1354–1369 *Nucleic Acids Research, 2016, Vol. 44, No. 3* *Published online 10 December 2015*
doi: 10.1093/nar/gkv1366

Targeted cross-linking-mass spectrometry determines vicinal interactomes within heterogeneous RNP complexes

Christian Trahan[1,2] and Marlene Oeffinger[1,2,3,*]

[1]Department for Systems Biology, Institut de recherches cliniques de Montréal, Montréal, Québec H2W 1R7, Canada, [2]Département de biochimie, Faculté de médecine, Université de Montréal, Montréal, Québec H3T 1J4, Canada and [3]Division of Experimental Medicine, Faculty of Medicine, McGill University, Montréal, Québec H3A 1A3, Canada

Received September 29, 2015; Revised November 18, 2015; Accepted November 24, 2015

ABSTRACT

Proteomic and RNomic approaches have identified many components of different ribonucleoprotein particles (RNPs), yet still little is known about the organization and protein proximities within these heterogeneous and highly dynamic complexes. Here we describe a targeted cross-linking approach, which combines cross-linking from a known anchor site with affinity purification and mass spectrometry (MS) to identify the changing vicinity interactomes along RNP maturation pathways. Our method confines the reaction radius of a heterobifunctional cross-linker to a specific interaction surface, increasing the probability to capture low abundance conformations and transient vicinal interactors too infrequent for identification by traditional cross-linking-MS approaches, and determine protein proximities within RNPs. Applying the method to two conserved RNA-associated complexes in *Saccharomyces cerevisae*, the mRNA export receptor Mex67:Mtr2 and the pre-ribosomal Nop7 subcomplex, we identified dynamic vicinal interactomes within those complexes and along their changing pathway milieu. Our results therefore show that this method provides a new tool to study the changing spatial organization of heterogeneous dynamic RNP complexes.

INTRODUCTION

Ribonucleoprotein particles (RNPs) are matured in the nucleus through a series of events that begin with transcription and end with an RNP being exported from the nucleus through a nuclear pore complex (NPC) (1,2). RNP maturation is driven by transient interactions with RNA biogenesis factors, which are assembled onto the RNA and facilitate individual steps in processes that are specific to the type of RNA (1,3). RNP assembly, which is believed to be performed in a hierarchical fashion to ensure correct RNA maturation but may also involve parallel biogenesis routes, is a highly dynamic process during which the association and dissociation of biogenesis factors create highly heterogeneous assembly intermediates (4–7). Given the integrated nature and dynamics of RNP assembly involving hundreds of proteins and thousands of different RNAs, researchers have put considerable effort into studying the composition of different RNP complexes. The advent of new and high-throughput approaches enabled us to determine RNP composition as well as RPB rearrangement occurring on RNPs during their maturation on a global scale. Techniques such as affinity purification-mass spectrometry (MS) as well as protein–RNA and RNA–RNA UV crosslinking approaches combined with RNA sequencing have provided deeper insight into RNP assembly, quality control and regulation, including many of the proteins involved, position-specific binding of known RNA-binding proteins (RBPs) to different transcripts and RNAs and glimpses into order and timing of assembly (5–6,8).

Yet despite the increasingly detailed picture we have gained about RNPs and their components, little is known about the vicinal protein-neighbourhoods, proximity or direct interactions of proteins within different RNPs, how these change along assembly intermediates, or the overall architecture of different intermediates (9,10). As RNPs are large, heterogeneous and dynamic complexes, proteomic analysis alone will not provide information on the composition of distinct assembly intermediates or the vicinal interactome of selected baits. Moreover, not all RNP associated proteins are necessary RBPs, and thus protein–RNA crosslinking approaches will fail to report on the positioning of these factors. As proteins are the central players and regulators, not only of RNP pathways but all cellular information pathways, determining the changing vicinal interactomes of factors as well as protein proximities within complexes throughout RNP assembly is imperative for discerning their individual roles in the regulation and advance-

*To whom correspondence should be addressed. Tel: +1 514 987 5668; Email: marlene.oeffinger@ircm.qc.ca

ment of a pathway. Moreover, it will allow us to gain information on the architecture of different RNP intermediates and, over time, build a more complete dynamic picture of RNP maturation pathways.

One way to stabilize the dynamic nature of a protein's vicinal interactors is to cross-link them. It has been shown that chemical cross-linking can yield low-resolution structure information on the distance constraints within a molecule, and advances in chemical cross-linking reagents combined with MS have created a powerful method for determining the composition and architecture of macromolecular complexes (11). However, conventional cross linking-MS methods, employing homobifunctional cross-linkers targeting lysines due to their high frequency within proteins, have been used to position and refine crystallographic structures of subcomplexes to and gain deeper insights into the three-dimensional (3D) architecture of known complex structures rather than dynamic heterogeneous macromolecular complexes of unknown structure (12–14). Applying this method to macromolecular complexes such as RNPs is extremely challenging due to their heterogeneity, dynamicity and complexity, making the discrimination between different intermediates and changing vicinal interactomes difficult. This is also hindered by poor signal-to-noise ratio of individual crosslinks, and the non-discriminate nature of homobifunctional cross-linkers, which makes the analysis of protein interactions within such complexes extremely challenging. To date, there is no label transfer reagent available that attaches to one unique site in a protein.

Here, we have developed a targeted cross-linking approach to identify near neighbours and vicinity interactomes as well as determine protein proximities within stable and dynamic RNPs. In this method, one end of the cross-linker is fixed to a single point within an *ex vivo* isolated RNP as another protein within the same complex contains a Protein A-tag, which renders the complex affinity-purifiable by IgG. The vicinal interactome is probed with a heterobifunctional cross-linker, which, attached to the single anchor site, has a restricted reaction radius thus limiting the number of possible crosslinks to a very defined region (Figure 1A). The approach takes advantage of the comparative rarity of surface cysteines on proteins as we are using a combination-affinity tag, which contains one cysteine immediately adjacent to a His_{10} affinity tag (CH-tag) (Figure 1B). The use of the well-studied, commercial heterobifunctional cross-linker SM(PEG)n, targeting cysteine and lysine, allows us to saturate the attachment to the cysteines prior to activating cross-linking to any nearby lysines by simply shifting the pH of our buffer (Figure 1C) (15,16).

We applied our approach to two independent protein assemblies: (i) the yeast mRNA export receptor heterodimer Mex67:Mtr2 (17) and (ii) the Nop7 subcomplex involved in yeast ribosome biogenesis (18). A large portion of the *Saccharomyces cerevisae* Mex67:Mtr2 heterodimer had recently been crystallized (19,20), which made it a suitable proof-of-principle model as it allowed us to predict potential cross-links based on the location of the CH-tag and cross-linker length. Moreover, the heterodimer is part of the dynamic mRNA maturation pathway, with changing interaction partners from the nucleus, through the NPC, to the cytoplasm, and cross-linking experiments may provide some insight in its changing interaction partners along the pathway (2). For the heterotrimeric Nop7 complex, potential interactions between Nop7, Erb1 and Ytm1 have been inferred from genetic and biochemical studies; it thus presents a suitable less understood model complex (18,21). Using our approach, we have further elucidated the spatial organization of the Mex67:Mtr2 heterodimer and close-proximity components of its dynamic vicinal interactome within the nuclear pore, as well as previously unknown dynamic interactions of the pre-ribosomal Nop7 subcomplex, demonstrating the strength of this method.

MATERIALS AND METHODS

More detailed protocols, plasmids, strains and primers are provided in the Supplementary Data file.

Cloning of pCH-KanMX6 and generation of yeast strains

Two complementary 51-mers containing the sequence for the CH-tag ('RCDPHHHHHHHHHH') with overhangs for the restriction sites SalI and AscI were phosphorylated and then slowly hybridized at RT. Plasmid pYM13 (Euroscarf) was cut with SalI and AscI to release the Tandem Affinity Purification (TAP) tag and dephosphorylated, and the CH-tag fragment was inserted. Positive clones were screened for on Ampicillin-containing plates and candidates were sequenced for confirmation. Growth and handling of *S. cerevisiae* were carried out using standard techniques. PrA- and CH-tagged strains were generated by a one-step polymerase chain reaction strategy in the wild-type strain W303 as described (22). Plasmids, strains and primers are listed in Supplementary Tables S1–3.

Cell harvest and lysis

Harvested cells were rapidly frozen in liquid nitrogen and cryolysis was performed by solid phase milling in a planetary ball mill (Retsch) producing a fine cell grindate (23). The grindate was stored at −80°C until processed.

Affinity purification

Affinity purification of Protein A-tagged bait proteins was carried out as described in (23). Frozen cell grindate was rapidly thawed into TBT150 buffer with 1 mM Dithiothreitol (DTT). The resulting lysate was vortexed for 1 min, polytroned for 30 s and then cleared by centrifugation at 2600 *g* for 5 min at 4°C. Magnetic beads (Dynabeads M-270, Invitrogen), conjugated with rabbit IgG Ab (Sigma), were washed three times with TBT150 + DTT and added to the cleared cell lysate at a concentration of 3.75 mg (25 μl slurry)/0.5 g of cell grindate. The samples were rotated for 30 min at 4°C. After binding, the beads were magnetically harvested and then quickly washed three times with 10ml of TBT150 + DTT, and finally once with 10ml of LWB while vortexing at very low speed for 5 min.

Sequential crosslinking

For removal of ammonium ions from the last wash and equilibration, the beads were washed with Maleimide Reaction Buffer (MRB) (1× 20 ml, 2× 10 ml). The $SM(PEG)_2$

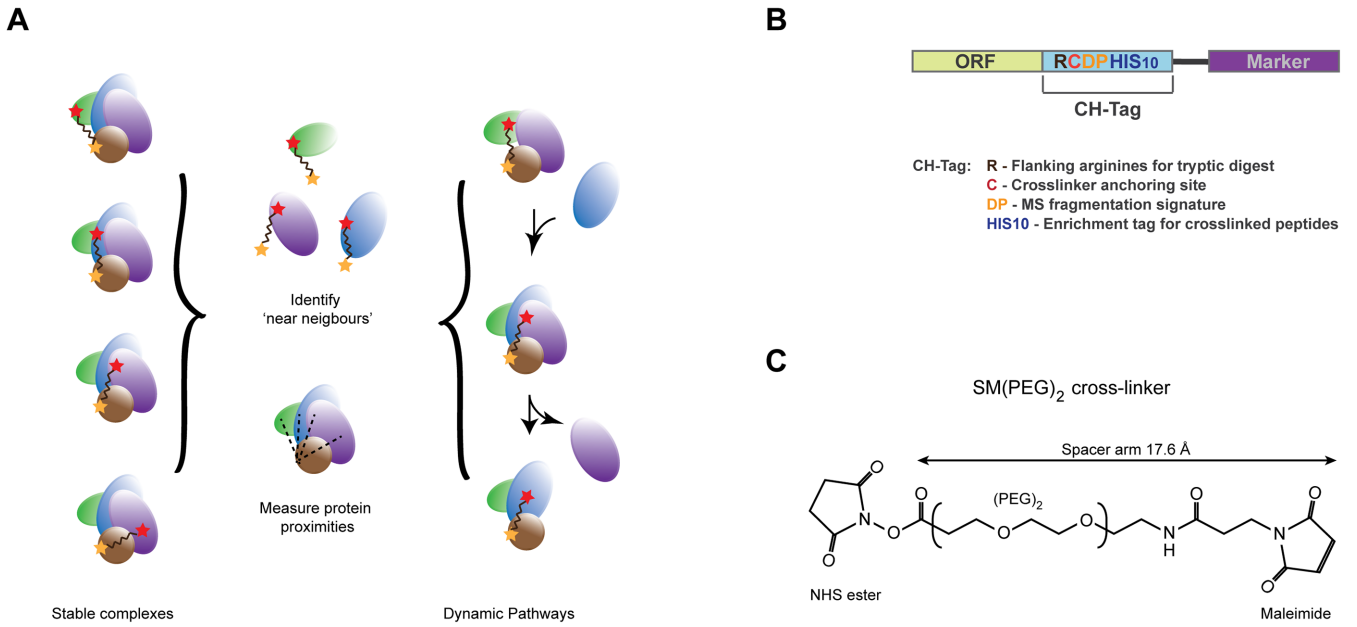

Figure 1. Schematic outline of the targeted crosslinking strategy and tools. (**A**) Determining protein proximities and identify neighbors vicinal to an epitope-tagged bait protein within either a stable complex or a dynamic pathway, using a targeted cross-linking approach. (**B**) Sequence arrangement of the CH-epitope tag for targeted cross-linking. The anchoring site for a heterobifunctional cross-linker, a single cysteine, is integrated into the epitope tag in addition to a His_{10} sequence for the enrichment of cross-linked peptides prior to mass spectrometric analysis. (**C**) $SM(PEG)_2$ cross-linker (Thermo Scientific).

cross-linker (Thermo Scientific) was diluted to 100 μM in 1 ml of MRB, and added to the beads, tapping the bottom of the tube and swirling the buffer until the beads are dispersed. The maleimide reaction was allowed to proceed for 30 min at room temperature while vortexing at the lowest speed (160 RPM). After 30 min, the cross-linker was diluted with 10 ml of MRB, the beads collected, and, after removal of the supernatant, 10 ml of N-Hydroxysuccinimide (NHS)-ester reaction Buffer (NRB) were quickly added to the beads for initial dispersion before adding an additional 35 ml of NRB. The NHS-ester reaction was allowed to proceed for 1 h at room temperature with the tube standing upright in a tube rack, inverting the tube every few minutes. The reaction was quenched by the addition of 5 ml of 1M Tris–HCl and the tube incubated on a nutator for 30 min at room temperature.

Tryptic digest and peptide enrichment

Bead suspension in Trypsin buffer was transferred to Lobind 1.5 ml tubes, and after addition of 355 μl Trypsin buffer and 25 μl trypsin (Promega; 0.2 μg/μl in 1 mM HCl) incubated overnight at 37°C on a rotating wheel. Next day 10 μl trypsin (0.3 μg/μl in Trypsin buffer) were added and the reaction incubated at 37°C, rotating for 4 h. Trypsin quantities and buffer volumes may be adjusted according to bead volume amount of material. A high trypsin to protein ratio of about 1:10 was used. Reduction/alkylation of cysteine was not necessary as the cysteine sulfhydryl groups within cross-linked CH-tags are no longer available for alkylation. Samples were magnetized and the supernatant transferred to a fresh tube and 20 μl of 5M NaCl (f.c. 250 mM) and 0.8 μl 2.5M imidazole pH 8.0 (f.c. 5 mM, high purity grade) (Semba Biosciences) added. Following tryptic digest, the peptides carrying the cross-linking anchoring tag were enriched for using 200 μl nickel-resin containing UptiTips (Interchim Innovations). The tips were rehydrated by washing 5× with 200 μl H_2O, then 3× with 200 μl of UptiTip Pre-Wash Buffer (UPWB) before being equilibrated 5× with 200 μl of UptiTip Binding Buffer (UBB). For each trypsin digests obtained above, two UptiTips are used. Half of the trypsin digest was transferred to a new 1.5 ml LoBind tube (Eppendorf), and the peptides are bound to the equilibrated UptiTips by pipetting the peptide mixture 100×. The UptiTips are then washed 30× with 20 μl of UBB, 3× with 20 μl UptiTip Wash Buffer (UWB), followed by three washes with 20 μl of MS grade H_2O. Once the remaining water is expelled, the His-tag-containing, cross-linked peptides were eluted by pipetting 20× in 10 μl UptiTip Elution Buffer (UEB). The elution step was repeated twice, all elution fractions were pooled into one 1.5 ml LoBind tube and the purified peptides lyophilized in a SpeedVac for ∼1 h. Lyophilized samples were resuspended in MS Sample Buffer (MSB) containing 1 mM Tris(2-carboxyethyl)phosphine (TCEP) to prevent formation of disulfide bonds between free cysteines, as uncross-linked peptides linked by disulfide bonds would elute later during the LC elution gradient, interfering with the MS detection of cross-linked peptides to the CH-tag. Uncross-linked CH-tag-containing peptides were not well retained at the high concentration of acetonitrile used at the beginning of the LC elution gradient (see 'Mass spectrometry' section below), preventing their co-elution with cross-linked peptides.

Mass spectrometry

The lyophilized peptides were resuspended in 10 μl MSB and loaded directly onto a PicoFrit fused silica capillary column (15 cm × 75 μm i.d; New Objective) packed with C-18 reverse-phase material (Jupiter 5 μm particles, 300 Å pore size; Phenomenex) using a high-pressure packing cell. This column was installed on the Easy-nLC II system (Proxeon Biosystems) and coupled to the Q-Exactive (ThermoFisher Scientific) equipped with a Proxeon nano-electrospray Flex ion source. The chromatographic separation of peptides was carried out in a 0.2% formic acid (buffer A) and 100% acetonitrile/0.2% formic acid (buffer B). Peptides are loaded on-column at a flow rate of 600 nl/min and eluted with a 2 slope gradient at a flow rate of 250 nl/min. Solvent B was increased from 10 to 40% B over 42 min, and then from 40 to 85% B over 18 min. LC-MS/MS data was acquired using a data-dependent top 12 method combined with a dynamic exclusion window of 5 s. The mass resolution for full MS scan was set to 70 000 (at m/z 400) and lock masses were used to improve mass accuracy. The mass range window was set to 330–2000 m/z for MS scanning with a target value at 1×10^6, the maximum ion fill time (IT) at 100 ms, the intensity threshold at 1.2×10^4 and the underfill ratio at 0.9%. The data dependent MS2 scan events were acquired at a resolution of 17 500 with the maximum ion fill time at 75 ms and the target value at 1×10^5. The normalized collision energy used was at 27 and the capillary temperature was 250°C. Nanospray and S-lens voltages were set to 1.3–1.7 kV and 50 V, respectively.

Data analysis

For data analysis, Thermo Excalibur .raw files were converted into mgf format using MASCOT software (24). The .mgf files were used as input for data searches with pLink software v1.21 (http://pfind.ict.ac.cn/software/pLink/) (25), set to use higher energy collisional dissociation (HCD) ion types. The pLink default 20 ppm error window for MS/MS fragment ion mass for HCD dataset in the instrument.ini file was used and missed cleavages were set to a maximum of 4. The $SM(PEG)_2$ cross-linker mass and sequence was added to the xlink.ini file for pLink for correct mass addition to peptides linked by the cross-linker. In addition, since the NHS ester group can result in O-acylation of serine, threonine and tyrosines (26), three more cross-linker masses corresponding to these reactions were added into xlink.ini. The default search window of +/− 5 Da on the precursor mass tolerance for combinatorial mode was used to cover all precursors monoisotopic and isotopic peaks, to better assign monoisotopic peaks to MS/MS spectra. A filter of +/− 10 ppm was used for all precursors isotopic mass accuracy, and the *E*-value was set to 0.001, which corresponds to false discover rate (FDR) of less than 0.05% according to the determined relation between *E*-values and FDR (25). The pLink analysis of MS data was carried out using simultaneously all four cross-linkers set in xlink.ini for each analysis. Identified peptides were run against a whole proteome of *S. cerevisiae* W303 fasta file in pLink as well as against custom fasta files from the BioGRID (http://thebiogrid.org/) lists of proteins known to interact with each CH-bait. All peptide spectra were manually validated including their ion fragmentation patterns, and peptides displaying >5 ppm difference between theoretic and observed masses were eliminated from the compiled results.

Complex modelling and cross-link predictions

In order to predict $SM(PEG)_2$ cross-linked sites within the Mex67/Mtr2 and Nop7-Ytm1 complexes, we used I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) (27) to model the CH-tag, either of Mtr2-CH using the crystal of Mex67/Mtr2 (4WWU.PDB) (20) as template, or of Ytm1-CH and Nop7-CH using protein sequences. The CH-tag of the best scoring PDB file generated by I-TASSER was then taken and added directly on the 4WWU.PDB structure file for Mex67/Mtr2. To simplify the visualization of the cross-links on the 4WWU structure, chains D through I were removed resulting in two heterodimers of Mex67/Mtr2 interacting with one another. The resulting PDB file was then submitted to XWalk (http://www.xwalk.org/) (28) for prediction, validation and visualization of obtained cross-links. For Nop7/Ytm1, XL-MS information for spatial restraints to create a Ytm1-Nop7 docking model were submitted to HADDOCK (29).

SDS PAGE Sample analysis

For analysis of cross-linked samples by western blot, 1/250th of material taken before and after cross-linking were separated on a 4–12% NuPAGE Novex Bis-Tris precast gel (Life Technologies) and transferred to PVDF membrane. The membrane was blocked in 5% milk/1×TBST. His-tagged proteins were detected using a mouse anti-His antibody (ABM), followed by an HRP-conjugated anti-mouse antibody (GE Healthcare). PrA-tagged proteins were detected using peroxidase-anti-peroxidase antibody (1:20 000; Sigma). Bands were visualized by chemiluminescence. For analysis by Silver staining, a 1/200th of material taken before and after cross-linking reactions was separated on a 4–12% NuPAGE Novex Bis-Tris precast gel and silver stained using the fast staining protocol of the SilverQuest staining kit (Invitrogen).

RESULTS

A strategy for the capture of direct protein interactions within RNPs

To study protein proximities and changing vicinal interactomes within heterogeneous RNPs along dynamic RNA maturation pathways, we have designed a novel targeted cross-linking MS approach. The approach uses a cross-linker anchor site encoded in an epitope tag within a multi-protein complex of interest, thus targeting one end of the cross-linker to a single known site in the complex. This combination-affinity tag, named CH (Cysteine—Histidine) tag, is comprised of a single cysteine, followed by a 'DP' (aspartic acid-proline) moiety, and a polyhistidine sequence; the tag is flanked by one arginine on either side for trypsin digestion (Figure 1B). Our approach takes advantage of the comparative rarity of cysteines, in particularly surface ones,

in proteins in many organisms, as the observed frequency of cysteines is only ∼1.2% in the yeast proteome, ∼1.9% in *Drosophila* and ∼2.3% in mouse and human (30), making this approach feasible in many organisms. The epitope tag-encoded cysteine serves as the anchoring site for a heterobifunctional cross-linker, while the poly-his sequence allows enrichment of the peptides cross-linked to the anchor site after proteolytic cleavage with trypsin. This enrichment, targeting the epitope tag within the complex of interest rather than the cross-linker, reduces the complexity of the cross-linked peptide mixture to manageable levels while increasing the yield of cross-linked species to a detectable level by MS. In addition, the DP moiety provides a secondary means of detecting cross-linked peptides in the subsequent MS/MS analysis through a characteristic neutral loss fragmentation pattern (31).

To further reduce data complexity and maximize cross-linking signal-to-noise ratios, we combined the CH-tag with an affinity purification step prior to chemical cross-linking via an Protein A (PrA)-tagged bait protein within the same complex, allowing for further targeting of specific protein assemblies and reduced background. We have previously shown that our single-step affinity purification protocol is highly effective in preserving transient interactions, while minimizing non-specific contaminants due to cryolytic cell lysis and short handling times (23). This first step is therefore suitable to provide intact complexes to determine protein proximities and complex organization under close to *in vivo* conditions. In this study, selected 'anchor' proteins were C-terminally tagged by homologous recombination of the CH-tag to the only endogenous copy of its gene in the *S. cerevisiae* genome. However, the CH-tag can also be fused to the N-terminus of the anchor protein and thus be used to probe different parts of a protein assembly and the spatial of organization and proximity of its different components.

Our approach uses a commercial heterobifunctional cross-linker, $SM(PEG)_n$, with a sulfhydryl-specific C-terminal maleic acid imide (maleimide) group targeting the cysteine anchor site within the CH-tag, while on its N-terminus it carries a NHS ester, targeting lysines in nearby proteins (Figure 1C) (15). Of additional advantage is that the reaction chemistries of maleimide and NHS ester are defined by distinct pH ranges, allowing for a more controllable two-stage reaction, while minimizing random targeting of the cross-linker (15,16). The $SM(PEG)_2$ used in this study contains two polyethylene glycol (PEG) molecules for a total spacer arm length of ∼17.6Å between the reactive groups. Therefore it makes for an appropriate cross-linking agent to study the extended milieu around individual complex components, and thus the architecture of macromolecular complexes.

Targeted crosslinking-MS analysis of Mex67-PrA:Mtr2-CH

Recently, a large portion of the *S. cerevisae* Mex67:Mtr2 heterodimer has been crystalized (19,20), which made it a suitable proof-of-principle model allowing the prediction of potential cross-links based on the location of the CH-tag and cross-linker length using XWalk (28). Moreover, the heterodimer is part of the mRNA maturation pathway, with changing interaction partners from the nucleus, through the NPC, to the cytoplasm and cross-linking experiments may provide some insight not only in the vicinal interactome of the heterodimer, but also into this dynamic pathway (2).

The Mex67:Mtr2 heterodimer was affinity purified via PrA-tagged Mex67 on magnetic beads, with the CH-tag fused to Mtr2. With the purified complexes still immobilized on the beads, cross-linking was carried out *ex vivo* in a two-step pH-dependent reaction, as maleimide reacts with cysteines at between pH 6.5–7.5, while NHS-ester only reacts with lysines at pH 7–9 (Figure 2) (15,16). First, $SM(PEG)_2$ was added at a concentration of 100 μM at pH6.6 for 30 min at room temperature for attachment to cysteines. After removal of excess cross-linker, the NHS ester reaction was activated in buffer at pH8.8, while simultaneously quenching further sulfhydryl reactivity (Figure 2). Cross-linked complexes were digested with trypsin in an 'on-bead' reaction for subsequent analysis by MS. As the CH-tag is flanked by one arginine on either side, digestion with trypsin released a 'R**C**DPHHHHHHHHHH' peptide cross-linked to lysine-carrying peptides via its cysteine. However, the peptides resulting from the trypsin digest were likely to resemble a mixture of uncross-linked, cross-linked (to internal cysteines, if present) and CH-tag containing cross-linked peptides, as well as containing multiple types of cross-links (32). To enrich specifically for CH-tag containing cross-linked peptides, the peptide mixture was incubated with histidine-binding silica coated nitrilotriacetic acid resin encased in micro-tips (UptiTip), specifically designed to purify polyhistidine peptides from small sample volumes (Figure 3A). After stringent washing, peptides were eluted and lyophilized. The re-suspended peptides were separated by reverse-phase chromatography and subjected to liquid-chromatography-tandem MS (LC-MS/MS). A typical dataset from a targeted cross-linking sample yielded ∼7000–16 000 MS/MS spectra. Samples taken prior to trypsin digest were analysed by western blot and Silver staining (Figure 4A; Supplementary Figure S1).

The Mex67-PrA/Mtr2-CH complex was purified under low or high salt conditions prior to sequential cross-linking. Under low stringency conditions (25 mM NaCl) ∼100 unique cross-links were initially identified in pLink, of which an average of ∼80 confirmed cross-links remained after filter application and final validation (Table 1). Under high salt conditions prior to cross-linking (1M NaCl), ∼160 unique crosslinks were initially identified, of which ∼130 were confirmed cross-links by filter application and final validation (Table 2). We also compared MS data from cross-linked Mex67-PrA/Mtr2-CH to those from cross-linked and non-cross-linked Mex67-PrA control samples, and non-cross-linked Mex67-PrA/Mtr2-CH samples in low stringency conditions, having undergone the exact same procedure as described above (Figure 4A; Supplementary Figure S1). The LC-MS/MS did not yield any spectra matching the cross-linked CH-tag from these control samples, suggesting that our approach has successfully minimized background below our detection range and has a robust low false positive rate (data not shown).

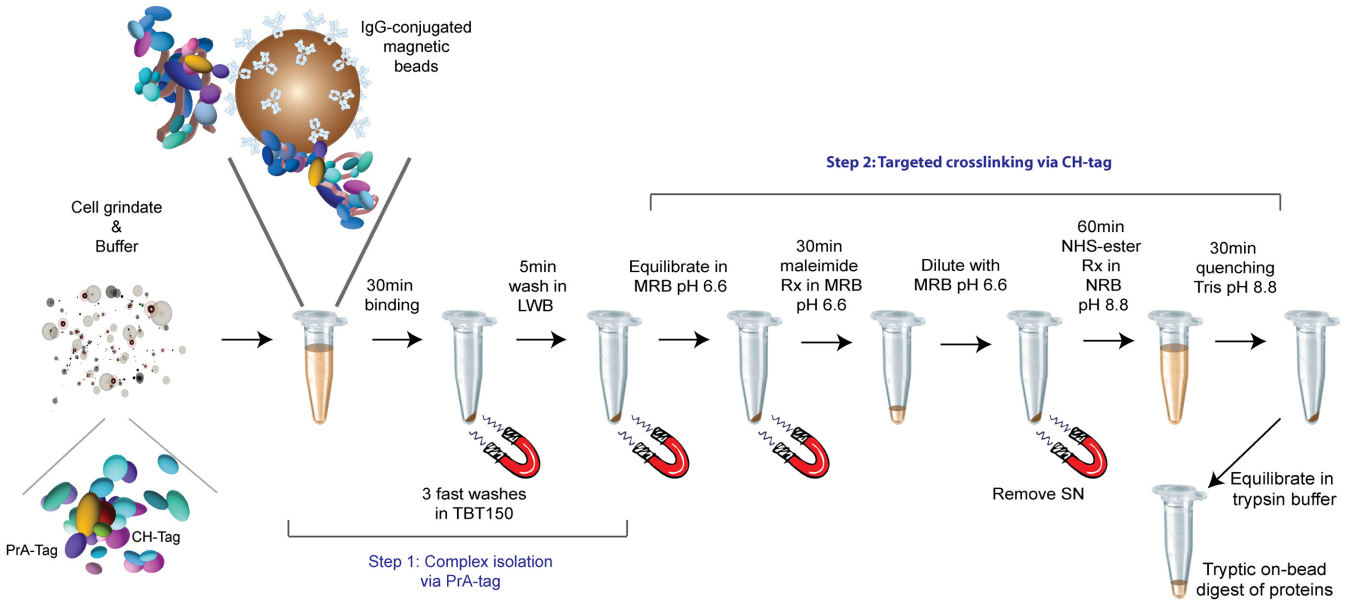

Figure 2. Single-step affinity purification and sequential crosslinking workflow. Cell grindate is mixed with an optimized buffer, and, after a short incubation with antibody-conjugated magnetic beads, the purified complexes are cross-linked in a pH-depended two-step reaction. Complexes are tryptically-digested 'on-bead' prior to enrichment of cross-linked peptides and mass spectrometry.

Table 1. Peptides cross-linked to Mtr2-CH under low stringency conditions

Proteins IDs	Peptides cross-linked	Total spectra	Best *E*-value	error ppm
MEX67(119)-MTR2(186)	RYDPQTKLLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	9	6.17E-10	1.28666
MEX67(119)-MTR2(186)	YDPQT**K**LLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	14	1.88E-09	0.26181
MEX67(120)-MTR2(186)	RYDPQT**K**LLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	7	1.25E-10	0.69694
MEX67(120)-MTR2(186)	YDPQT**K**LLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	66	7.31E-15	1.70396
MEX67(228)-MTR2(186)	ELLMTNNPIT**T**DKLYR-**C**DPHHHHHHHHHH	3	1.98E-07	−3.14549
MEX67(230)-MTR2(186)	ELLMTNNPITTD**K**LYR-**C**DPHHHHHHHHHH	7	1.25E-10	−0.35063
MEX67(435)-MTR2(186)	YNHGYNS**T**SNNKLSK-CDPHHHHHHHHHH	1	1.19E-06	−1.29123
MTR2(186)-MEX67(210)	**C**DPHHHHHHHHHH-SLEVW**K**NK	5	2.29E-06	0.17896
MTR2(186)-MEX67(24)	**C**DPHHHHHHHHHH-I**K**ISVR	5	2.96E-06	0.35174
MTR2(186)-MEX67(256)	**C**DPHHHHHHHHHH-DEQ**K**LQTVYSLPMK	3	4.19E-08	0.24637
MTR2(186)-MEX67(75)	**C**DPHHHHHHHHHH-AEAESLM**K**WNGVR	9	5.97E-12	0.27108
NSP1(336)-MTR2(186)/ NSP1(431)-MTR2(186)	SNEDKQDGTAKPA**FSFG**AK-CDPHHHHHHHHHH	1	2.68E-04	2.86407
NSP1(338)-MTR2(186)/ NSP1(433)-MTR2(186)	SNEDKQDGTA**K**PA**FSFG**AK-CDPHHHHHHHHHH	1	8.82E-08	1.89692
NSP1(384)-MTR2(186)	DGDASKPA**FSFG**A**K**PDENK-CDPHHHHHHHHHH	1	7.01E-04	−1.94757
NSP1(395)-MTR2(186)	ASATS**K**PA**FSFG**AKPEEK-CDPHHHHHHHHHH	1	1.17E-04	−0.99050

Proteins and positions of cross-linked amino acids are shown, as well as sequences of cross-linked peptides identified by MS with cross-linked amino acids marked in bold. FG-region within Nup159 and Nsp1 are underlined. The total number of spectra for each cross-linked peptide pair, best *E*-value and corresponding mass error in ppm are also indicated.

Proximities and spatial organization within the Mex67:Mtr2 heterodimer

Based on a recently published partial structure of the Mex67:Mtr2 heterodimer (20), we used XWalk (28) to predict four lysines and one threonine on Mex67 as potential cross-link sites, taking the C-terminal position of the CH-tag on Mtr2 and the length of SM(PEG)$_2$ as constraints. Out of these five predicted sites, we identified four in our spectra: T119, K120, K343 and K348, of which the first two were observed with high frequencies (Figure 4C; Tables 1 and 2); the spectrum for K120 is shown in Figure 4B. Moreover, we identified nine additional cross-linked sites within Mex67 with varying peptide frequencies (Figure 5A, frequency indicated by line thickness; Tables 1 and 2); out of these, seven were cross-links to lysine (K24, K75, K154, K210, K230, K256, K439) and two to threonine residues (T228, T435), the latter making up ∼ 3% of NHS ester reactions (Figures 4C and 5A; Tables 1 and 2) (26). We further compared cross-links identified between Mex67-PrA and Mtr2-CH under high (1M NaCl) and low (25 mM NaCl) stringency affinity purification conditions, and out of the 13

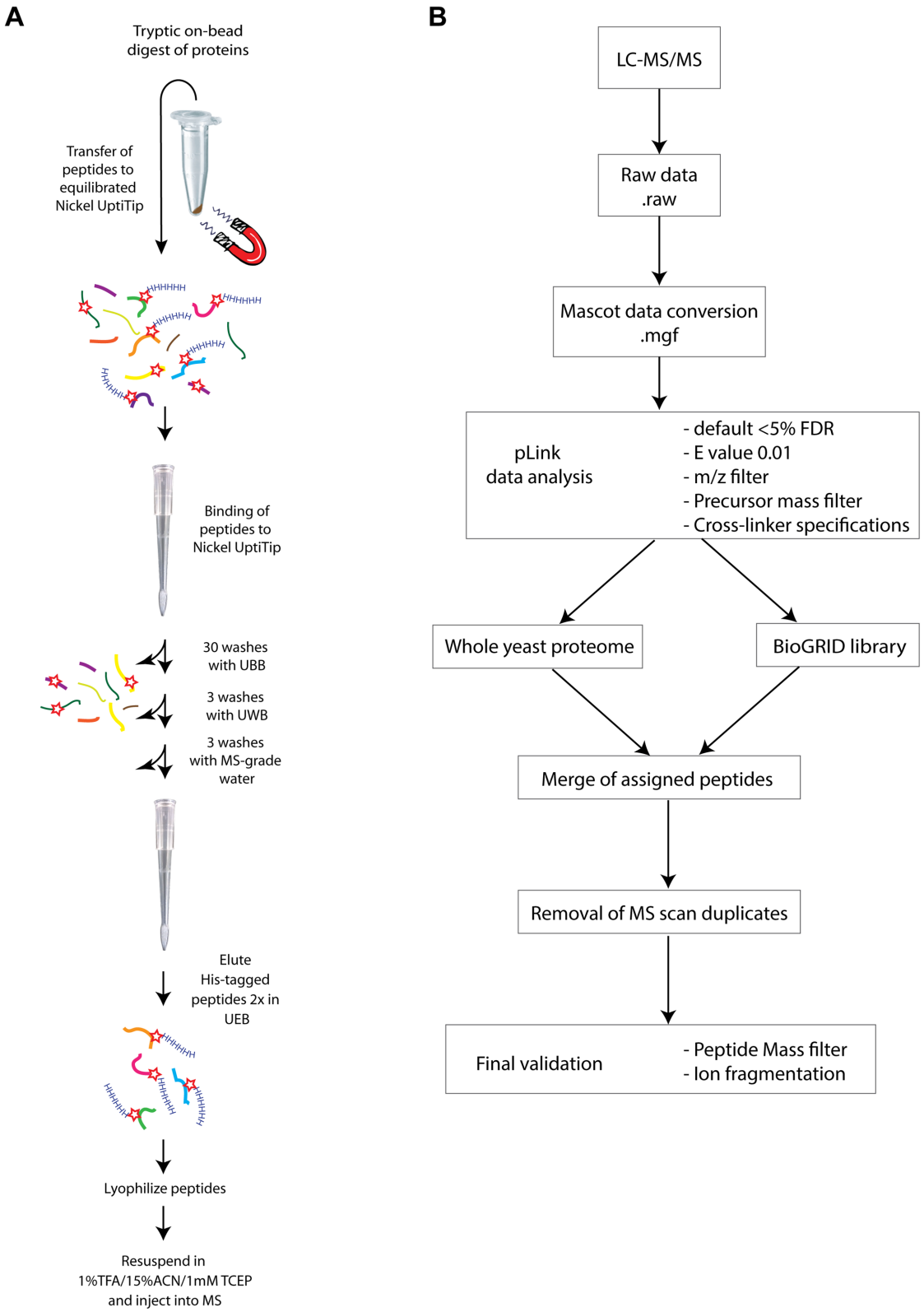

Figure 3. Cross-linked peptide enrichment and MS data analysis. (**A**) Enrichment of cross-linked peptides. The peptide mixture is incubated on nickel coated micropipette tips to enrich for polyhistidine-containing peptides (His_{10} encoded in the CH-tag). Eluted peptides are lyophilized prior to LC-MS/MS analysis. (**B**) Outline of XL-MS data analysis workflow. Raw LC-ESI-MS/MS data files are converted to Mascot format (.mgf) and used for pLink analysis (25). The overall data is reduced by filtering (false discovery rate and *E*-value thresholds, alignment score, precursor mass, isotopic mass, cross-linker specifications, e.i. reactivity and mass), prior to peptide alignment against two libraries (yeast whole proteome and BioGRID interactome for the bait protein). The results are merged and MS scan duplicates removed prior to final peptide validation based on peptide mass and ion fragmentation.

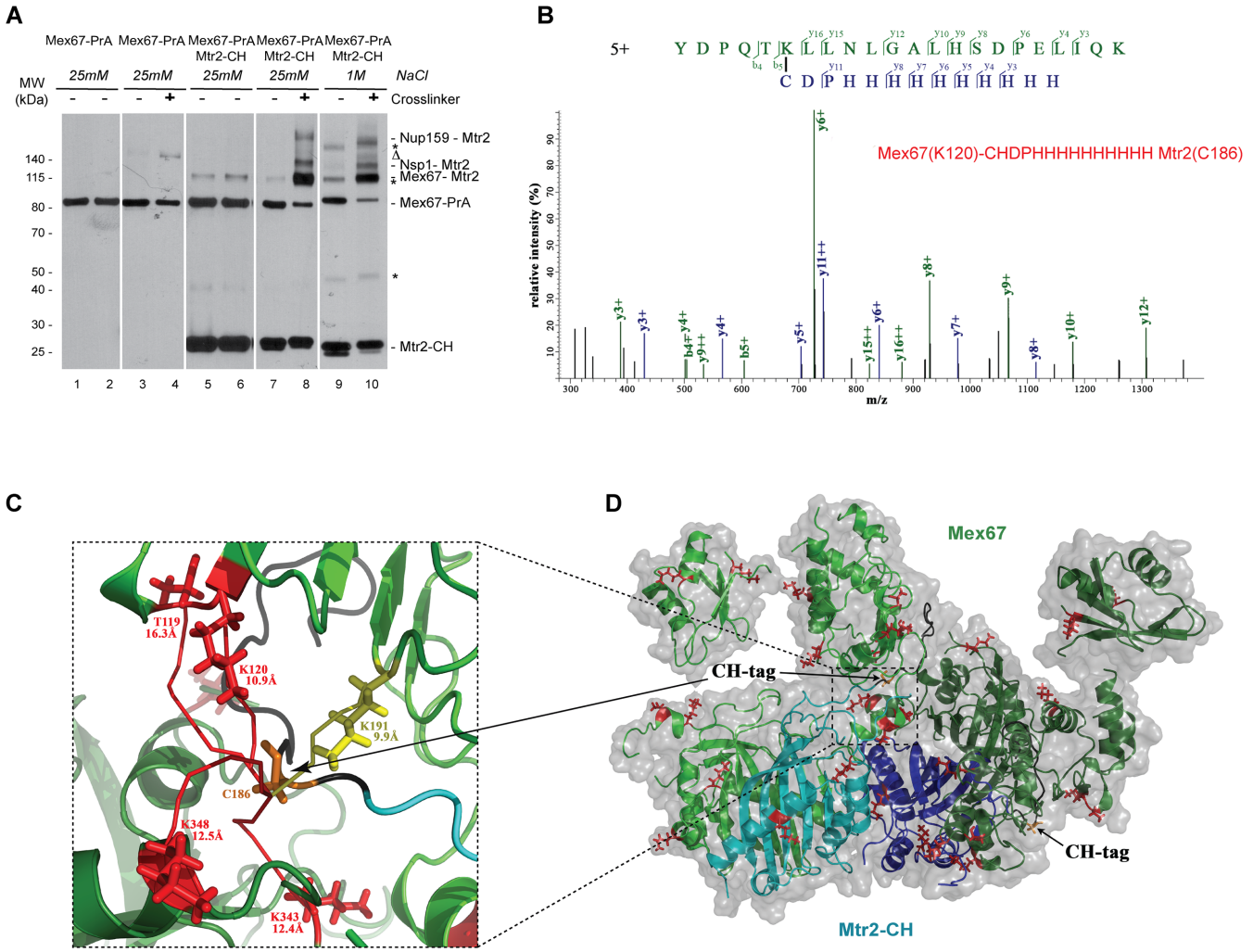

Figure 4. Prediction and validation of Mtr2-CH anchored cross-links on the Mex67-PrA/Mtr2-CH complex. (**A**) Mex67-PrA and Mex67-PrA/Mtr2-CH complexes were purified under low salt (lanes 1–8) or high salt conditions (lanes 9 and 10). Samples were taken before (lanes 1, 3, 5, 7, 9) and after cross-linking (lanes 3, 8, 10), or prior to trypsin digest in control samples (lanes 2, 6). Proteins cross-linked to Mtr2-CH were resolved by western blotting and visualized by anti-His antibody (lanes 5–10). Anti-His antibody cross-reactive bands are indicated with an asterisk (*). Mex67-PrA was visualized with peroxidase anti-peroxidase antibody (lanes 1–10). A Mex67-Mex67 dimer in the cross-linked control sample (Lane 4) is denoted with Δ. Cross-linked proteins are inferred by molecular weight of bands and listed in Tables 1 and 2. (**B**) MS/MS spectra of the validated cross-link between C186 of Mtr2-CH and Mex67 K120. Peptides and fragment ions derived from Mex67 or Mtr2 shown in green and blue, respectively. (**C**) Close-up view of XWalk-predicted cross-links (28) between the cysteine of the Mtr2-CH (C186; orange) and nearby lysines on Mex67 (PDB ID: 4WWU.PDB) (19). Predicted and confirmed crosslinks are depicted in red (T119, K120, (K343, K348), and predicted cross-links that were not identified in yellow (K191). The lines are the solvent accessible surface distances (SASD) measurements, which correspond to the length of the shortest path between Mtr2 (C186) and nearby lysines, where the path leads through solvent occupied space without penetrating the protein surface. (**D**) Representation of a Mex67:Mtr2 tetramer complex obtained recently (PDB ID: 4WWU.PDB) (19), to which the cross-linking anchoring CH-tag (dark grey) containing the cross-linker anchoring cysteine (orange) has been added at the C-terminal of Mtr2 in I-TASSER (27). Mex67 molecules are coloured light and dark-green, Mtr2 in cyan and blue. Identified cross-links within Mex67and Mtr2 are depicted in red.

peptides detected in high salt, we detected nine also under low salt conditions (Tables 1 and 2). The observed peptide frequencies for those nine peptides remained the same in both conditions. It is possible that the residues not identified in low salt were inaccessible due to either steric reasons or additional protein interactions under these conditions (23). We also detected 'intra-cross-links' within Mtr2 in high salt (K25, K111), possibly due to a conformation change in the absence of other complex components under these conditions (Table 2).

Mex67 contains four different structurally distinct domains, an N-terminal RNA recognition motif (RRM) domain followed by a leucine rich repeat (LRR), a nuclear-transport factor2-like domain (NTF2L) and a C-terminal ubiquitin-associated domain (UBA), with the Mex67-NFT2L domain forming the interface with Mtr2 (3). A recent X-ray crystallography and small-angle X-ray scattering (SAXS) study of Mex67$^{\Delta UBA}$:Mtr2 suggests a more stable interaction between the Mex67-LRR domain and Mex67-NTF2L:Mtr2 interface (20), which lies within the proximity of the Mtr2 C-terminus and CH-tag. Consistent with this

Table 2. Peptides cross-linked to Mtr2-CH-tag under high salt conditions

Proteins IDs	Peptides cross-linked	Total spectra	Best *E*-value	error ppm
MEX67(119)-MTR2(186)	RYDPQ**T**KLLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	9	3.60E-08	1.54399
MEX67(119)-MTR2(186)	YDPQ**T**KLLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	3	7.66E-09	0.74327
MEX67(120)-MTR2(186)	RYDPQT**K**LLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	4	4.00E-10	0.15226
MEX67(120)-MTR2(186)	YDPQT**K**LLNLGALHSDPELIQK-**C**DPHHHHHHHHHH	48	3.98E-15	0.86751
MEX67(228)-MTR2(186)	ELLMTNNPITTD**K**LYR-**C**DPHHHHHHHHHH	6	3.51E-10	0.67839
MEX67(230)-MTR2(186)	ELLMTNNPITTD**K**LYR-**C**DPHHHHHHHHHH	11	2.85E-13	−0.50307
MEX67(435)-MTR2(186)	YNHGYNS**T**SNNKLSK-**C**DPHHHHHHHHHH	2	4.74E-07	−0.36090
MEX67(439)-MTR2(186)	YNHGYNSTSNN**K**LSK-**C**DPHHHHHHHHHH	2	3.74E-10	0.30476
MTR2(186)-MEX67(154)	**C**DPHHHHHHHHHH-MFPAMM**K**LASTEK	1	5.31E-04	0.48592
MTR2(186)-MEX67(210)	**C**DPHHHHHHHHHH-SLEVW**K**NK	17	4.48E-06	0.17896
MTR2(186)-MEX67(24)	**C**DPHHHHHHHHHH-I**K**ISVR	18	1.01E-07	0.30411
MTR2(186)-MEX67(256)	**C**DPHHHHHHHHHH-DEQ**K**LQTVYSLPMK	3	1.97E-09	0.16515
MTR2(186)-MEX67(343)	**C**DPHHHHHHHHHH-NIS**K**VSSEK	3	1.19E-06	1.92656
MTR2(186)-MEX67(348)	**C**DPHHHHHHHHHH-VSSE**K**SIQQR	2	5.82E-04	−0.04724
MTR2(186)-MEX67(75)	**C**DPHHHHHHHHHH-AEAESLM**K**WNGVR	21	9.27E-12	−0.28820
MTR2(186)-**MTR2**(25)	**C**DPHHHHHHHHHH-**K**ILAHLDDPDSNK	30	1.85E-18	2.00599
MTR2(111)-MTR2(186)	D**K**MGQDATVPIQPNNTGNR-**C**DPHHHHHHHHHH	2	1.18E-11	0.28748
MTR2(186)-**GLE1**(273)	**C**DPHHHHHHHHHH-**K**ADVNVR	2	2.88E-04	0.28417
MTR2(186)-NSP1(319)/ MTR2(186)-NSP1(414)	CDPHHHHHHHHHH-KDDNSS**K**PA**FSFG**AK	5	9.03E-09	0.74739
MTR2(186)-NSP1(346)/ MTR2(186)-NSP1(441)	**C**DPHHHHHHHHHH-PA**FSFG**A**K**PAEK	1	7.04E-06	0.51479
MTR2(186)-NSP1(355)/ MTR2(186)-NSP1(450)	**C**DPHHHHHHHHHH-NNNE**T**SKPA**FSFG**AK	5	1.13E-05	−0.08550
NSP1(191)-MTR2(186)	**T**EPDKPA**FS**FNSSVGNK-**C**DPHHHHHHHHHH	2	6.80E-05	−0.99253
NSP1(195)-MTR2(186)	TEPD**K**PA**FS**FNSSVGNK-**C**DPHHHHHHHHHH	4	3.69E-05	0.25790
NSP1(298)-MTR2(186)	AGA**T**SKPA**FSFG**AKPEEK-**C**DPHHHHHHHHHH	1	1.27E-04	0.55777
NSP1(300)-MTR2(186)	AGATS**K**PA**FSFG**AKPEEK-**C**DPHHHHHHHHHH	1	1.18E-05	0.40138
NSP1(338)-MTR2(186)/ NSP1(433)-MTR2(186)	SNEDKQDGTA**K**PA**FSFG**AK-**C**DPHHHHHHHHHH	5	2.03E-21	1.28622
NSP1(376)-MTR2(186)	DGDAS**K**PA**FSFG**AKPDENK-**C**DPHHHHHHHHHH	2	2.70E-09	−1.10896
NSP1(509)-MTR2(186)	DSGSS**K**PA**FSFG**AKPDEK-**C**DPHHHHHHHHHH	2	2.01E-06	0.15251
NUP159(515)-MTR2(186)	**PAFG**AIA**K**EPSTSEYA**FG**K-CDPHHHHHHHHHH	1	4.28E-04	2.43676
NUP159(526)-MTR2(186)	EPSTSEYA**FG**K**P**S**FG**APS**FG**SGK-CDPHHHHHHHHHH	2	2.85E-05	−0.15866
MTR2(186)-NUP159(829)	**C**DPHHHHHHHHHH-LTETI**K**K	1	5.66E-07	0.46376
MTR2(186)-NUP159(849)	**C**DPHHHHHHHHHH-NPV**FG**NHV**K**AK	4	5.61E-05	0.26671
NUP159(751)-MTR2(186)	SPFSSF**T**KDDTENGSLSK-**C**DPHHHHHHHHHH	1	1.78E-04	−0.96953
NUP159(752)-MTR2(186)	SPFSSFT**K**DDTENGSLSK-**C**DPHHHHHHHHHH	1	9.64E-07	0.17926
NUP159(840)-MTR2(186)	SANIDMAGL**K**NPV**FG**NHVK-**C**DPHHHHHHHHHH	2	1.85E-05	−0.58125

Proteins and positions of cross-linked amino acids are shown, as well as sequences of cross-linked peptides identified by MS with cross-linked amino acids marked in bold. FG-regions within Nup159 and Nsp1 are underlined. The total number of spectra for each cross-linked peptide pair, best *E*-value and corresponding mass error in ppm are also indicated.

structural data, we identified cross-linked peptides within the Mex67-LRR helix and domain that form part of the Mex67:Mtr2 interaction surface (T228, K230, K256) (Figures 4D and 5A; Tables 1 and 2) (20). Moreover, the same study observed multiple arrangements of the Mex67-RRM domain, positioning the RRM in either close or more distant proximity to the Mex67:Mtr2 interface, suggesting a flexible linker between the RRM and LRR domains (20). It has been proposed that this linker may be sufficiently flexible to accommodate various binding partners such as mRNA or other proteins. Our data resembles a closer proximity of the RRM domain to the Mex67:Mtr2 interface and CH-tag under the conditions tested, as shown by cross-links to K24, K75, T119 and K120, all of which are located within the RRM and were observed with high frequency (Figures 4D and 5A; Tables 1 and 2). This could be due the isolation of a RNA-bound Mex67:Mtr2 complex, as previous work has shown that Mex67-PrA assemblies affinity-purified under these conditions contain intact RNAs (23). Alternatively, the RRM domain could be trapped in a closed confirmation as a result of cross-linking.

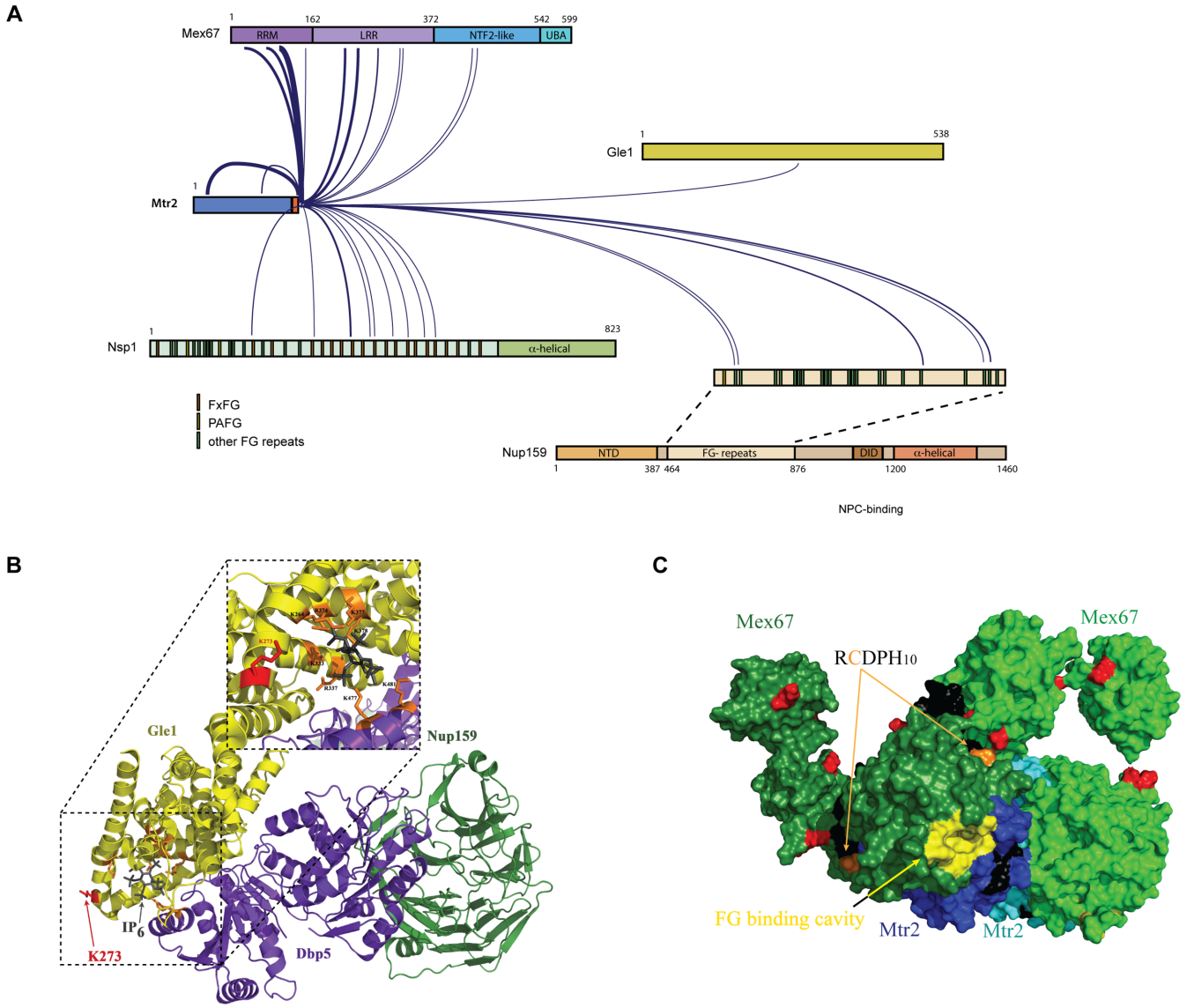

Figure 5. Cross-linking of Mtr2 to nuclear pore components. (**A**) Visualization of the validated cross-links on the affinity-purified Mex67-PrA-Mtr2-CH complex using SM(PEG)$_2$. The primary structure of all identified proteins is shown, and specific regions are indicated. Line thickness represents frequency of cross-linked peptides. (**B**) Structure of Nup159-Dbp5-Gle1-InsP$_6$ obtained previously (PDB ID: 3RRM) generated in PyMol (38). Nup159 is depicted in green, Dbp5 in purple, Gle1 in yellow and InsP$_6$ as grey sticks. InsP$_6$ interacting residues on Dbp5 and Gle1 are labelled in orange; the residue cross-linked in Gle1 (K273) is highlighted in red. Inlet on top shows a close-up of cross-link identified adjacent to Dbp5-Gle1-InsP$_6$ binding interface (PDB ID: 3RRM) (38) generated in PyMol. Dbp5 is depicted in purple, Gle1 in yellow and InsP$_6$ as grey sticks. InsP$_6$ interacting residues on Dbp5 and Gle1 are shown in orange, and the identified cross-link in Gle1 (K273) within proximity of the Mtr2-CH is highlighted in red. (**C**) FG binding cavity position relative to the CH tags within the Mex67-Mtr2 structure (PDB ID: 4WWU). Both Mex67 and Mtr2 are colour coded as described in Figure 4D. Cross-linked amino acids (red) anchored to the cysteine (orange) are also shown. The FG binding cavity represented in yellow corresponds to a predicted FG binding cavity from *Candida albicans* (*Ca*). Based on sequence alignments, amino acids properties and surface accessibility, the cavity would be composed of Mtr2 amino acids I88, P89 (*Ca* L92 And F99, respectively) as well as Mex67 amino acids S304, Q306, I392, V452, M454, A461 and I459 (*Ca* Q332, E334, L423, I482, G484, A492 and I490, respectively).

Mex67:Mtr2—nuclear pore interactome

High salt affinity purification conditions prior to cross-linking appeared to stabilize what are believed to be more transient interactors of the Mex67:Mtr2 heterodimer, such as terminal nuclear pore proteins (nucleoporins or nups) (Tables 1 and 2). In yeast, the Mex67:Mtr2 complex is the principal export receptor that mediates mRNA transport across the NPC (33). In agreement with the role of Mex67:Mtr2 as a nuclear export factor for RNA, we identified peptides from three nucleoporins cross-linked to Mtr2-CH: Nup159, Nsp1 and Gle1, presumably caught in close proximity to Mex67:Mtr2 complexes in the process of transporting cargos across the NPC (Figure 5A; Tables 1 and 2). *In vitro* reconstitution studies have previously shown that the dimer interacts with phenylalanine-glycine (FG) repeat motifs of nuclear pore proteins, the FG nups (17). FG nups consist of differing subtypes of FG repeats (e.g. FxFG, PSFG, SAFG, PAFG or GLFG), which provide

docking sites for various transport receptors that shuttle proteins and macromolecular complexes across the NPC (34–36). A preference of interaction of Mex67:Mtr2 with specific sites or even subtypes of FG repeats within FG nups has not been determined so far. Targeted cross-linking from the C-terminus of Mtr2-CH resulted in eleven peptides cross-linked to the FG region of the central channel nucleoporin Nsp1, between residues 191 and 521 (Figure 5A). Interestingly, the cross-linked residues are all located in the middle portion of Nsp1, in particular around the FxFG repeats, and nine of them are located within two to five residues from FxFG moieties (Tables 1 and 2). No cross-linked residues were identified near the N-terminal PAFG, or the NPC-anchoring α-helical C-terminus, suggesting a Mex67:Mtr2 preference for the FxFG region within Nsp1 during its translocation through the NPC.

Similarly, the seven cross-links identified within the nucleoporin Nup159, which forms part of the fibrils on the cytoplasmic face of the NPC, were also restricted to a limited region within the protein: the FG region between amino acids 515 and 851 (Figure 5A, Table 2). Nup159 plays an important role during the release of exporting mRNPs from the NPC, together with the DEAD-box ATPase Dbp5 and the NPC-associated protein Gle1. Nup159's propeller domain was shown to interact with a Dbp5-Gle1 complex, and its FG domain has been implicated in regulating terminal mRNP release (37,38). No cross-links were identified in either the N-terminal beta-propeller or the NPC-anchoring C-terminal α-helical domain of Nup159, suggesting a specific positioning of the Mex67:Mtr2 heterodimer in proximity to Nup159 *in vivo* for mRNP release. In line, we identified one cross-linked site within Gle1 (K273) (Figure 5A) directly adjacent to the inositol hexakisphosphate ($InsP_6$) binding pocket of Gle1 (Figure 5B) (37,38). Binding of Dbp5 to Gle1 and its cofactor $InsP_6$ activates the ATPase function of Dbp5 and stimulates a conformational change in the protein that leads to the dissociation of mRNA binding factors, including Mex67:Mtr2 (38–39). It is thus conceivable that the formation of a Mex67:Mtr2–Gle1–$InsP_6$–Dbp5–Nup159 complex is required for the remodelling and release of mRNA.

Targeted cross-linking-MS analysis of the pre-ribosomal Nop7 subcomplex

As a second complex to test our approach, we selected the trimeric Nop7 subcomplex. The heterotrimer is highly conserved among eukaryotes and each of its members, Nop7, Erb1 and Ytm1, is essential (40). It is located primarily in the nucleolus and required for the assembly and maturation of 60S ribosomal subunits (41,42). While a number of genetic and biochemical studies have determined aspects of the functional relationship of the subcomplex, direct interactions between its components have only been inferred from yeast two-hybrid, *in vitro* GST pulldown assays and co-immunoprecipitation experiments (20). To determine if our method could capture the proximities between some of the subcomplexes' components, we carried out reciprocal targeted cross-linking in which the CH-tag was placed on either the C-terminus of Ytm1 or Nop7 to determine the spatial proximity between these two proteins. The Nop7-PrA/Ytm1-CH and Ytm1-PrA/Nop7-CH complexes were affinity purified under medium stringency conditions (150 mM NaCl; Supplementary Figure S2) prior to sequential cross-linking, peptide enrichment and MS. pLink initially identified ∼105 and 20 unique cross-links, respectively. Out of these ∼100 and 20 were confirmed following filtering and manual validation (Tables 3 and 4). Using Ytm1-CH as the SM $(PEG)_2$ anchoring site, we identified a total of nine cross-linked peptides: five mapped close to the C-terminus of Ytm1 itself (K365, T408, K437, K453, K459), one to the Ytm1 N-terminal region (K74), two to sites within Nop7 (K342, K600), and one within the ribosomal protein Rpl8 (K63) (Figure 6A-blue lines and Table 3). Ytm1 contains seven WD40 repeats, which are located between residues S101 and K453 and which are predicted to fold into a circularized beta-propeller structure common to WD40 repeat-containing proteins, bringing the C-terminal residues in close proximity of each other (Figure 6B–Ytm1 in blue) (43). Five of the six Ytm1-Ytm1 intra-cross-links were located in WD40 domains 5–7, and within proximity of the C-terminus and the CH-tag, supporting the Ytm1 structure prediction (Figure 6B and C). Based on an I-TASSER model, the Ytm1 N-terminus is comprised of a more disordered region; hence, cross-linking within the N-terminal region of Ytm1 (K74) suggests that this disordered region is within close vicinity of the C-terminus of the protein (Figure 6B) (17,20).

While it was previously proposed that the C-terminus of Ytm1, in particular WD40 domains 6 and 7 between A362 and N460, interact with Erb1, we did not find Erb1 located within the proximity of the Ytm1 C-terminus. Interestingly, we did however identify cross-links between Ytm1-CH and Nop7, indicating a distance close enough for a direct interaction that had not been previously observed. One of the identified cross-linked peptides was located at the very C-terminal end of Nop7 (K600) (Figure 6A–C-Nop7 in green; Table 3), and was supported by a reciprocal cross-link from Nop7-CH (C607) to the C-terminus of Ytm1 (K459), positioning the C-termini of Nop7 and Ytm1 in close proximity to one another (Figure 6A-red lines, Figure 6C and Table 4). Another cross-link within Nop7 (from Ytm1-CH) was identified at K342, which falls between the first predicted coiled-coil region (residues 298–339) and the BRCT domain (residues 357–439) (44). Orientation of this region of Nop7 towards the C-terminus of Ytm1, and thus within spatial proximity of the Ytm1-CH, has previously been suggested by Tang *et al*. (20) (Figure 6B and C and Table 4). The functional significance of the orientation and close proximity of Ytm1 and Nop7 to one another remains to be determined in future experiments (Figure 6B and C and Table 4). Besides one cross-linked region within Ytm1, all other identified cross-links, using Nop7-CH as anchor, were located within the C-terminal region of Nop7 itself (K533, K571, K593, K597, K600), which is part of a lysine-rich coiled-coil region (Figure 6A–C and Table 4). We also observed cross-linking of Ytm1-CH to Rpl8 (K63), a large ribosomal subunit protein that has recently been implicated in the recruitment of the Nop7 subcomplex to the pre-60S ribosomes (Figure 6A and Table 3) (42).

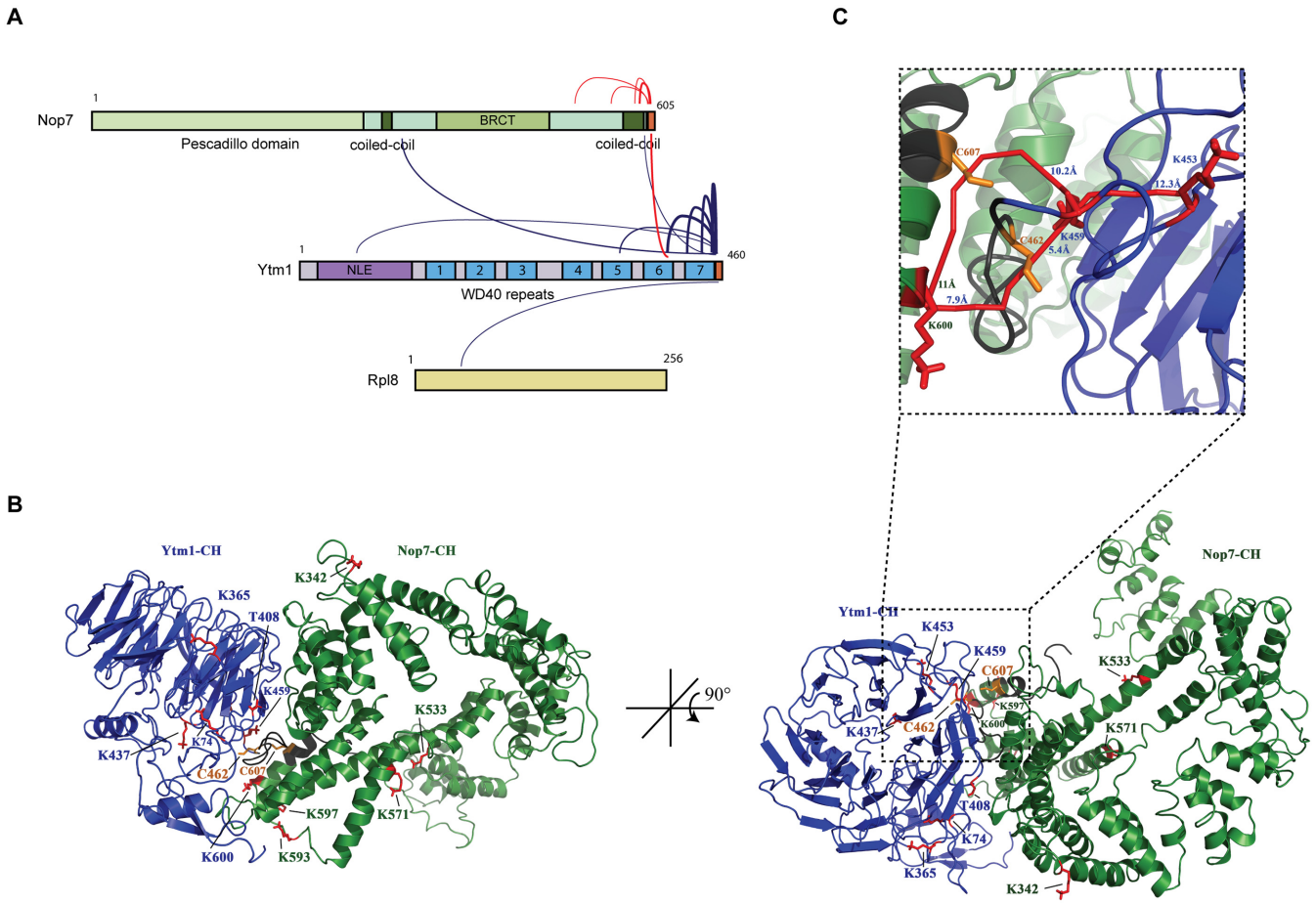

Figure 6. Cross-linking of the pre-ribosomal Nop7 subcomplex. (**A**) Visualization of the validated cross-links on the affinity-purified Nop7-PrA/Ytm1-CH and Ytm1-PrA/Nop7-CH complexes using $SM(PEG)_2$. The primary structure of all identified proteins is shown and specific regions are indicated. Line thickness represents frequency of cross-linked peptides. Cross-links originating from Ytm1-CH are depicted as blue lines, cross-links originating from Nop7-CH as red lines. (**B**) Structural models of Nop7 (green) and Ytm1 (blue) were generated using I-TASSER, and a docking model for their interaction was built translating XL-MS information into spatial restraints using identified intermolecular cross-links in HADDOCK (29). The CH-tags are depicted in grey, cross-linker anchor cysteines are indicated. Two views of the Ym1-Nop7 complex are presented, rotated by 90° on the *x*-axis. (**C**) Predicted and confirmed cross-links between CH-tagged Nop7 and Ytm1. Close-up view of XWalk-predicted cross-links (28) between the CH-tags (dark grey, embedded cysteines in orange) of Ytm1 (blue) and Nop7 (green) and nearby lysines are shown. Predicted and confirmed crosslinks are depicted in red. The lines are the solvent accessible surface distances (SASD) measurements, which correspond to the length of the shortest path between the CH-tag encoded cysteine and nearby lysines, where the path leads through solvent occupied space without penetrating the protein surface.

Table 3. Peptides cross-linked to Ytm1-CH

Proteins IDs	Peptides cross-linked	Total spectra	*E*-value	error ppm
Ytm1(462)-Ytm1(459)	**C**DPHHHHHHHHHH-GDNIF**K**NR	36	3.94E-08	0.11756
Ytm1(462)-Ytm1(453)	**C**DPHHHHHHHHHH-IQIN**K**GDNIFK	15	5.10E-09	−0.41168
Ytm1(462)-Ytm1(408)	**C**DPHHHHHHHHHH-S**T**SPMYTITR	29	2.22E-11	−0.23341
Ytm1(462)-Ytm1(74)	**C**DPHHHHHHHHHH-TSLHDYLT**K**K	1	4.94E-06	0.10250
Ytm1(462)-Ytm1(365)	**C**DPHHHHHHHHHH-VGASS**K**VTQQQLIGHK	4	1.57E-09	0.46282
Ytm1(462)-Ytm1(437)	**C**DPHHHHHHHHHH-WAE**K**VGIISAGQDK	11	1.98E-10	0.34134
Ytm1(462)-**Nop7**(342)	**C**DPHHHHHHHHHH-N**K**GDILIQPSK	3	4.10E-07	−0.54858
Ytm1(462)-Nop7(600)	**C**DPHHHHHHHHHH-LN**K**LDSK	1	1.08E-05	0.35673
Ytm1(462)-**Rpl8**(63)	**C**DPHHHHHHHHHH-**K**ILSIR	2	1.16E-05	0.05832

Proteins and the positions of the amino acids cross-linked together are shown, as well as the sequences of cross-linked peptides identified by MS with the position of each amino acids cross-linked together. The total number of spectra for each cross-linked peptide pair, *E*-value and corresponding mass error in ppm are also indicated.

Table 4. Peptides cross-linked to Nop7-CH

Proteins IDs	Peptides cross-linked	Total spectra	*E*-value	error ppm
Nop7(607)-Nop7(597)	**C**DPHHHHHHHHHH-A**K**LNK	2	1.08E-05	−1.46492
Nop7(607)-Nop7(600)	**C**DPHHHHHHHHHH-LN**K**LDSK	8	8.40E-07	0.08830
Nop7(607)-Nop7(571)	**C**DPHHHHHHHHHH-M**K**YSNAK	1	1.64E-06	−0.27318
Nop7(607)-Nop7(593)	**C**DPHHHHHHHHHH-QIA**K**QK	1	2.30E-04	−2.07383
Nop7(607)-Nop7(533)	**C**DPHHHHHHHHHH-YSETSEAD**K**DVNK	1	4.73E-08	−0.32862
Nop7(607)-**Ytm1**(459)	R**C**DPHHHHHHHHHH-GDNIF**K**NR	7	5.51E-06	−0.12427

Proteins and positions of cross-linked amino acids are shown, as well as sequences of cross-linked peptides identified by MS with cross-linked amino acids marked in bold. The total number of spectra for each cross-linked peptide pair, *E*-value and corresponding mass error in ppm are also indicated.

DISCUSSION

Here we described a targeted cross-linking-MS approach that allowed us to study the vicinal interactomes and changing milieu of three RNP components. As RNPs are large, heterogeneous and dynamic complexes, affinity purification followed by MS is not sufficient to gain information on the composition of vicinal protein-neighbourhoods or proximities of their individual components. However, as RNA and RNPs are being established as having significant importance in cellular regulation (45), determining the changing vicinal interactome of factors and protein proximities within complexes throughout RNP assembly is imperative for discerning their individual roles in the regulation and advancement of, and connections between pathways. Chemical cross-linking-MS methods, employing homobifunctional lysine–lysine cross-linkers, are commonly used to refine subcomplexes of known architecture to gain deeper insights into sub-complexes of known structure (11–13). However, due to the non-discriminate nature of homobifunctional cross-linkers and the heterogeneity, dynamicity and complexity, determining changing vicinal interactomes within RNPs is extremely challenging using this method, and no label transfer reagents are available that attach to one unique site in a protein. Conversely, anchoring a cross-linker to a known target site within a complex, however, allows us to probe the interactomes around selected RNP components. Using a low-frequency residue as anchor site increases the probability of cross-linking to this anchor at optimized concentrations.

First, we have applied this method to the mRNP export adapter heterodimer Mex67:Mtr2 with the CH-tag located on the C-terminal end of Mtr2. The availability of the heterodimer's crystal structure (19), provided an ideal proof-of-principle complex enabling us predict cross-linked lysines based on anchor tag (CH-tag) position and cross-linker length using XWalk (28). Validating our approach we identified four out of five XWalk-predicted cross-linked lysine residues in Mex67, all of which were within a calculated distance of 10.9–16.3Å (Figure 4C). The fifth cross-link was within a calculated distance of <9.9Å and it is conceivable that either the distance was too close for the flexibility of the $(PEG)_2$ spacer arm that separates the NHS ester and maleimide groups in the cross-linker (Figure 1C), or since shorter distance cross-links were identified between Nop7 and Ytm1, that the CH-tag is in an unfavourable confirmation towards this residue. In addition, we identified nine additional cross-linked peptides within Mex67 with varying spectra frequencies (between 1 and 66; Tables 1 and 2), and two within Mtr2 itself. Overall, the peptides identified within the Mex67-LRR helix and the domain that form part of the Mex67:Mtr2 interaction surface are consistent with recent structural data on the complex (19). In addition, high frequency cross-links identified within the RRM domain (K24, K75, T119 and K120) suggest a close, RNA-bound confirmation of the *ex vivo* Mex67:Mtr2 complex (19). Alternatively, the RRM domain could have been trapped in such a confirmation as a result of cross-linking, or resemble an intermediate state; so far no intermediate confirmations have been identified, moreover, the residues of the RRM had to have been sufficiently close for crosslinking to occur.

In agreement with the role of Mex67:Mtr2 as the mRNA nuclear export factor, we also identified peptides from three nucleoporins cross-linked to Mtr2-CH: Nsp1, Nup159 and Gle1, presumably caught in the process of transporting Mex67:Mtr2 complexes across the NPC (Figure 5). Previous *in vitro* studies have shown that Mex67:Mtr2 interacts with FxFG repeats of Nsp1, GLFG repeats of both Nup116 and Nup100, as well as FG repeats of Nup159 (16). A preference of interaction of Mex67:Mtr2 with specific sites or even subtypes of FG repeats within FG nups *in vivo* has not been determined so far. Using targeted cross-linking, we provide first illustrations of close proximity between Mex67:Mtr2 and (Fx)FG repeats within Nsp1 and Nup159. We identified nine cross-linked peptides within Nsp1 all of which were located in immediate (2–7 residues) proximity of FxFG repeats (Figure 5A). No cross-linked residues were identified near the N-terminal PAFG and GLFG sequences or the NPC-anchoring α-helical C-terminus, suggesting a closer proximity of Mex67:Mtr2 to the FxFG region within Nsp1 during its translocation through the NPC, which could indicate a preference for this FG-repeat subtype across this region. A preference for FxFG over other FG repeats has previously been shown for NTF2, the import receptor of RanGDP and the UBA-domain of Tap/NFX1, the mammalian homologue of Mex67 (46). Mtr2 contains an NTF2-like fold and furthermore interacts with the NTF2-like domain within Mex67 as further demonstrated by our crosslinking data (Figures 4D and 5A) (47). Mtr2 also forms an FG binding pocket similar to that of NTF2, which involves residues well conserved among its homologues, and, in support of our data, is positioned to place FG nups upon binding within the vicinity of the CH-tag (Figure 5C) (46). Thus, it is conceivable that Mex67:Mtr2 may have a similar preference for FxFG over other FG subtypes. Previous studies implied that transport across the NPC, in particular that of large mRNA-containing complexes, requires transport receptors to break

bonds that were suggested to form between specific FG subtypes, such as GLFG, that otherwise create a meshwork barrier within the channel of the NPC (48,49). Interestingly, the FxFG repeats of Nsp1 that are in close proximity to Mex67:Mtr2 are so-called non-bond forming FGs (35,36), suggesting that (despite the huge size of their mRNP cargoes) no such bonds are involved or need to be broken by the transporting Mex67:Mtr2-mRNPs to traverse the NPC (21). However, it cannot be discounted that other regions of the Mex67:Mtr2 heterodimer could be in close proximity to other FG subtypes. In addition, it can also not be excluded that the Mtr2-Nsp1 cross-links may perhaps reflect Mtr2 being imported through the pore.

In line with exporting Mex67:Mtr2 RNPs, seven cross-linked peptides were identified between Mtr2-CH and FG repeats of Nup159, in addition to a residue within Gle1, which clearly indicate a close positioning of the Mex67:Mtr2 heterodimer to Nup159-Dbp5-Gle1 mRNP release platform at the cytoplasmic face of the nuclear pore, and we speculate that this specific positioning of Mex67:Mtr2-mRNP cargo may be the prelude to mRNP remodelling by activated Dbp5 (50). Binding of Dbp5 to Gle1 and its cofactor inositol hexakisphosphate ($InsP_6$) activates the ATPase function of Dbp5 that catalyzes mRNP remodelling where a subset of factors such as Mex67:Mtr2 and Nab2 are removed from mRNAs, releasing the cargo in the cytoplasm for translation (37,39). Interaction of the heterodimer with the Nup159 FG region, located between the propeller domain and the NPC targeting region, would position it in close proximity to the proposed Dbp5 activity (Figure 5B). This idea is further supported by the identified cross-linked peptide within Gle1, directly adjacent to its $InsP_6$ binding pocket.

As was our initial aim, we did identify a number of components of the dynamic vicinity interactome of the Mex67:Mtr2 heterodimer. Previous studies have identified a genetically and biochemically interaction between Mtr2 and Nup85 (33). While we did not find Nup85 in the C-terminal vicinity neighbourhood of Mtr2, it is conceivable that protein was either beyond the distance of the cross-linker, or not within the proximity of the C-terminus. Besides the proteins described above, we also identified low-level cross-linking from Mtr2-CH to a number of known mRNP components and interactors: Yra1, Nab2, Sac3, Mlp1 and Mlp2 (data not shown) (2–3,51). For Yra1 and Nab2, the cross-linked peptides located within the domains identified for Mex67:Mtr2 interaction (3). However, for all five proteins, the spectra obtained were of low quality and did not pass our stringent validation.

Dynamic vicinal interactions within the pre-ribosomal Nop7 complex

The second complex we subjected to our targeted cross-linking-MS approach is part of nucleolar 60S pre-ribosomes, the Nop7 complex, whose components are required for correct ribosome assembly and pre-rRNA processing (41,42). It had previously been suggested that Nop7 and Ytm1, two components of the trimeric subcomplex, while interacting directly with Erb1, do not interact with one another (17,20). However, supported by reciprocal targeted cross-linking, here we demonstrate that the C-termini of both Nop7 and Ytm1 are within close proximity of one another (5.4–12.3Å; Figure 6). Previous functional studies have suggested an interaction between Erb1 and Ytm1 involving Ytm1's WD40 domains 6 and 7, which was confirmed by a recent crystal structure in *Chaetomium thermophilum* (52). Interestingly, the interaction between ChYtm1 and ChErb1 was shown to occur on the face opposite to that of the predicted Ytm1–Nop7 interaction (based on the Ytm1–Nop7 cross-links reported here) thus placing Ytm1–Erb1 interacting residues outside the solvent accessible path of the cross-linker. Thus, in line with our data, the ChErb1–ChYtm1 crystal structure supports a docking model prediction between Nop7 and Ytm1 illustrating the close proximity of both C-termini (Figure 6B and C) (52); however, it should be noted that the model may not represent folding of either protein with 100% accuracy. While the cross-links observed between Nop7 and Ytm1-CH also support the previously suggested orientation of the C-terminus of Ytm1 towards the region between the first predicted coiled-coil region and BRCT domain of Nop7 (20), any functional significance still remains to be determined. We also identified a cross-linked peptide within Rpl8, originating from Ytm1-CH (Figure 6A). Rpl8 was shown to bind near the proximal stem formed by base-pairing between the 3′ end of 5.8S ribosomal RNA (rRNA) and the 5′ end of 25S rRNA (53). UV-induced protein–RNA cross-linking (CRAC) showed that Erb1 and Nop7 are bound to the pre-rRNA in a region adjacent to that bound by Rpl8 (54). Moreover, the Nop7 subcomplex was found absent from pre-60S ribosomes in cells depleted of Rpl8, suggesting that Rpl8 is required for the association of the subcomplex with pre-ribosomes (42). A binding site for Ytm1 on the pre-rRNA could not be determined (Granneman, personal communications) and it is possible that Ytm1 does not interact with the rRNA but rather establishes and regulates protein interactions, as has been suggested for WD40 repeat proteins (17,43). In this context, it may be conceivable that an interaction of Ytm1 with Rpl8 could be involved in the recruitment of the Nop7 subcomplex to pre-ribosomes; however, this remains to be determined. Overall, our data supports a model for a Nop7-Ytm1 interaction, while suggesting that Erb1 is not within close proximity of either the Ytm1 or the Nop7 C-terminus, or beyond the radius of the cross-linker distance. Moreover, it also provides evidence for a close proximity of Rpl8 and Ytm1.

A novel tool to study the organization of heterogeneous complexes

The distinct and reproducible specificity of cross-links within discrete regions of near-neighbour proteins with functional relevance (i.e. FG domains within Nsp1 and Nup159) yet not within their structural domains, provides strong evidence that our approach reads out functional interactions. In addition, we demonstrated that the method allows discovery of so far unidentified near-neighbor relationships that had not previously been identified. The consistently observed variability in frequencies for different identified cross-linked peptides leads to the question whether it is possible that these frequencies may correlate

with the residency time of a protein within the vicinity of the 'anchor' protein: lower frequencies reflecting short residency times (i.e. due to more transient interactions, complex dynamics or rapid conformational changes within the complex), while higher peptide numbers indicate longer residency times (i.e. due to more stable interactions/complexes, fewer conformational changes). Examples are the cross-linked peptides between Mtr2 and the nuclear pore components Nsp1, Nup159 and Gle1, which occurred with a frequency of 1–5 spectra/peptide and are presumably of transient nature, compared to peptides of the more stable regions within Mex67:Mtr2 heterodimer, for which up to 66 spectra were identified (Tables 1 and 2).

While the RNP complexes in this work were chosen from *S. cerevisiae*, the presented approach is not limited to yeast, but can readily be adapted to other systems including mammalian-derived ones using CRISPR/Cas9 integration of the CH-tag. It is also not limited to the C-terminal region of bait proteins, and the tag can be placed N-terminal. Moreover, one could envisage CH-tagging multiple proteins of the same complex by removing the 5′ arginine from the tag, leaving a terminal peptide attached to each tagged protein to identify from which proteins each cross-link originates. SM(PEG)$_n$ cross-linkers corresponding to various length are also available, with spacer arms ranging from 17.6 to 95.2Å. While traditional cross-linking/MS methods have been invaluable to position domains within known crystal structures to refine 3D architecture of complexes, they are less suited to uncover protein proximities and dynamic interactomes within heterogeneous RNPs. Our approach does not seek to determine the structural organization of complexes but rather to investigate dynamic vicinal interactomes and protein neighbourhoods within RNPs and along their changing pathway milieu. Given its ease, wide applicability and high discovery potential, we believe our method presents a highly useful and important addition to the technical repertoire of the scientific community for the study of dynamic and heterogeneous macromolecular complexes.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank D. Faubert from the IRCM Proteomics platform for his support and help provided during this project. We also thank S. Fan from the Institute of Computing Technology at the Chinese Academy of Sciences for his modifications of pLink. D. Zenklusen (Université de Montréal), M. Rout and B. Chait (Rockefeller University) and B. Montpetit (University of Alberta) for feedback and critical reading of the manuscript, and S. Granneman (University of Edinburgh) for personal communications, and all the members of the Oeffinger laboratory for their support.

FUNDING

Canadian Institutes of Health Research [FRN 106628]; National Science and Engineering Research Council of Canada [RGPIN 386315]; CIHR New Investigator Award (to M.O.); Fonds de recherche du Québec—Santé Chercheur Boursier Junior I (to M.O.). Funding for open access charge: National Science and Engineering Research Council of Canada [RGPIN 386315].
Conflict of interest statement. None declared.

REFERENCES

1. Woolford,J.L. and Baserga,S.J. (2013) Ribosome biogenesis in the yeast Saccharomyces cerevisiae. *Genetics*, **195**, 643–681.
2. Oeffinger,M. and Zenklusen,D. (2012) To the pore and through the pore: A story of mRNA export kinetics. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, **1819**, 494–506.
3. Niño,C.A., Hérissant,L., Babour,A. and Dargemont,C. (2013) mRNA nuclear export in yeast. *Chem. Rev.*, **113**, 8523–8545.
4. Perez-Fernandez,J., Roman,A., De Las Rivas,J., Bustelo,X.R. and Dosil,M. (2007) The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell Biol.*, **27**, 5414–5429.
5. Talkish,J., Zhang,J., Jakovljevic,J., Horsey,E.W. and Woolford,J.L. (2012) Hierarchical recruitment into nascent ribosomes of assembly factors required for 27SB pre-rRNA processing in Saccharomyces cerevisiae. *Nucleic Acids Res.*, **40**, 8646–8661.
6. Gamalinda,M., Ohmayer,U., Jakovljevic,J., Kumcuoglu,B., Woolford,J., Mbom,B., Lin,L. and Woolford,J.L. (2014) A hierarchical model for assembly of eukaryotic 60S ribosomal subunit domains. *Genes Dev.*, **28**, 198–210.
7. Baßler,J., Kallas,M., Pertschy,B., Ulbrich,C., Thoms,M. and Hurt,E. (2010) The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Mol. Cell*, **38**, 712–721.
8. Bradatsch,B., Leidig,C., Granneman,S., Gnädig,M., Tollervey,D., Böttcher,B., Beckmann,R. and Hurt,E. (2012) Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. *Nat. Struct. Mol. Biol.*, **19**, 1234–1241.
9. Matsuo,Y., Granneman,S., Thoms,M., Manikas,R.-G., Tollervey,D. and Hurt,E. (2013) Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. *Nature*, **505**, 112–116.
10. Rappsilber,J., Siniossoglou,S., Hurt,E.C. and Mann,M. (2000) A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. *Anal. Chem.*, **72**, 267–275.
11. Jennebach,S., Herzog,F., Aebersold,R. and Cramer,P. (2012) Crosslinking-MS analysis reveals RNA polymerase I domain architecture and basis of rRNA cleavage. *Nucleic Acids Res.*, **40**, 5591–5601.
12. Hurt,E. and Beck,M. (2015) Towards understanding nuclear pore complex architecture and dynamics in the age of integrative structural analysis. *Curr. Opin. Cell Biol.*, **34**, 31–38.
13. Greber,B.J., Bieri,P., Leibundgut,M., Leitner,A., Aebersold,R., Boehringer,D. and Ban,N. (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. *Science*, **348**, 303–308.
14. Smyth,D.G., Blumenfeld,O.O. and Konigsberg,W. (1964) Reactions of N-ethylmaleimide with peptides and amino acids. *Biochem. J.*, **91**, 589–595.
15. Lomant,A.J. and Fairbanks,G. (1976) Chemical probes of extended biological structures: synthesis and properties of the cleavable protein cross-linking reagent [35S]dithiobis(succinimidyl propionate). *J. Mol. Biol.*, **104**, 243–261.
16. Strasser,K., Bassler,J. and Hurt,E. (2000) Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. *J. Cell Biol.*, **150**, 695–706.
17. Miles,T.D., Jakovljevic,J., Horsey,E.W., Harnpicharnchai,P., Tang,L. and Woolford,J.L. (2005) Ytm1, Nop7, and Erb1 form a complex necessary for maturation of yeast 66S preribosomes. *Mol. Cell Biol.*, **25**, 10419–10432.

18. Fribourg,S. and Conti,E. (2003) Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. *EMBO Rep.*, **4**, 699–703.
19. Aibara,S., Valkov,E., Lamers,M. and Stewart,M. (2015) Domain organization within the nuclear export factor Mex67:Mtr2 generates an extended mRNA binding surface. *Nucleic Acids Res.*, **43**, 1927–1936.
20. Tang,L., Sahasranaman,A., Jakovljevic,J., Schleifman,E. and Woolford,J.L. Jr (2008) Interactions among Ytm1, Erb1, and Nop7 required for assembly of the Nop7-subcomplex in yeast preribosomes. *Mol. Biol. Cell*, **19**, 2844–2856.
21. Rout,M., Aitchison,J., Suprapto,A., Hjertaas,K., Zhao,Y. and Chait,B. (2000) The yeast nuclear pore complex: composition, architecture, and transport mechanism. *J. Cell Biol.*, **148**, 635–651.
22. Oeffinger,M., Wei,K.E., Rogers,R., Degrasse,J.A., Chait,B.T., Aitchison,J.D. and Rout,M.P. (2007) Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat. Methods*, **4**, 951–956.
23. Granneman,S., Petfalski,E., Swiatkowska,A. and Tollervey,D. (2010) Cracking pre-40S ribosomal subunit structure by systematic analyses of RNAi[ndash]protein cross-linking. *EMBO J.*, **29**, 2026–2036.
24. Hirosawa,M., Hoshida,M., Ishikawa,M. and Toya,T. (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. *Comput. Appl. Biosci.*, **9**, 161–167.
25. Yang,B., Wu,Y.-J., Zhu,M., Fan,S.-B., Lin,J., Zhang,K., Li,S., Chi,H., Li,Y.-X., Chen,H.-F. *et al.* (2012) Identification of cross-linked peptides from complex samples. *Nat. Methods*, **9**, 904–906.
26. Kalkhof,S., Haehn,S., Ihling,C., Paulsson,M., Smyth,N. and Sinz,A. (2008) Determination of disulfide bond patterns in laminin beta1 chain N-terminal domains by nano-high-performance liquid chromatography/matrix-assisted laser desorption/ionization time-of-flight/time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.*, **22**, 1933–1940.
27. Zhang,Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
28. Kahraman,A., Malmström,L. and Aebersold,R. (2011) Xwalk: computing and visualizing distances in cross-linking experiments. *Bioinformatics*, **27**, 2163–2164.
29. de Vries,S.J., van Dijk,M. and Bonvin,A.M.J.J. (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.*, **5**, 883–897.
30. Miseta,A. and Csutora,P. (2000) Relationship between the occurrence of cysteine in proteins and the complexity of organisms. *Mol. Biol. Evol.*, **17**, 1232–1239.
31. Martin,D.B., Eng,J.K., Nesvizhskii,A.I., Gemmill,A. and Aebersold,R. (2005) Investigation of neutral loss during collision-induced dissociation of peptide ions. *Anal. Chem.*, **77**, 4870–4882.
32. Tang,X. and Bruce,J.E. (2010) A new cross-linking strategy: protein interaction reporter (PIR) technology for protein-protein interaction studies. *Mol. Biosyst.*, **6**, 939–947.
33. Santos-Rosa,H., Moreno,H., Simos,G., Segref,A., Fahrenkrog,B., Pante,N. and Hurt,E. (1998) Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Mol. Cell Biol.*, **18**, 6826–6838.
34. Strambio-De-Castillia,C., Niepel,M. and Rout,M.P. (2010) The nuclear pore complex: bridging nuclear transport and gene regulation. *Nat. Rev. Mol. Cell Biol.*, **11**, 490–501.
35. Patel,S.S., Belmont,B.J., Sante,J.M. and Rexach,M.F. (2007) Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell*, **129**, 83–96.
36. Yamada,J., Phillips,J.L., Patel,S., Goldfien,G., Calestagne-Morelli,A., Huang,H., Reza,R., Acheson,J., Krishnan,V.V., Newsam,S. *et al.* (2010) A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Mol. Cell Proteomics*, **9**, 2205–2224.
37. Adams,R.L., Terry,L.J. and Wente,S.R. (2014) Nucleoporin FG domains facilitate mRNP remodeling at the cytoplasmic face of the nuclear pore complex. *Genetics*, **197**, 1213–1224.
38. Montpetit,B., Thomsen,N.D., Helmke,K.J., Seeliger,M.A., Berger,J.M. and Weis,K. (2011) A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export. *Nature*, **472**, 238–242.
39. Ledoux,S. and Guthrie,C. (2011) Regulation of the Dbp5 ATPase cycle in mRNP remodeling at the nuclear pore: a lively new paradigm for DEAD-box proteins. *Genes Dev.*, **25**, 1109–1114.
40. Rohrmoser,M., Holzel,M., Grimm,T., Malamoussi,A., Harasim,T., Orban,M., Pfisterer,I., Gruber-Eber,A., Kremmer,E. and Eick,D. (2007) Interdependence of Pes1, Bop1, and WDR12 controls nucleolar localization and assembly of the PeBoW complex required for maturation of the 60S ribosomal subunit. *Mol. Cell Biol.*, **27**, 3682–3694.
41. Harnpicharnchai,P., Jakovljevic,J., Horsey,E., Miles,T., Roman,J., Rout,M., Meagher,D., Imai,B., Guo,Y., Brame,C. *et al.* (2001) Composition and functional characterization of yeast 66s ribosome assembly intermediates. *Mol. Cell*, **8**, 505–515.
42. Jakovljevic,J., Ohmayer,U., Gamalinda,M., Talkish,J., Alexander,L., Linnemann,J., Milkereit,P. and Woolford,J.L. (2012) Ribosomal proteins L7 and L8 function in concert with six A3 assembly factors to propagate assembly of domains I and II of 25S rRNA in yeast 60S ribosomal subunits. *RNA*, **18**, 1805–1822.
43. Wall,M.A., Coleman,D.E., Lee,E., Iñiguez-Lluhi,J.A., Posner,B.A., Gilman,A.G. and Sprang,S.R. (1995) The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell*, **83**, 1047–1058.
44. Hölzel,M., Grimm,T., Rohrmoser,M., Malamoussi,A., Harasim,T., Gruber-Eber,A., Kremmer,E. and Eick,D. (2007) The BRCT domain of mammalian Pes1 is crucial for nucleolar localization and rRNA processing. *Nucleic Acids Res.*, **35**, 789–800.
45. Narla,A., Narla,A., Ebert,B.L. and Ebert,B.L. (2010) Ribosomopathies: human disorders of ribosome dysfunction. *Blood*, **115**, 3196–3205.
46. Bayliss,R., Leung,S.W., Baker,R.P., Quimby,B.B., Corbett,A.H. and Stewart,M. (2002) Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO J.*, **21**, 2843–2853.
47. Senay,C., Ferrari,P., Rocher,C., Rieger,K.-J., Winter,J., Platel,D. and Bourne,Y. (2003) The Mtr2-Mex67 NTF2-like domain complex. Structural insights into a dual role of Mtr2 for yeast nuclear export. *J. Biol. Chem.*, **278**, 48395–48403.
48. Frey,S., Richter,R.P. and Görlich,D. (2006) FG-rich repeats of nuclear pore proteins form a three-dimensional meshwork with hydrogel-like properties. *Science*, **314**, 815–817.
49. Ribbeck,K. and Gorlich,D. (2002) The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO J.*, **21**, 2664–2671.
50. Lund,M. and Guthrie,C. (2005) The DEAD-box protein Dbp5p is required to dissociate Mex67p from exported mRNPs at the nuclear rim. *Mol. Cell*, **20**, 645–651.
51. Lei,E., Lei,E.P., Stern,C.A., Stern,C., Fahrenkrog,B., Fahrenkrog,B., Krebber,H., Krebber,H., Moy,T., Moy,T.I. *et al.* (2003) Sac3 is an mRNA export factor that localizes to cytoplasmic fibrils of nuclear pore complex. *Mol. Biol. Cell*, **14**, 836–847.
52. Wegrecki,M., Rodríguez-Galán,O., de la Cruz,J. and Bravo,J. (2015) The structure of Erb1-Ytm1 complex reveals the functional importance of a high-affinity binding between two β-propellers during the assembly of large ribosomal subunits in eukaryotes. *Nucleic Acids Res.*, doi:10.1093/nar/gkv1043.
53. Ben-Shem,A., Garreau de Loubresse,N., Melnikov,S., Jenner,L., Yusupova,G. and Yusupov,M. (2011) The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science*, **334**, 1524–1529.
54. Granneman,S., Petfalski,E. and Tollervey,D. (2011) A cluster of ribosome synthesis factors regulate pre-rRNA folding and 5.8S rRNA maturation by the Rat1 exonuclease. *EMBO J.*, **30**, 4006–4019.