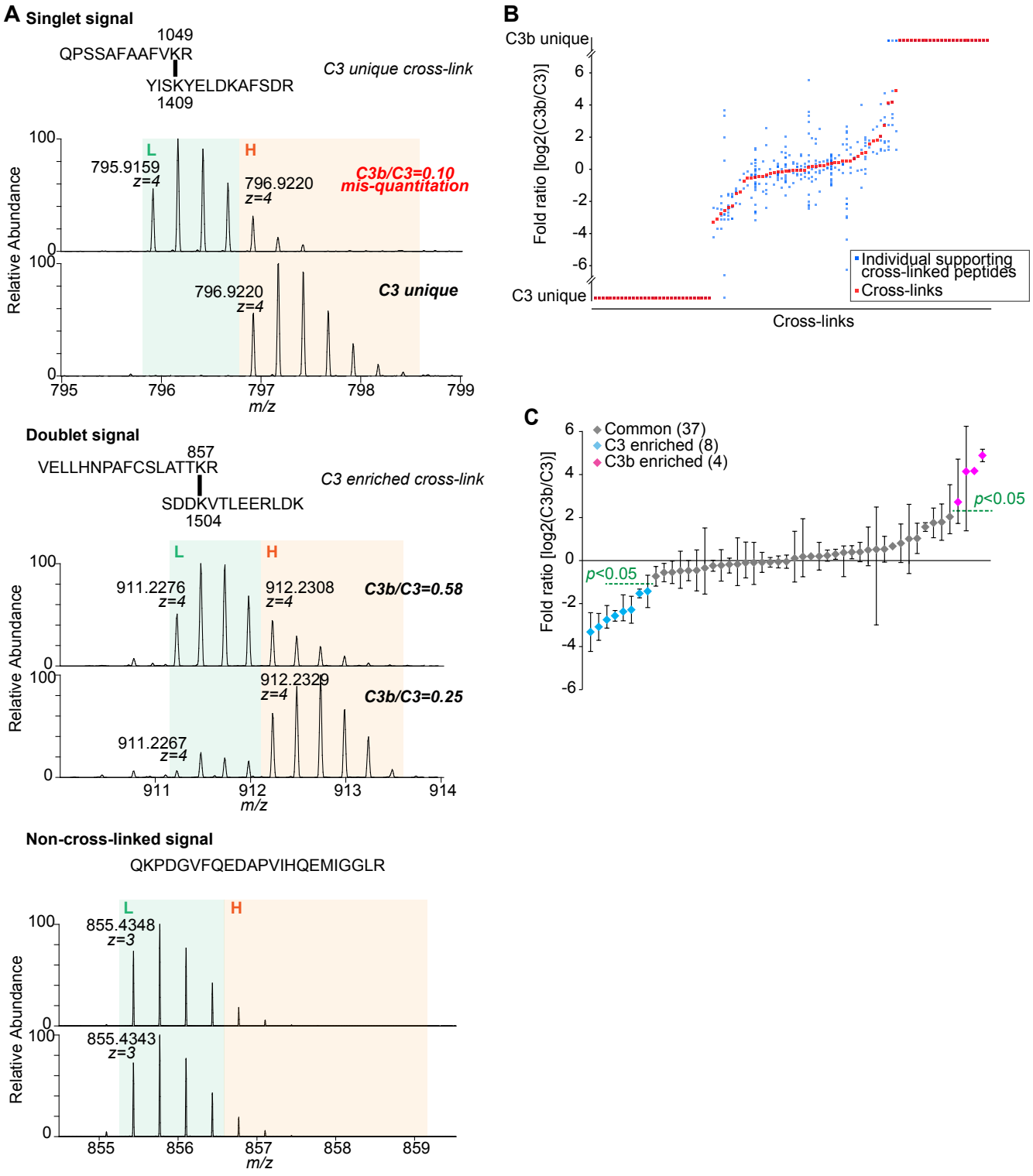

A Singlet signal
1049
QPSSAFAAFVKR
C3 unique cross-link
YISKYELDKAFSDR
1409
L
H
795.9159
z=4
796.9220
z=4
C3b/C3=0.10
mis-quantitation
796.9220
z=4
C3 unique
Relative Abundance
m/z
Doublet signal
857
VELLHNPAFCSLATTKR
C3 enriched cross-link
SDDKVTLEERLDK
1504
911.2276
z=4
912.2308
z=4
C3b/C3=0.58
912.2329
z=4
C3b/C3=0.25
911.2267
z=4
Non-cross-linked signal
QKPDGVFQEDAPVIHQEMIGGLR
855.4348
z=3
855.4343
z=3
B
C3b unique
C3 unique
Fold ratio [log2(C3b/C3)]
Individual supporting cross-linked peptides
Cross-links
Cross-links
C
Common (37)
C3 enriched (8)
C3b enriched (4)
p<0.05
p<0.05
Fold ratio [log2(C3b/C3)]

Figure 4

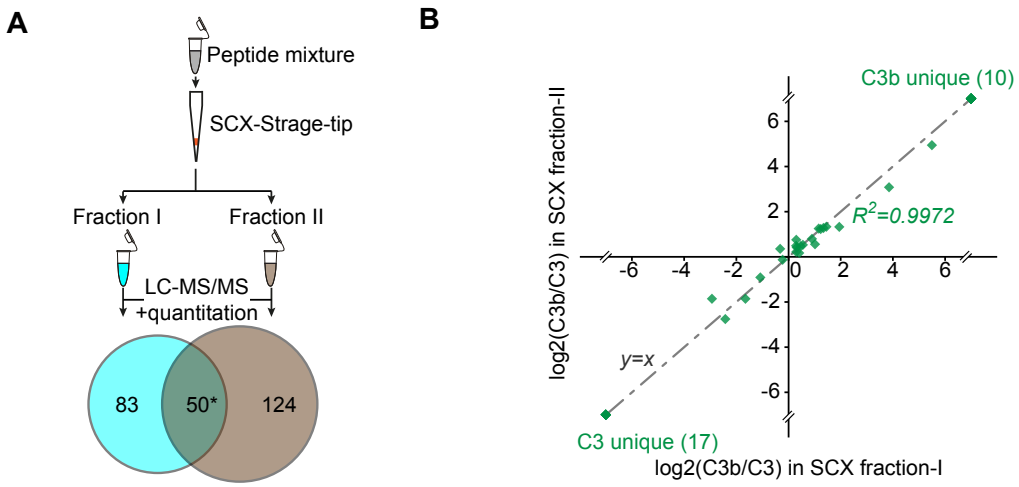

A
Peptide mixture
SCX-Strage-tip
Fraction I
Fraction II
LC-MS/MS
+quantitation
83
50*
124
B
C3b unique (10)
R^2=0.9972
y=x
C3 unique (17)
log2(C3b/C3) in SCX fraction-II
log2(C3b/C3) in SCX fraction-I
6
4
2
-2
-4
-6
-6
-4
-2
0
2
4
6

Figure 5

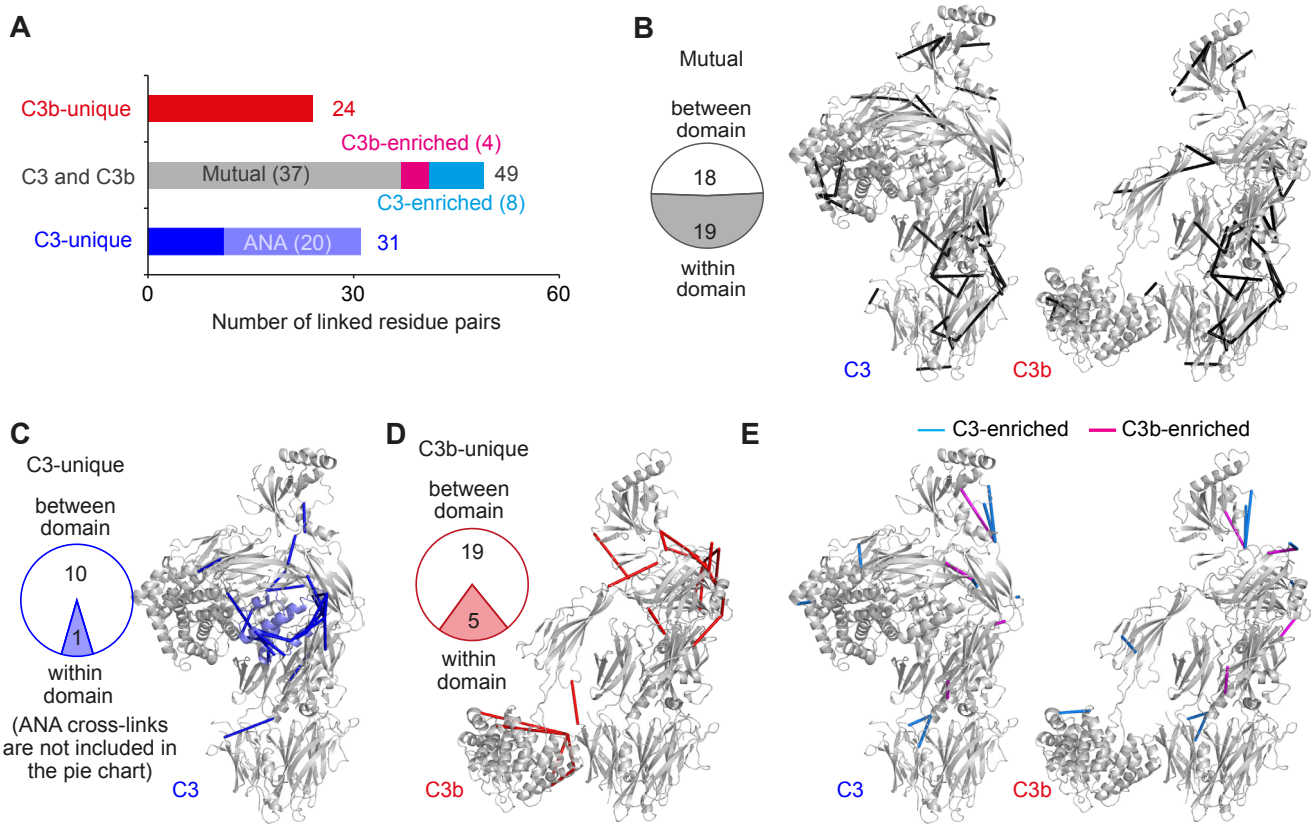

A
C3b-unique
24
C3b-enriched (4)
C3 and C3b
Mutual (37)
49
C3-enriched (8)
C3-unique
ANA (20)
31
0
30
60
Number of linked residue pairs
B
Mutual
between domain
18
19
within domain
C3
C3b
C
C3-unique
between domain
10
1
within domain
(ANA cross-links are not included in the pie chart)
C3
D
C3b-unique
between domain
19
5
within domain
C3b
E
C3-enriched
C3b-enriched
C3
C3b

Figure 6

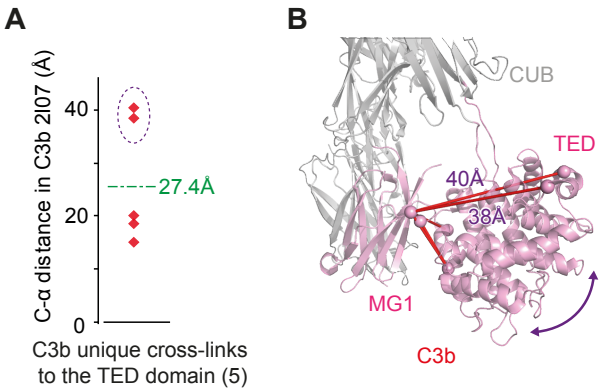
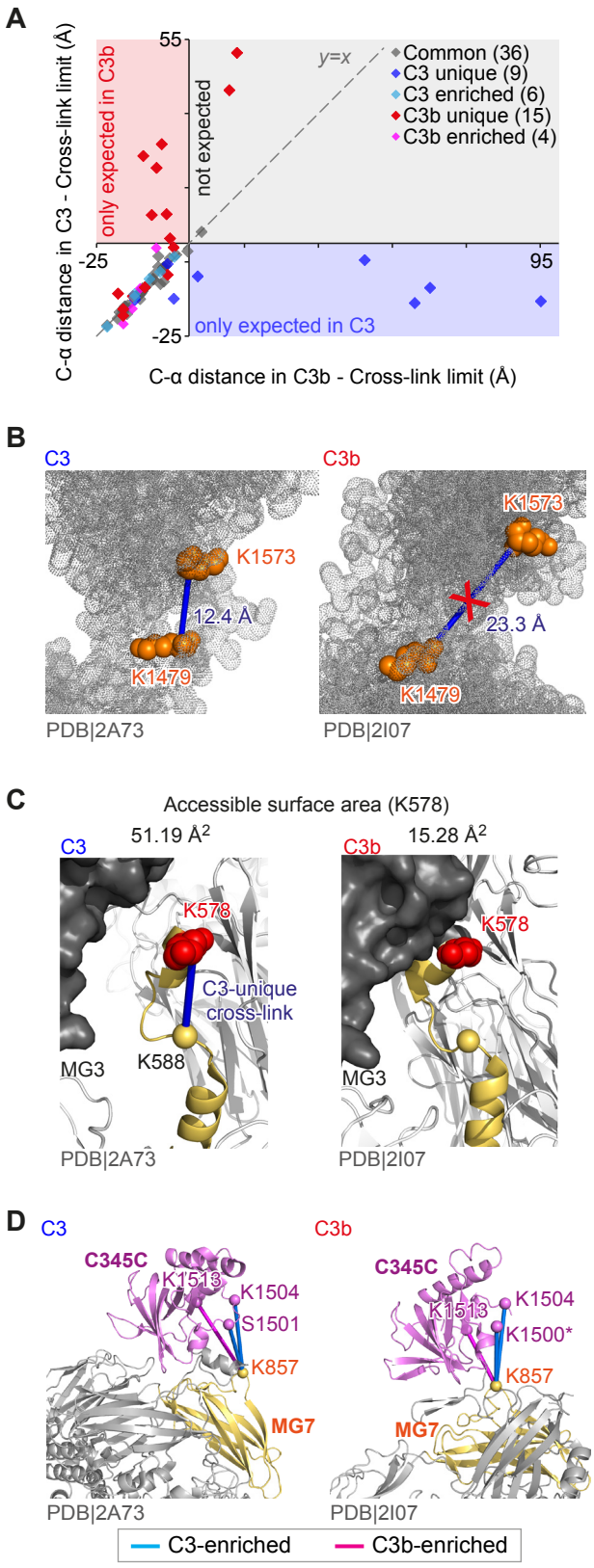

A
B
C-α distance in C3b 2I07 (Å)
40
20
0
27.4Å
C3b unique cross-links
to the TED domain (5)
CUB
TED
40Å
38Å
MG1
C3b

Figure 7


A
C-α distance in C3 - Cross-link limit (Å)
C-α distance in C3b - Cross-link limit (Å)
55
-25
95
-25
only expected in C3b
not expected
only expected in C3
y=x
Common (36)
C3 unique (9)
C3 enriched (6)
C3b unique (15)
C3b enriched (4)
B
C3
C3b
K1573
12.4 Å
K1479
K1573
23.3 Å
K1479
PDB|2A73
PDB|2I07
C
Accessible surface area (K578)
C3
51.19 Å2
C3b
15.28 Å2
K578
C3-unique cross-link
K588
MG3
PDB|2A73
K578
MG3
PDB|2I07
D
C3
C3b
C345C
K1513
K1504
S1501
K857
MG7
C345C
K1513
K1504
K1500*
K857
MG7
PDB|2A73
PDB|2I07
C3-enriched
C3b-enriched