Driving integrative structural modeling with serial capture affinity purification

Xingyu Liu[a], Ying Zhang[a], Zhihui Wen[a], Yan Hao[a], Charles A. S. Banks[a], Jeffrey J. Lange[a], Brian D. Slaughter[a], Jay R. Unruh[a], Laurence Florens[a], Susan M. Abmayr[a], Jerry L. Workman[a], and Michael P. Washburn[a,b,1]

[a]Stowers Institute for Medical Research, Kansas City, MO 64110; and [b]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, KS 66045

Edited by Andrej Sali, University of California, San Francisco, CA, and approved October 19, 2020 (received for review April 28, 2020)

Streamlined characterization of protein complexes remains a challenge for the study of protein interaction networks. Here we describe serial capture affinity purification (SCAP), in which two separate proteins are tagged with either the HaloTag or the SNAP-tag, permitting a multistep affinity enrichment of specific protein complexes. The multifunctional capabilities of this protein-tagging system also permit in vivo validation of interactions using acceptor photobleaching Förster resonance energy transfer and fluorescence cross-correlation spectroscopy quantitative imaging. By coupling SCAP to cross-linking mass spectrometry, an integrative structural model of the complex of interest can be generated. We demonstrate this approach using the Spindlin1 and SPINDOC protein complex, culminating in a structural model with two SPINDOC molecules docked on one SPIN1 molecule. In this model, SPINDOC interacts with the SPIN1 interface previously shown to bind a lysine and arginine methylated sequence of histone H3. Our approach combines serial affinity purification, live cell imaging, and cross-linking mass spectrometry to build integrative structural models of protein complexes.

chromatin | epigenetics | integrative structural modeling | cross-linking mass spectrometry | quantitative imaging

Estimates of the number of human protein–protein interactions (HPPIs) continue to grow. In 2017, 625,641 HPPIs were predicted (1), while a subsequent estimate yielded nearly 1 million HPPIs (2). This number will likely continue to increase considering the vast number of cells in the human body and myriad of different cellular conditions in normal and diseased states. New methods are needed to tackle the enormous challenge of validating and determining the structural and functional significance of proposed HPPIs and how they are organized into larger networks inside of cells. One way to tackle this challenge is to develop computational methods to identify potential direct protein–protein interactions (3) and then use these predictions to guide further experimental studies. Another established experimental approach is affinity purification followed by mass spectrometry (APMS), in which an affinity-tagged protein is purified along with its interactors, which are then identified by liquid chromatography coupled to mass spectrometry (LCMS). This APMS approach has been successful, but there are computational challenges associated with distinguishing nonspecific interactions (4). Foundational work by Rigaut et al. (5) described an affinity purification approach for the study of *Saccharomyces cerevisiae* protein complexes, in which a protein was fused to two affinity tags, enabling a two-step enrichment method that resulted in protein complexes of higher purity. This TAP-tag method has played a key role in the analysis of *S. cerevisiae* protein complexes and protein interaction networks (6). Recently, more advanced multifunctional tags, such as the HaloTag (7) and SNAP-tag (8), have been developed and can be used for both affinity purification and microscopy imaging methods.

Combining these concepts, we have devised a strategy to study any pair of proteins that might directly associate to 1) validate their interactions via proteomics, 2) assess their interactions in live cells, and 3) build a structural model of the complex. Here we describe the development of reagents for serial capture affinity purification (SCAP), an approach fundamentally different from conventional affinity purification. SCAP uses a combination of two separately tagged bait proteins to reduce the complexity and increase the purity of protein complexes. SCAP can be followed by label-free quantitative mass spectrometry analysis (SCAP-MS) to identify protein complexes containing the two interacting proteins of interest. Furthermore, the SCAP constructs can be used to validate their interaction in vivo using quantitative imaging techniques. Finally, the SCAP pipeline can also include a cross-linking (XL) step followed by MS to identify cross-linked peptides defining interaction interfaces and intermolecular distance constraints, which can be used to build integrative molecular models of protein complexes.

We demonstrate the striking capabilities of this technology using a pair of proteins, Spindlin1 (SPIN1) and SPINDOC (c11orf84), previously proposed to directly interact in biochemical (9) and computational (3) studies. SPIN1 is a well-characterized histone methylation reader (9–21), while SPINDOC has only been defined by its ability to bind SPIN1 (9). In this study, we first characterized the direct interaction and codiffusion of SPIN1 and SPINDOC in live cells using imaging methods. We next used SCAP-MS to generate a sample enriched in SPIN1

Significance

Determining the three-dimensional structures of protein complexes is critically important to guide biological research. Structural models of complexes can be built using powerful integrative approaches that combine emerging technologies in mass spectrometry, molecular modeling, and protein docking; however, preparing enriched biochemical samples suitable for analysis remains a major challenge. Here we describe serial capture affinity purification (SCAP), which can be used for the study of protein interactions in live cells and, when combined with cross-linking mass spectrometry, contribute distance restraints for integrative structural modeling. This broadly applicable technology can be used to study any protein complex in human tissue culture cells. We demonstrate SCAP capabilities on a poorly characterized epigenetic protein complex with roles in human cancer.

Author contributions: X.L., Y.Z., L.F., J.L.W., and M.P.W. designed research; X.L., Y.Z., Z.W., Y.H., C.A.S.B., J.J.L., B.D.S., and J.R.U. performed research; X.L., Y.Z., Y.H., C.A.S.B., J.J.L., B.D.S., J.R.U., and L.F. contributed new reagents/analytic tools; X.L., Y.Z., Z.W., J.J.L., B.D.S., J.R.U., L.F., S.M.A., J.L.W., and M.P.W. analyzed data; and X.L., L.F., and M.P.W. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence may be addressed. Email: mpwashburn70@gmail.com.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2007931117/-/DCSupplemental.

First published November 30, 2020.

and SPINDOC. This fraction was then analyzed using advance XLMS techniques followed by molecular modeling. The culmination of these studies resulted in an integrated structural model of a complex formed by one molecule of SPIN1 and two molecules of SPINDOC.

Results

Building Reagents for SCAP. In a typical APMS study, a protein of interest (POI) is affinity-tagged and transiently expressed in cells. The tagged POI is then used to capture protein complexes from cell extracts, and proteins copurifying with the POI are identified by mass spectrometry. Although this is a well-proven approach, it essentially analyses ex vivo complexes that might not reflect genuine interactions within a cell. As an example, we expressed Halo-SPIN1 (*SI Appendix*, Fig. S1*C*) and Halo-SPINDOC (*SI Appendix*, Fig. S1*D*) separately in HEK293 cells, affinity-purified the associated proteins, and analyzed them by label-free quantitative proteomics (Dataset S1 *B* and *C*). The resulting data demonstrate that the bait protein was the most abundant protein in each sample, with approximately sevenfold less SPINDOC than SPIN1 in the Halo-SPIN1 purification (*SI Appendix*, Fig. S1*C*) and approximately 20-fold less SPIN1 than SPINDOC in the Halo-SPINDOC purification (*SI Appendix*, Fig. S1*D*). There were many additional proteins in both purifications (Dataset S1 *B* and *C*).

There are several challenges in interpreting such APMS datasets: first, defining which proteins genuinely interact with the POI; second, determining which proteins interact directly with the POI; and third, establishing the stoichiometry of the proteins in purified complexes. Considering such challenges, new approaches are needed to better characterize protein complexes in a streamlined and efficient manner. Therefore, we devised the SCAP approach using two orthogonal affinity tags: the HaloTag (22) and the SNAP-tag (8). These tags are multifunctional and facilitate multiple types of analyses using a single construct. For example, the tags can be labeled with fluorophores for live cell imaging in addition to their use for affinity purification. Both the HaloTag and SNAP-tag covalently bind to their respective substrates immobilized on beads. There is an established system for eluting immobilized Halo-tagged proteins isolated from mammalian cell extracts. A linker sequence between the Halo tag and the POI contains a tobacco etch virus (TEV) protease cleavage site (22–24), allowing protease-mediated release of the POI and associated proteins from the beads while leaving the HaloTag bound to the beads (22). In contrast, a system for purifying SNAP-tagged proteins from mammalian cells is not well established.

To develop a SNAP purification strategy that could be used together with Halo purification, allowing independent cleavage of Halo- and SNAP-tagged baits, we needed to choose a protease that recognized a cleavage sequence different from TEV protease. To evaluate the suitability of proteases for eluting SNAP isolated proteins, we constructed N′-terminal SNAP-tag versions of SPIN1 with different protease cleavage sites (*SI Appendix*, Fig. S1*A*). We expressed these constructs in HEK293/FRT cells, isolated SPIN1 using the SNAP-tag, and used the indicated protease to elute SPIN1 for analysis by Western blotting (*SI Appendix*, Fig. S1*B*). The quantity and purity of SNAP-SPIN1 isolated using PreScission protease were comparable to those of SNAP-SPIN1 isolated using TEV protease (*SI Appendix*, Fig. S1*B*). In contrast, enterokinase and Factor Xa protease have shorter recognition sequences than TEV and PreScission protease, and not only cleave the linker, but also appear to cleave SPIN1 (*SI Appendix*, Fig. S1*B*). Therefore, we chose a PreScission protease-based cleavage system as our elution method for SNAP-tag purification. We tagged SPIN1 and SPINDOC at their N termini using Halo or SNAP tag, since N-terminally tagged SPIN1 (25, 26) and SPINDOC (9, 20) have been used in previous functional studies. As the Halo tag (33 kDa) and the SNAP tag (20 kDa) are relatively large tags, we have used an ∼25-aa linker peptide between the tag and bait protein to reduce interference between the tag and the bait protein as they fold.

Developing a Serial Capture Affinity Purification. To perform SCAP, we first constructed an expression vector based on pcDNA5/FRT that would enable us to express both Halo- and SNAP-tagged proteins from the same plasmid. We inserted sequences coding for both the HaloTag and SNAP-tag, each followed by convenient restriction sites for subcloning our two bait proteins (Fig. 1*A* and *SI Appendix*, Fig. S2*A*). As our plasmid generates an mRNA coding for two tagged proteins, we also engineered an internal ribosomal entry site (IRES) between the sequences coding for each protein. We then inserted the open reading frames (ORFs) for the two POIs into the ORF1 and ORF2 regions. These plasmids can be used for either transient or stable expression. The expression of tagged proteins was driven by the cytomegalovirus (CMV) promoter. An anti-SPIN1 antibody recognized both the endogenous SPIN1 (lower band) and tagged versions of SPIN1 (upper band) in the input lanes (*SI Appendix*, Fig. S1*B*). Comparing the intensity of the upper and lower bands, the CMV-driven expression of tagged SPIN1 appears to be close to its endogenous level. The expression levels of recombinant proteins are usually lower in stably transfected cell lines than in transiently transfected cells, as stably transfected lines bear only one copy of the recombinant construct in each cell.

To develop a sequential purification system, we next generated a stable cell line expressing Halo-SPIN1 and SNAP-SPINDOC (Fig. 1*B*). First, proteins from whole-cell extracts were isolated on SNAP affinity beads and then eluted using PreScission protease (fraction E1). We then used 80% of fraction E1 for further affinity purification of tagged complexes using Halo affinity beads, retaining 20% for MS analysis. The unbound supernatant of the Halo purification was collected as fraction UB2. The proteins captured by the Halo affinity beads were eluted using the TEV protease as fraction E2. Proteins from the E1, UB2, and E2 fractions were analyzed by silver-stained sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS/PAGE) (Fig. 1*C*). This analysis clearly indicated a more complex protein mixture in the E1 and UB2 fractions compared with the E2 fraction. The analysis of fraction E2 generated two major bands consistent with enrichment of SPIN1 and SPINDOC proteins, a minor band (25 kDa) consistent with TEV protease, and two other minor unidentified bands (∼70 kDa). The SCAP purification thus generated a sample with high concentrations of the interacting proteins of interest, SPIN1 and SPINDOC, removing most contaminants.

To confirm the presence of SPIN1 and SPINDOC and assess their enrichment at each stage of the purification, the E1, UB2, and E2 fractions were subjected to label-free quantitative proteomic analysis using MudPIT (Dataset S2). The averaged dNSAF values of the top 20 proteins in each fraction are shown in *SI Appendix*, Fig. S2 *B–D*. An overall comparison of the enrichment of the SPINDOC and SPIN1 pair in each fraction is summarized in Fig. 1*D*. Over the course of the two-step SCAP protocol, the enrichment of SPINDOC and SPIN1 was increased by more than 2.5-fold; after SNAP purification, the spectral counts matching these two proteins contributed 15.2% and 8.1% of the total spectral counts, respectively, while after the second purification step, their enrichment was measured as 40.4% and 21%. In both E1 and E2 eluates (Fig. 1*D*), the $dNSAF_{SPINDOC}$:$dNSAF_{SPIN1}$ ratio was ∼2:1, suggesting a stoichiometry of two SPINDOC molecules for each SPIN1 molecule.

Assessing the Interaction of SPIN1 and SPINDOC in Live Cells. Next, using the multifunctional capability of the Halo and SNAP tagging systems, we analyzed the interaction of Halo-SPIN1 and

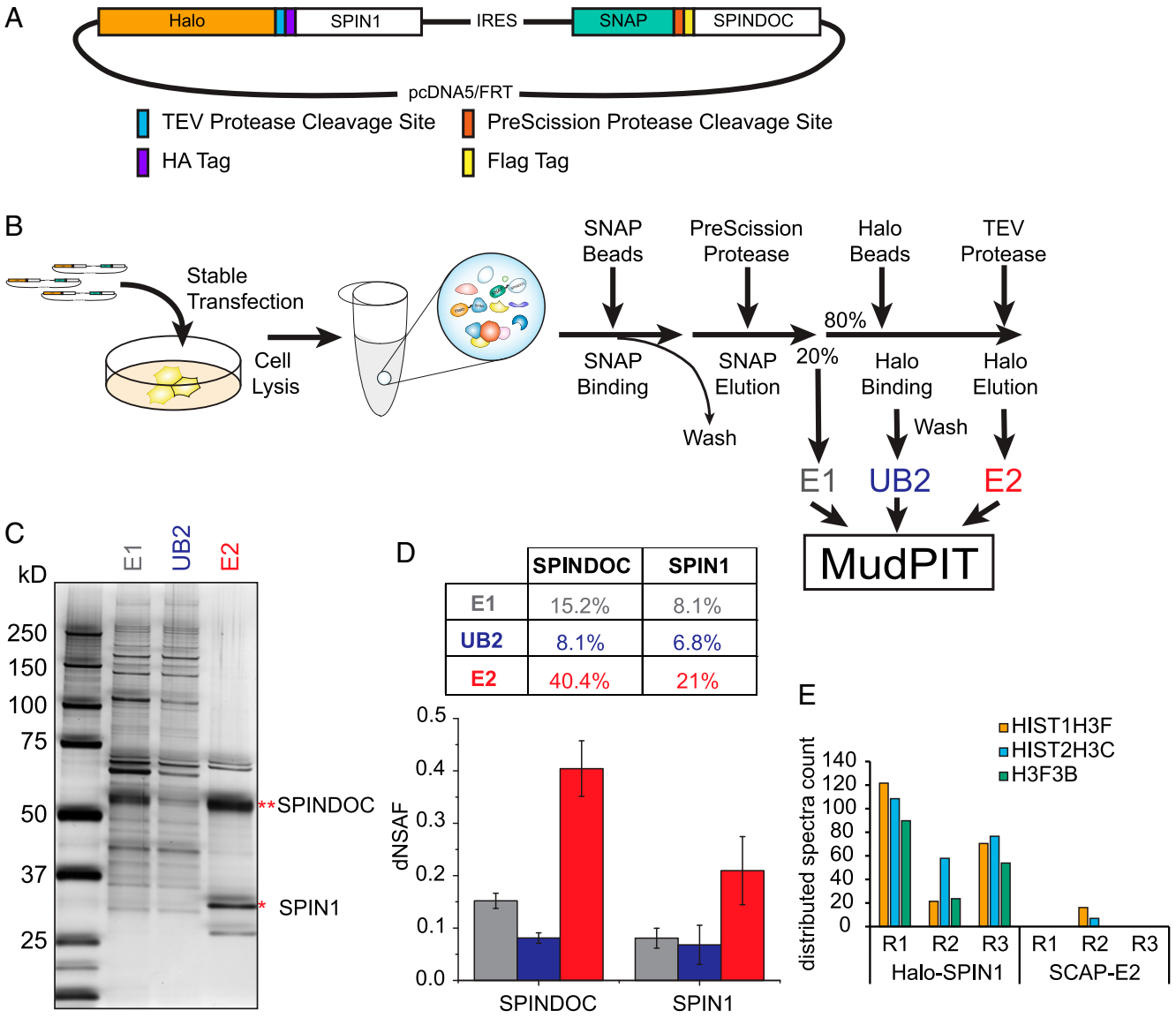

Fig. 1. SCAP of SPIN1 and SPINDOC complexes. (*A*) Schema of the expression vector designed for coexpression of HaloTag and SNAP-tag protein pairs, with the detailed vector map provided in *SI Appendix*, Fig. S2*A*. (*B*) Workflow of the SCAP quantitative proteomics method. E1 (the elution from SNAP purification), UB2 (the unbound proteins after binding to Halo bead resin), and E2 (the elution from the Halo beads) were separately analyzed by MudPIT, and dNSAF values were calculated for all identified proteins (Dataset S2). The abundance of SPIN1 and SPINDOC were calculated as dNSAF × 100%. (*C*) Silver-stained SDS/PAGE of the proteins eluted from the E1, UB2, and E2 fractions. (*D*) dNSAF plot of SPINDOC and SPIN1 in the E1, UB2, and E2 fractions. The value is the mean of three biological replicates; error bars represent SD. (*E*) Spectral counts measured for histones associated with SPIN1 in a single Halo purification and after SCAP.

SNAP-SPINDOC in live cells (Fig. 2). To confirm the interaction between SPIN1 and SPINDOC in vivo, we implemented two different imaging-based approaches: acceptor photobleaching Förster resonance energy transfer (AP-FRET) (27) and fluorescence cross-correlation spectroscopy (FCCS) (28). Both techniques can be applied in live cells, and both benefit from low concentrations of labeled proteins expressed in cells. These features reduce the possibility that SPIN1 capture of SPINDOC is due either to overexpression of the bait protein or to breakdown of subcellular segregation during cell lysis.

To apply AP-FRET on SPIN1 and SPINDOC, we coexpressed Halo-SPIN1 and SNAP-SPINDOC in HEK293/FRT cells. Then we labeled the HaloTag with TMRDirect ligand as an acceptor and the SNAP-tag with 505Star ligand as the donor. We photobleached the acceptor cell-by-cell with a 561-nm laser and measured the average donor intensity change of each bleached cell before and after bleaching of the acceptor (Fig. 2*A*). The FRET efficiency of each cell was then calculated from the increased donor intensity after photobleaching of the acceptor to the donor intensity after bleaching. Multiple cells were analyzed (*SI Appendix*, Fig. S3*B* and Dataset S3 *A* and *B*), and the average FRET efficiency was calculated (Fig. 2*B*). Protein pairs serving as positive and negative controls were also tested in parallel to determine the upper limit of FRET efficiency (*SI Appendix*, Fig. S3*A*). Halo-SPIN1 showed a significantly higher FRET efficiency with SNAP-SPINDOC than with SNAP itself (Fig. 2*B*), indicating a direct interaction between SPIN1 and SPINDOC in live cells.

Next, we used FCCS (28) to investigate whether Halo-SPIN1 and SNAP-SPINDOC codiffuse in live cells (*SI Appendix*, Fig. S3*C*).

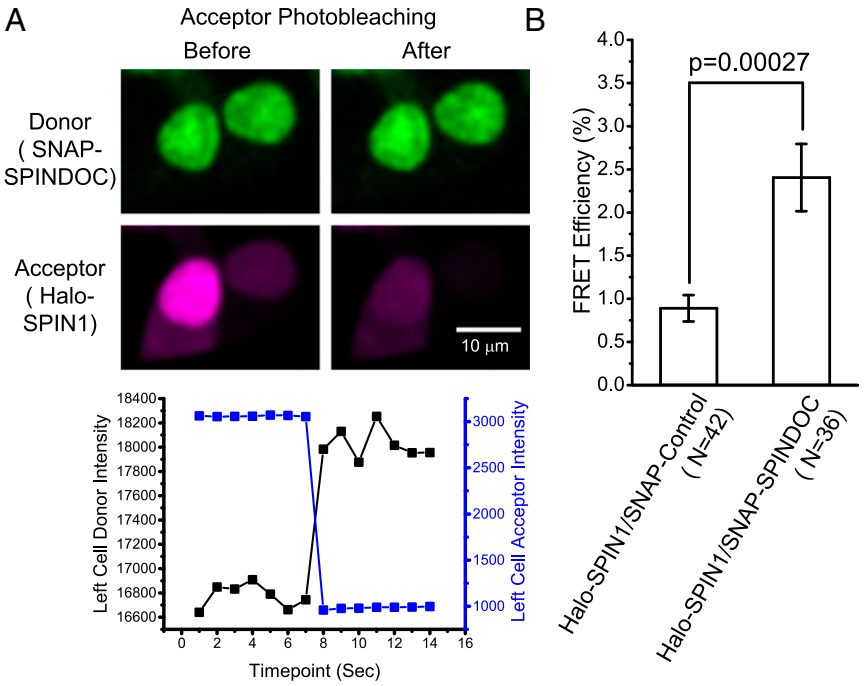

Fig. 2. SPIN1 and SPINDOC interaction in live cells. (*A*) Example image and intensity measurement of AP-FRET. Halo-SPIN1 was labeled with HaloTag TMRDirect, and SNAP-SPINDOC was stained with SNAP-Cell 505-Star ligand. (*B*) Averaged FRET efficiencies measured for Halo-SPIN1 and SNAP-SPINDOC in live HEK293FRT cells (Dataset S3 *A* and *B*). Error bars represent SE of means for the datapoints defined in *SI Appendix*, Fig. S3*B*, and *P* values were calculated using the two-tailed *t* test.

As with AP-FRET, we coexpressed SNAP-SPINDOC with Halo-SPIN1 or Halo-tag only in live cells. We then labeled the two tags with ligands conjugated to distinct fluorophores and measured both the auto self-correlation function of each species and the cross-correlation functions between HaloTag and SNAP-tag (*SI Appendix*, Fig. S3*D* and Dataset S3 *C* and *D*). From the curves, we observed a cross-correlation between Halo-SPIN1 and SNAP-SPINDOC, but not between Halo-Control and SNAP-SPINDOC (*SI Appendix*, Fig. S3*E*). The fraction of Halo-SPIN1 binding to SNAP-SPINDOC was then calculated using $G(\tau)$, the respective amplitudes of the autocorrelation and cross-correlation functions (*SI Appendix*, Fig. S3*F*), which revealed a significantly larger fraction of SPINDOC binding to SPIN1 than to the HaloTag by itself. These findings suggest that Halo-SPIN1 and SNAP-SPINDOC codiffuse and thus interact in live cells.

Implementing SCAP-XL to Derive a Structural Model of the SPIN1:SPINDOC Complex. Given the significant enrichment of the SPINDOC and SPIN1 proteins in the E2 fraction (Fig. 1*D*), we reasoned that this highly purified complex would be an excellent candidate for further structural analysis. Taking advantage of the availability of MS-cleavable cross-linking reagents and highly sensitive mass spectrometers (29, 30), we added a chemical XL step to further improve the SCAP pipeline (SCAP-XL). The MS-cleavable disuccinimidyl sulfoxide (DSSO) cross-linker (29) was added to the SCAP isolated proteins bound to Halo beads before TEV protease elution, resulting in a cross-linked E2 fraction (Fig. 3*A*). Analysis of the proteins on SDS/PAGE confirmed the presence of higher molecular weight cross-linked species (*SI Appendix*, Fig. S4*A*). We then analyzed these fractions in multiple replicates (Dataset S1*A*) on an Orbitrap Fusion Lumos Tribrid mass spectrometer, which identified cross-linked peptides using MS1, MS2, and MS3 information (*SI Appendix*, Fig. S4*B*). The resulting MS datasets were analyzed with the XlinkX search engine implemented through Proteome Discoverer (31) (Dataset S4 *A* and *B*). Although there were 140 proteins for which cross-links to SPIN1 or SPINDOC were detected, all but four of these proteins had only one or two cross-links to SPIN1 or SPINDOC (*SI Appendix*, Fig. S4*C* and Dataset S5*A*). Three proteins had three or four cross-links to the baits. In contrast, 21 cross-links were observed between SPIN1 and SPINDOC, suggesting a particularly close association between these two proteins. Consistent with the evidence from the FRET experiments (Fig. 2), these cross-linking results support a direct interaction between SPIN1 and SPINDOC.

To assess the distribution of both intermolecular cross-links (between SPINDOC and SPIN1) and intramolecular cross-links (within each protein), we visualized the cross-links using xiNET (32) (Fig. 3*B*). While many cross-linked sites were distributed throughout the three Tudor domains of SPIN1, suggesting an extended interaction interface, there were distinct regions of SPINDOC either with or without extensive cross-linking. As the N-terminal half of SPINDOC (1 to 180) was sparse in cross-linked sites, we predicted that this region would not be sufficient to bind to SPIN1 alone. In contrast, as the C-terminal half of SPINDOC (181 to 381) contained extensive cross-links to SPIN1, we predicted that this region would be sufficient to bind SPIN1. To test this, we performed mutational analysis on SPINDOC and found that SNAP-SPIN1 indeed failed to capture Halo-SPINDOC 1 to 180 but could capture Halo-SPINDOC 181 to 381 (*SI Appendix*, Fig. S4*D*). Consistent with these results, Devi et al. (20) also reported that SPINDOC lost interaction with SPIN1 on C-terminal deletion. Since cross-links to SPIN1 were not present in the approximately 30 C-terminal amino acids of SPINDOC, we further truncated the C-terminal mutant to a region from amino acids 181 to 350. Consistent with our cross-linking data, the results in *SI Appendix*, Fig. S4*D* suggest that SPINDOC 181 to 350 alone is sufficient to mediate the interaction with SPIN1.

The structure of SPIN1 had been previously solved by X-ray crystallography (35–37). To assess whether SPIN1 self-cross-links

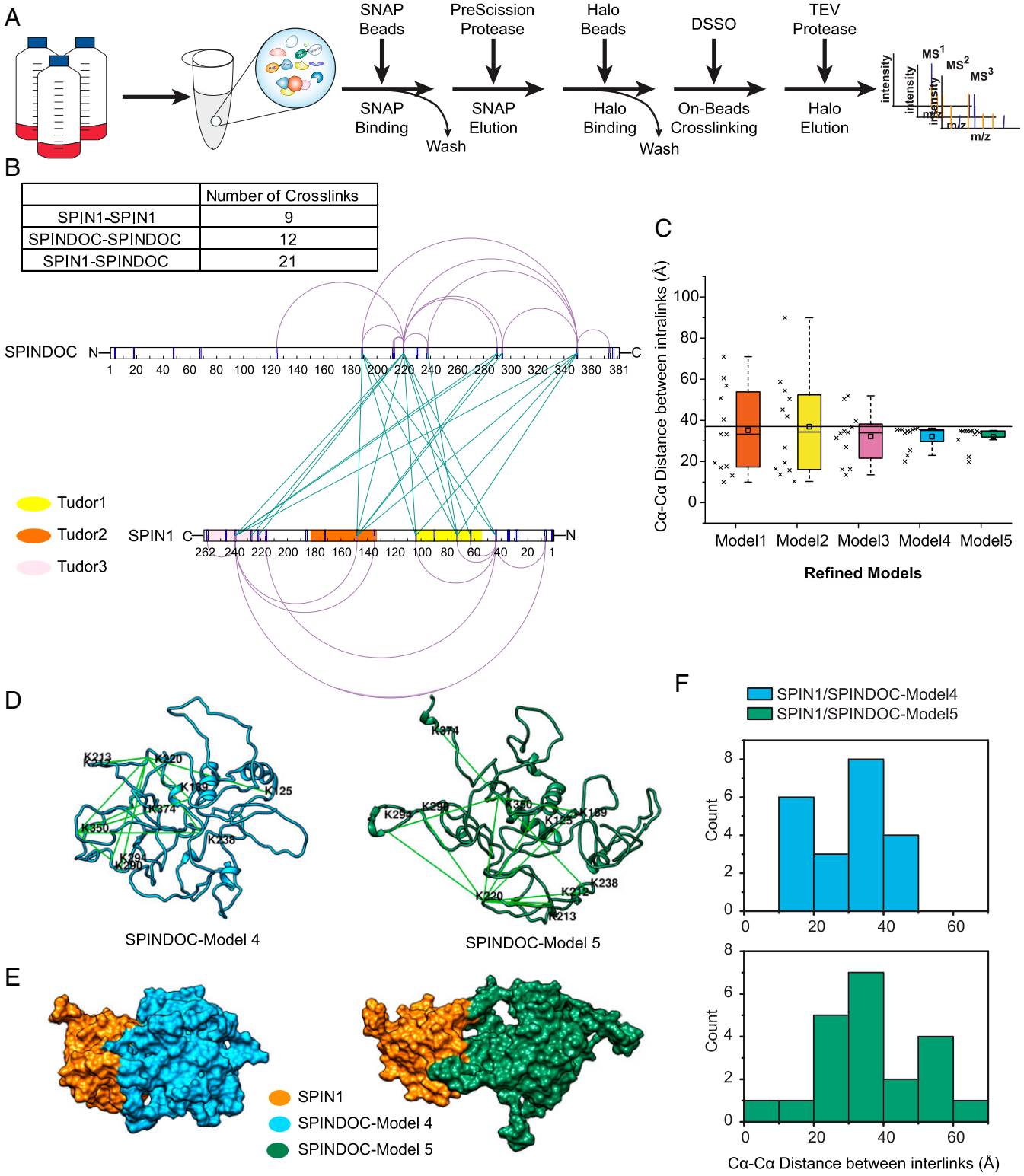

Fig. 3. (*A*) Workflow of the SCAP-XL method in which a DSSO cross-linking reaction was performed before the TEV protease elution while the purified proteins were still on Halo beads. (*B*) A two-dimensional visualization of SPIN1 and SPINDOC cross-links via XiNET (32). Intracross-links are shown in purple and intermolecular cross-links are shown in blue. (*Inset*) Total numbers of cross-links detected and identified in replicate SCAP-XL analyses (Dataset S4 *A* and *B*). (*C*) Cα–Cα distances between SPINDOC intracross-link sites in refined structural models predicted by I-TASSER (33). These models were predicted with the guidance of SPINDOC intracross-links. (*D*) SPINDOC refined structural models 4 and 5 defined by I-TASSER (33). Intracross-links are shown as green lines. (*E*) Docking models of SPIN1/SPINDOC-model 4 and SPIN1/SPINDOC-model 5 generated by HADDOCK (34). (*F*) Cα–Cα distances between SPIN1/SPINDOC intercross-linked sites for SPIN1/SPINDOC-model 4 and SPIN1/SPINDOC-model 5 complex models.

determined in solution were consistent with existing structural models, we mapped them to structures 4mzf (37), 4h75 (36), and 2ns2 (35) and found that all cross-linked pairs were within 30 Å, in agreement with the distance allowed by DSSO spacer arm length (Dataset S5*B*). However, the structure of its dominant interacting protein, SPINDOC, remains to be elucidated, and the structural nature of the SPINDOC:SPIN1 complex is not well understood.

To gain a better understanding of the SPINDOC:SPIN1 complex architecture, we used the I-Tasser structure prediction method to predict the three-dimensional structure of SPINDOC using its amino acid sequence (33, 38). We initially modeled SPINDOC structures using the amino acid sequence alone and obtained five preliminary models. When we mapped intracross-links onto the predicted structures, many of the Cα–Cα distances between pairs of cross-linked lysines exceeded the ∼35 Å limit suggested by the length of the DSSO cross-linker (*SI Appendix*, Fig. S4*E*), although for preliminary model 3, most cross-links did not exceed this distance limit. Therefore, to refine our model, we used the 12 intracross-links as distance restraints to determine five additional predicted structures (Fig. 3*C*). This approach generated two models for SPINDOC (models 4 and 5) for which all distances between the cross-linked lysine residues were less than the estimated 35 Å distance allowed by DSSO.

We selected these two models for further analysis to test their suitability for assembly with SPIN1 (Fig. 3 *D* and *E*). Although we obtained two predicted structures for SPINDOC that are consistent with the observed intracross-links resulting from individual molecules of SPINDOC, we cannot be certain that these observed cross-links indeed result from cross-linking between two lysines within a single SPINDOC molecule or, alternatively, between two copies of SPINDOC. Future development of methods aimed at resolving this question would enhance our ability to use cross-linking data to guide modeling of protein structures (39).

We next used the intermolecular cross-links between SPIN1 and SPINDOC to guide complex assembly. For SPIN1, we used 4mzf.pdb, which is the most recently published structure of SPIN1 binding a K4 trimethylated and R8 dimethylated H3 N′ peptide (37). For SPINDOC, as both model 4 and model 5 (refined models in Fig. 3*C*) satisfied our cross-linking data, we initially decided to assess the docking of both models to SPIN1. To do so, we first performed interaction space analysis using DisVis web portal to filter the SPIN1-SPINDOC cross-links (30, 40, 41). Cross-links that violate the accessible interaction space between SPIN1 and each SPINDOC model were excluded in docking of heterodimers. The filtering of cross-links was based on the *z*-score in DisVis output at a cutoff of 1 (Dataset S5*C*). A higher *z*-score indicates that the cross-link is more likely to be false. We excluded only two cross-links for model 4 and excluded six cross-links for model 5. This analysis suggested that refined SPINDOC model 4 forms a complex with SPIN1 that is better supported by the cross-linking evidence.

After filtering by DisVis analysis, the intercross-links were next used as unambiguous restraints for docking either SPINDOC refined model 4 or model 5 to SPIN1 4mzf. The docking was performed using the HADDOCK2.2 webserver (34, 42) with default parameters (Dataset S6*A*). For each docking, multiple clusters of SPIN1:SPINDOC dimer models were generated with a HADDOCK score (Dataset S6 *B* and *C*). A more negative HADDOCK score indicates that the cluster is more reliable. The best model in the top cluster was chosen (Fig. 3*E*). The Cα–Cα distances between pairs of cross-linked lysines were measured for each model (Fig. 3*F*). Consistent with the initial DisVis analysis, the model using SPINDOC refined model 4 docked to SPIN1 is more consistent with the structural limitations suggested by the cross-linking evidence.

For SPIN1 and SPINDOC refined model 4 dimer models, we further evaluated the four models in the top cluster using XL scores (Dataset S5*D*) following guidelines outlined by Orbán-Németh et al. (43). In brief, a higher XL score indicates that a model satisfies more cross-links with shorter Cα–Cα distances. Of the four models in cluster 1, the first SPIN1-SPINDOC dimer model had the highest XL score (Fig. 4 *A* and *B*; overlaid ensemble representations of all models are reported in *SI Appendix*, Fig. S5 *A* and *B*). However, even for this SPIN1-SPINDOC model, five cross-links were not satisfied. This suggested that a binary model was not the optimal solution supported by the interlink data. Consistent with this, a stoichiometry of two SPINDOC molecules to one SPIN1 molecule in the complex had previously been suggested by the quantitative proteomics analysis of the E2 fraction in the SCAP purification (Fig. 1 *C* and *D*).

Therefore, we used HADDOCK to assemble two SPINDOC molecules and one SPIN1 molecule to generate heterotrimer models. For the initial modeling (Dataset S6*D*), we used all intercross-links as restraints between SPIN1 and both copies of SPINDOC. We chose the first model in the top cluster as a preliminary trimer model. We distributed each restraint to the SPINDOC chain that gave the shortest distance when mapped to the preliminary trimer model (Dataset S5*E*), then used HADDOCK to generate refined trimer models using the updated restraints list. To select the best HADDOCK model, we then calculated XL scores for all best four models in the top HADDOCK clusters (Dataset S5*F*). The third model in cluster 1, which satisfies 18 of the 21 cross-links measured by SCAP-XL, was chosen as a representative structure for the trimeric complex based on XL scoring criteria (Fig. 4*C*); however, it should be considered in the context of the other top-scoring HADDOCK clusters of models. The top two clusters in our refined trimer model had very close HADDOCK scores (cluster 2, −205.6 ± 6.5; cluster 1, −203.4 ± −2.8). Thus, the top models in each of these clusters (*SI Appendix*, Fig. S5*D*) should be considered as potential modeling solutions. Datasets S5*F* and S6*E* provide a full analysis comparing all high-scoring models in the top clusters, and additional possible solutions for both dimer and trimer models shown in *SI Appendix*, Fig. S5.

Discussion

We designed SCAP by taking advantage of two multifunctional and orthogonal affinity tags, the HaloTag (22) and SNAP-tag (8) and, as a proof of principle, built a cell line stably expressing both Halo-SPIN1 and SNAP-SPINDOC for characterization in multiple experiments. The Tudor domain containing protein SPIN1 is a histone methylation reader and has been found to specifically bind H3K4me3-containing peptide with high affinity (25, 35–37). In addition, SPIN1 has been reported to promote cancer proliferation and progression (13, 26). The structure of SPIN1 has been solved by X-ray crystallography (35–37), and consequently, multiple researchers are pursuing the development of a SPIN1 inhibitor (14, 15, 44–47). C11orf84, recently renamed SPINDOC by Bae et al. (9), is a SPIN1-interacting protein that is less well understood, and the complex containing these two proteins remains poorly characterized.

In a multistep sequential affinity purification that we term SCAP, proteins associated with SNAP-SPINDOC were first isolated on SNAP beads, and then the population of such proteins also associating with Halo-SPIN1 was further enriched on Halo beads. An important feature of this approach is the use of distinct proteases for the Halo and SNAP purification steps, with PreScission Protease used to elute from the SNAP beads and TEV used to cleanly elute proteins from the Halo beads. This protocol resulted in a >2.5-fold increase in the concentration of each protein after elution from the second affinity step. Although our analysis of purified complexes might not necessarily reflect the complex stoichiometry in its native environment in vivo, the quantitative proteomic analysis of the enriched

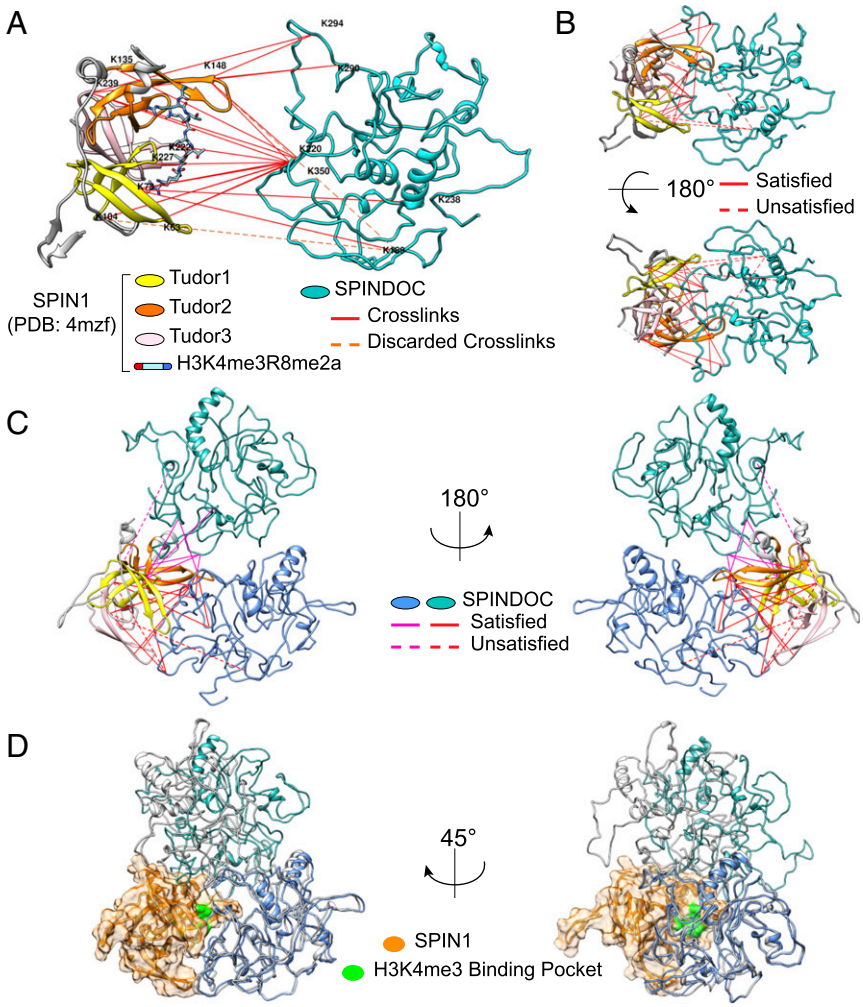

Fig. 4. Integrative structural modeling of the SPIN1:SPINDOC complex. (*A*) Visualization of the intermolecular cross-links between SPIN1 and SPINDOC-model 4. (*B*) A HADDOCK-generated docking model with one copy of SPIN1 and one copy of SPINDOC-model 4. (*C*) Docking model with one copy of SPIN1 and two copies of SPINDOC-model 4. Cross-links distributed to each copy of SPINDOC are in different colors. (*D*) Comparison of the refined SPIN1:SPINDOC heterotrimer model with the preliminary model. The SPIN1 subunits in the two trimer models were aligned. The SPINDOC subunits are shown in gray in the preliminary model.

complex suggested a 2:1 ratio of SPINDOC to SPIN1 molecules. The multifunctional features of the HaloTag (22) and the SNAP-tag (8) were also used to validate the interaction between SPIN1 and SPINDOC, with imaging approaches in live cells using the same expression constructs used for protein purification.

After optimization of the purification protocol for serial capture, the enriched population of Halo-SPIN1 and SNAP-SPINDOC complexes was the ideal candidate for further structural characterization using state-of-the-art cross-linking MS and computational approaches (29, 30). Therefore, we used SCAP-XL to purify DSSO–cross-linked protein complexes for analysis on an Orbitrap Fusion Lumos mass spectrometer, which has advanced capabilities for the study of collision-induced dissociation (CID)-cleavable cross-linked peptides (30). The culmination of this analysis was the identification of a reliable tridimensional model for SPINDOC, for which no structural information is available, followed by refinement of a structural model of a heterotrimer in which two SPINDOC molecules are docked on one SPIN1 molecule. In this study, we applied the SCAP approach to analyze the interactions between two proteins, SPIN1 and SPINDOC. In principle, the approach could be extended to prepare samples of larger protein complexes for analysis and structural modeling by tagging two components of a larger multisubunit complex. Our recent study modeling a substructure of the relatively abundant chromatin remodeler Sin3 suggests that this would be feasible (48). Although we were able to model a three-subunit substructure using a single-step Halo affinity purification in this case, the improvements in sample preparation offered by the SCAP approach might make similar studies with less abundant larger complexes more feasible.

The SPIN1 structure that we used to assemble the complex contained a histone H3-K4me3-R9me2a peptide (37). Intriguingly, the SPINDOC:SPIN1 interaction surface identified by cross-linking overlaps with the binding pocket for H3-K4me3-R8me2a (Fig. 4*A*). In the heterotrimer model, the pocket was also blocked by one of the two SPINDOC molecules (Fig. 4*D*). Comparing the final representative model with the preliminary model (shown in gray in Fig. 4*D*), we noticed that the major refinement appeared to be the positioning of the top molecule of SPINDOC (shown in green), while the second molecule of SPINDOC blocking the binding pocket (in blue) remained close to its position in the preliminary model. Furthermore, even when considering alternate structural solutions, such as the two trimeric models deemed best based on different scoring criteria (*SI*

Appendix, Fig. S5*D*), the positioning of the SPINDOC chain shown in the lower right (in blue in *SI Appendix*, Fig. S5*C*), which blocks the H3 tail-binding pocket, is conserved between all possible solutions generated by the docking calculations. This result suggests that the binding of SPIN1 to SPINDOC could disrupt and/or compete with its binding to modified histone H3, which is consistent with findings reported by Bae et al. (9). Finally, comparing the proteins recovered by proteomics analyses of the SCAP E2 elution with a Halo-SPIN1 purification alone revealed significantly fewer histone H3 interactions in the SCAP E2 elution (Fig. 1*E*), also supporting the possibility that SPINDOC disrupts SPIN1 interaction with histone methylation sites, which warrants further study.

The SCAP and SCAP-XL pipelines described herein were designed to be generic approaches that can realistically be applied to any pair of interacting proteins. We devised a plasmid-based system for making stable cell lines in HEK293 cells, which allowed us to generate enough starting material for the SCAP-XL pipeline, since cross-linked peptides can be of low abundance in a sample. This system should allow for medium throughput analysis of predicted direct protein interactions that are part of protein complexes of varying sizes (3), and these predicted directly interacting proteins are likely good candidates for incorporation into the SCAP and SCAP-XL pipelines for ex vivo complex characterization, in vivo interaction validation, and the building of structural models of protein complexes.

The larger concept of ProteoCellomics is defined by coupling of quantitative proteomic analysis of affinity purified complexes on the one hand and quantitative spectroscopy techniques to image these complexes in live cells on the other hand, (Fig. 5). Therefore, SCAP enables the practical application of ProteoCellomics in which quantitative proteomics, an ex vivo approach, and quantitative microscopy, an in vivo approach, are integrated to gain molecular insight into protein–protein interactions. Finally, when coupling the SCAP-XL pipeline to computational modeling approaches, an integrative structural model of a protein complex can be generated, further advancing the understanding of poorly characterized protein–protein interactions and protein complexes.

Materials and Methods

Critical Reagents. Magne HaloTag beads, sequencing-grade modified trypsin, rLys-C, and HaloTag ligands were purchased from Promega. All restriction endonucleases, SNAP-cell ligands, and SNAP-capture magnetic beads were purchased from New England BioLabs. PreScission protease was purchased from GE Healthcare Life Sciences. AcTEV protease and DSSO were purchased from Thermo Fisher Scientific. Salt active nuclease was purchased from ArcticZymes Technologies.

Plasmids and Cell Lines. The sequences of SPIN1 ORF was obtained from Kazusa Genome Technology, and the sequence of SPINDOC ORF was obtained from RT-PCR using mRNA extracted from 293FRT cells. Vectors containing SNAP-tag with different protease recognition sites were modified from pSNAPf (New England BioLabs). SPIN1 coding sequence were subcloned into pSNAPf-F/T/P/X vector for transient expression of SNAP-F/T/P/X-SPIN1. pcDNA5FRT vector was purchased from Invitrogen. pcDNA5FRT-Halo was generated by inserting a Halo-TEV sequence (obtained from the pFN21A vector purchased from Promega) downstream of the CMV promoter. pcDNA5FRT-SNAP was generated by inserting the SNAP-PP sequence (obtained from the pSNAP-P vector) downstream of the CMV promoter. Positive control

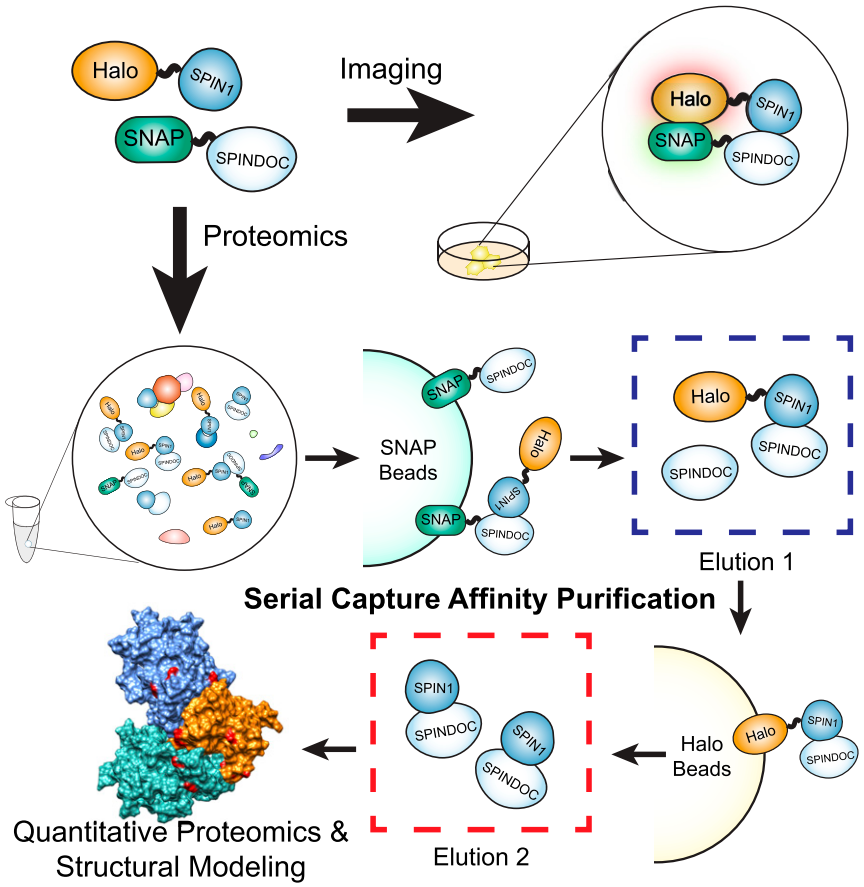

Fig. 5. General concept of ProteoCellomics. In ProteoCellomics, protein interactions are characterized ex vivo using APMS and in vivo using live cell imaging. This is accomplished by tagging two separate proteins with distinct multifunctional affinity tags like the HaloTag and the SNAP-tag. The interaction of two proteins can then be studied in a live cell using imaging techniques. Enriched protein complexes can then be isolated using SCAP and analyzed using mass spectrometry techniques.

vector pcDNA5FRT-Halo-NLS-SNAP was generated by linking Halo and SNAP tag sequences by an nuclear localization signal sequence and inserted downstream of the CMV promoter of the pcDNA5FRT vector. The dual expression vector was modified from the pcDNA5FRT-Halo vector by inserting an IRES sequence downstream of Halo-TEV-HA sequence, followed by the SNAP-PP--Flag sequence. Then SPIN1 or SPINDOC sequences were subcloned into each vector for expression of corresponding Halo- or SNAP-tagged protein. All oligos used in cloning were synthesized by Integrated DNA Technologies.

293FRT cells were purchased from Invitrogen and maintained in DMEM with GlutaMAX and 10% FBS. All stable expression cell lines were maintained in DMEM with GlutaMAX, 10% FBS, and 100 μg/mL hygromycin B. For transient expression, FuGene6 (Promega) was used to transfect the expression vector into 293FRT cells. Stable expression cell lines were generated using the Flp-In System (Invitrogen) according to the manufacturer's instruction. All the aforementioned vectors contain a CMV promoter to drive expression. All stable expression cell lines used in this study were generated from 293FRT cells. All cells were maintained at 37 °C in a 5% humidified incubator.

SCAP and SCAP-XL. For regular SCAP, both baits were stably expressed in 293FRT cells. Cells were collected and lysed with high salt lysis buffer. The lysate was centrifuged, and supernatant was incubated with SNAP-Capture magnetic beads (New England BioLabs) at 4 °C for 2 h. The beads were then washed three times with high salt wash buffer, followed by two washes in wash buffer. Bound proteins were eluted with elution buffer (containing PreScission protease). Twenty percent of the eluate was aliquoted as E1, and the remaining 80% was subjected to further Halo purification. Eluate thus obtained was then incubated with Magne HaloTag Beads (Promega) at 4 °C for 2 h. Unbound supernatant was collected as UB2. Beads were then washed five times with wash buffer. Bound proteins were eluted with elution buffer (containing TEV protease), and the eluate was collected as E2. For all three replicates, E1, UB2, and E2 were subjected to MudPIT analysis. For SCAP-XL, cells were collected from three roller bottles. The cell pellet was homogenized with a dounce tissue grinder in high salt lysis buffer and lysed at 4 °C. Lysate was centrifuged, and supernatant was subjected to SNAP purification. Then the eluate was bound to Magne HaloTag Beads (Promega). After washes, bound proteins were cross-linked on beads with 5 mM DSSO (Thermo Fisher Scientific) at 4 °C for 1 h. The cross-linking reaction was quenched by 50 mM Tris·HCl, and cross-linked proteins were eluted by TEV protease.

AP-FRET. AP-FRET was performed as described by Weems et al. (49). In detail, data were acquired with a PerkinElmer Life Sciences UltraVIEW VoX spinning-disk microscope controlled by Volocity software. The microscope was equipped with a Yokogawa CSU-X1 spinning-disk scanner, an ORCA-R2 camera (Hamamatsu, C10600-10B), and an EM-CCD camera (Hamamatsu, C9100-23B), and bleaching studies were conducted with the included PhotoKinesis accessory. The base of the microscope was a Carl Zeiss Axiovert 200 M. A main dichroic passing 405-, 488-, 561-, and 640-nm laser line was used. HaloTag Ligands TMRDirect labeled proteins were excited by a 561-nm laser, and their emission was collected through a dual bandpass 445 (W60)-615 (W70) filter. SNAP-Cell 505-Star was excited by a 488-nm laser, and emission was collected through a 525 (W50) bandpass filter.

To collect AP-FRET data, time-lapse movies were recorded to collect at least 10 timepoints before and after acceptor photobleaching. The movies were recorded at a speed of one image per second. For SPIN1/SPINDOC experiments, images were recorded using the EM-CCD with an objective of 40× (water, NA = 1.2). For each cell, acceptors were bleached by a 561-nm laser at 100% laser power for 10 cycles. The donor intensity before (I_{Before}) and after (I_{After}) were averaged separately over time. FRET efficiency, E, was calculated as $E = 1 - (I_{Before}/I_{After})$. E values were calculated in batches using in-house ImageJ plugin accpb FRET analysis jru v1. Control images verified that the acceptor was bleached effectively with the number of iterations.

LC-MS3 Data Acquisition and Analysis of SCAP-XL Samples. Cross-linked peptides were analyzed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher Scientific) coupled to a Dionex UltiMate 3000 RSCLnano System. Peptides were loaded on an Acclaim PepMap 100 C18 0.3-mm-i.d., 5-mm-long trap cartridge (Thermo Fisher Scientific) with a loading pump at 2 μL/min via an autosampler. An analytical column (50 μm i.d., 150 mm long) was packed in-house with ReproSil-Pur C18-AQ 1.9-μm resin (Dr. Masch GmbH). The organic solvent solutions were water/acetonitrile/formic acid at 95:5:0.1 (vol/vol/v) for buffer A (pH 2.6), and at 20:80:0.1 (vol/vol/v) for buffer B. When cross-linked peptides were analyzed, the chromatography gradient was a 20-min column equilibration step in 2% B; a 10-min ramp to reach 10% B; 120 min from 10% to 40% B; 5 min to reach 95% B; a 14-min wash at 95% B; 1 min to 2% B; followed by a 10-min column reequilibration step at 2% B. The nano pump flow rate was 120 nL/min. An MS3 method was developed specifically for the analysis of DSSO cross-linked peptides. Full MS scans were performed at 60,000-*m/z* resolution in the orbitrap with a 1.6-*m/z* isolation window, and the scan range was 375 to 1,500 *m/z*. The top three peptides with charge state 4 to 8 were selected for MS2 fragmentation with 20% CID energy. MS2 scans were detected in the Orbitrap with 30,000-*m/z* resolution and a dynamic exclusion time of 40 s. Among MS2 fragments, if two peptides with an exact mass difference of 31.9720 with 20 ppm mass tolerance were present, both were selected for MS3 fragmentation at CID energy of 35%. MS3 scans were performed in the ion trap at rapid scan with an isolation window of 3 *m/z* and a maximum ion injection time of 200 ms. Each MS2 scan was followed by maximum of four MS3 scans.

Data analysis of DSSO cross-linked peptides was performed with Proteome Discoverer 2.2 (Thermo Fisher Scientific), with the XlinkX add-on cross-linking node used in peptide identification and cross-linked peptide searches. The following settings were used: precursor ion mass tolerance, 10 ppm; fragment ion mass tolerance, 0.6 Da; fixed modification, Cys carbamidomethylation; variable modification, Met oxidation, Lys DSSO amidated, and Lys DSSO hydrolyzed; maximum number of dynamic modifications, 3. The protein false discovery rate was set at 0.01. Cross-linking data visualization and structural modeling are described in detail in *SI Appendix*.

Data Availability. All in-house written Fiji or ImageJ plugins can be downloaded at: https://research.stowers.org/imagejplugins/zipped_plugins.html. UCSF Chimera was downloaded from: http://www.rbvi.ucsf.edu/chimera. The MS dataset may be obtained from the MassIVE database with accession nos. MSV000084678 (50), MSV000084679 (51), MSV000084713 (52), and MSV000084719 (53). The heterotrimer model of SPIN1:SPINDOC with the highest XL score may be accessed at Protein Data Bank (PDB)-Dev (54) using accession no. PDBDEV_00000061 (55). Original data underlying this manuscript may be accessed from the Stowers Original Data Repository at https://www.stowers.org/research/publications/libpb-1496 (56).

ACKNOWLEDGMENTS. We thank Dr. Brinda Vallat for her assistance in depositing our model to PDB-Dev. The research reported in this publication was supported by the Stowers Institute for Medical Research and the National Institute of General Medical Sciences of the NIH (R35 GM118068, to J.L.W. and R01 GM112639, to M.P.W.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

1. T. Li *et al*., A scored human protein-protein interaction network to catalyze genomic interpretation. *Nat. Methods* **14**, 61–64 (2017).
2. M. Kotlyar, C. Pastrello, Z. Malik, I. Jurisica, IID 2018 update: Context-specific physical protein-protein interactions in human, model organisms and domesticated species. *Nucleic Acids Res.* **47**, D581–D589 (2019).
3. M. E. Sardiu *et al*., Topological scoring of protein interaction networks. *Nat. Commun.* **10**, 1118 (2019).
4. D. Mellacheruvu *et al*., The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).
5. G. Rigaut *et al*., A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* **17**, 1030–1032 (1999).
6. A. C. Gavin *et al*., Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141–147 (2002).
7. G. V. Los, K. Wood, The HaloTag: A novel technology for cell imaging and protein analysis. *Methods Mol. Biol.* **356**, 195–208 (2007).
8. A. Keppler *et al*., A general method for the covalent labeling of fusion proteins with small molecules in vivo. *Nat. Biotechnol.* **21**, 86–89 (2003).
9. N. Bae *et al*., A transcriptional coregulator, SPIN·DOC, attenuates the coactivator activity of Spindlin1. *J. Biol. Chem.* **292**, 20808–20817 (2017).
10. Y. Gao *et al*., Spindlin1, a novel nuclear protein with a role in the transformation of NIH3T3 cells. *Biochem. Biophys. Res. Commun.* **335**, 343–350 (2005).
11. T. G. Chew *et al*., A tudor domain protein SPINDLIN1 interacts with the mRNA-binding protein SERBP1 and is involved in mouse oocyte meiotic resumption. *PLoS One* **8**, e69764 (2013).
12. A. Ducroux *et al*., The Tudor domain protein Spindlin1 is involved in intrinsic antiviral defense against incoming hepatitis B Virus and herpes simplex virus type 1. *PLoS Pathog.* **10**, e1004343 (2014).
13. H. Franz *et al*., The histone code reader SPIN1 controls RET signaling in liposarcoma. *Oncotarget* **6**, 4773–4789 (2015).
14. N. Bae *et al*., Developing spindlin1 small-molecule inhibitors by using protein microarrays. *Nat. Chem. Biol.* **13**, 750–756 (2017).
15. R. Drago-Ferrante *et al*., Suppressive role exerted by microRNA-29b-1-5p in triple negative breast cancer through SPIN1 regulation. *Oncotarget* **8**, 28939–28958 (2017).

16. X. Chen, Y. W. Wang, P. Gao, SPIN1, negatively regulated by miR-148/152, enhances Adriamycin resistance via upregulating drug metabolizing enzymes and transporter in breast cancer. *J. Exp. Clin. Cancer Res.* **37**, 100 (2018).
17. Z. Fang *et al*., SPIN1 promotes tumorigenesis by blocking the uL18 (universal large ribosomal subunit protein 18)-MDM2-p53 pathway in human cancer. *eLife* **7**, e31275 (2018).
18. W. Chen *et al*., LINC00473/miR-374a-5p regulates esophageal squamous cell carcinoma via targeting SPIN1 to weaken the effect of radiotherapy. *J. Cell. Biochem.* **120**, 14562–14572 (2019).
19. J. W. Choi *et al*., Spindlin1 alters the metaphase to anaphase transition in meiosis I through regulation of BUB3 expression in porcine oocytes. *J. Cell. Physiol.* **234**, 8963–8974 (2019).
20. M. S. Devi *et al*., Spindlin docking protein (SPIN.DOC) interaction with SPIN1 (a histone code reader) regulates Wnt signaling. *Biochem. Biophys. Res. Commun.* **511**, 498–503 (2019).
21. V. Fagan *et al*., A chemical probe for tudor domain protein Spindlin1 to investigate chromatin function. *J. Med. Chem.* **62**, 9008–9025 (2019).
22. M. Urh, D. Hartzell, J. Mendez, D. H. Klaubert, K. Wood, Methods for detection of protein-protein and protein-DNA interactions using HaloTag. *Methods Mol. Biol.* **421**, 191–209 (2008).
23. D. L. Daniels *et al*., Examining the complexity of human RNA polymerase complexes using HaloTag technology coupled to label free quantitative proteomics. *J. Proteome Res.* **11**, 564–575 (2012).
24. C. A. S. Banks *et al*., A structured workflow for mapping human Sin3 histone deacetylase complex interactions using halo-MudPIT affinity-purification mass spectrometry. *Mol. Cell. Proteomics* **17**, 1432–1447 (2018).
25. W. Wang *et al*., Nucleolar protein spindlin1 recognizes H3K4 methylation and stimulates the expression of rRNA genes. *EMBO Rep.* **12**, 1160–1166 (2011).
26. J. X. Wang *et al*., SPINDLIN1 promotes cancer cell proliferation through activation of WNT/TCF-4 signaling. *Mol. Cancer Res.* **10**, 326–335 (2012).
27. E. A. Jares-Erijman, T. M. Jovin, FRET imaging. *Nat. Biotechnol.* **21**, 1387–1395 (2003).
28. K. Bacia, S. A. Kim, P. Schwille, Fluorescence cross-correlation spectroscopy in living cells. *Nat. Methods* **3**, 83–89 (2006).
29. A. Kao *et al*., Development of a novel crosslinking strategy for fast and accurate identification of crosslinked peptides of protein complexes. *Mol. Cell. Proteomics* **10**, M110 002212 (2011).
30. O. Klykov *et al*., Efficient and robust proteome-wide approaches for cross-linking mass spectrometry. *Nat. Protoc.* **13**, 2964–2990 (2018).
31. F. Liu, P. Lössl, R. Scheltema, R. Viner, A. J. R. Heck, Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* **8**, 15473 (2017).
32. C. W. Combe, L. Fischer, J. Rappsilber, xiNET: Crosslink network maps with residue resolution. *Mol. Cell. Proteomics* **14**, 1137–1147 (2015).
33. J. Yang, Y. Zhang, I-TASSER server: New development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).
34. G. C. P. van Zundert *et al*., The HADDOCK2.2 web server: User-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.* **428**, 720–725 (2016).
35. Q. Zhao *et al*., Structure of human spindlin1. Tandem tudor-like domains for cell cycle regulation. *J. Biol. Chem.* **282**, 647–656 (2007).
36. N. Yang *et al*., Distinct mode of methylated lysine-4 of histone H3 recognition by tandem tudor-like domains of Spindlin1. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17954–17959 (2012).
37. X. Su *et al*., Molecular basis underlying histone H3 lysine-arginine methylation pattern readout by Spin/Ssty repeats of Spindlin1. *Genes Dev.* **28**, 622–636 (2014).
38. Y. Zhang, I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 40 (2008).
39. A. Leitner, M. Faini, F. Stengel, R. Aebersold, Crosslinking and mass spectrometry: An integrated technology to understand the structure and function of molecular machines. *Trends Biochem. Sci.* **41**, 20–32 (2016).
40. G. C. van Zundert, A. M. Bonvin, DisVis: Quantifying and visualizing accessible interaction space of distance-restrained biomolecular complexes. *Bioinformatics* **31**, 3222–3224 (2015).
41. G. C. van Zundert *et al*., The DisVis and PowerFit web servers: Explorative and integrative modeling of biomolecular complexes. *J. Mol. Biol.* **429**, 399–407 (2017).
42. T. A. Wassenaar *et al*., WeNMR: Structural Biology on the grid. *J. Grid Comput.* **10**, 743–767 (2012).
43. Z. Orbán-Németh *et al*., Structural prediction of protein models using distance restraints derived from cross-linking mass spectrometry data. *Nat. Protoc.* **13**, 478–494 (2018).
44. X. Chen *et al*., Suppression of SPIN1-mediated PI3K-Akt pathway by miR-489 increases chemosensitivity in breast cancer. *J. Pathol.* **239**, 459–472 (2016).
45. D. Robaa *et al*., Identification and structure-activity relationship studies of small-molecule inhibitors of the methyllysine reader protein Spindlin1. *ChemMedChem* **11**, 2327–2338 (2016).
46. T. Wagner *et al*., Identification of a small-molecule ligand of the epigenetic reader protein Spindlin1 via a versatile screening platform. *Nucleic Acids Res.* **44**, e88 (2016).
47. Y. Xiong *et al*., Discovery of a potent and selective fragment-like inhibitor of methyllysine reader protein spindlin 1 (SPIN1). *J. Med. Chem.* **62**, 8996–9007 (2019).
48. C. A. S. Banks *et al*., Integrative modeling of a Sin3/HDAC complex sub-structure. *Cell Rep.* **31**, 107516 (2020).
49. J. C. Weems *et al*., Assembly of the elongin A ubiquitin ligase is regulated by genotoxic and other stresses. *J. Biol. Chem.* **290**, 15030–15041 (2015).
50. X. Liu *et al*, Triplicate MudPIT analyses of the proteins co-purified by Halo purifications of Halo-SPINDOC/C110rf84 and Halo-SNAP controls. MassIVE. https://doi.org/doi:10.25345/C5NX0Q. Deposited 11 December 2019.
51. X. Liu *et al*, Triplicate MudPIT analyses of the proteins co-purified by Halo purifications of 293FRT clones #9 and #11 stably expressing Halo-SPIN1 and Halo controls. MassIVE. https://doi.org/doi:10.25345/C5J676. Deposited 11 December 2019.
52. X. Liu *et al*, Triplicate MudPIT analyses of the proteins co-purified in the E1, UB2, and E2 fractions during SCAP purifications of Halo-SPIN1:SNAP-SPINDOC. MassIVE. https://doi.org/doi:10.25345/C51097. Deposited 17 December 2019.
53. X. Liu *et al*, Replicate LC/XLMS analyses of the peptides derived from DSSO crosslinking of the E2 fraction from SCAP purifications of Halo-SPIN1:SNAP-SPINDOC. MassIVE. https://doi.org/doi:10.25345/C57H64. Deposited 18 December 2019.
54. B. Vallat, B. Webb, J. D. Westbrook, A. Sali, H. M. Berman, Development of a prototype system for archiving integrative/hybrid structure models of biological macromolecules. *Structure* **26**, 894–904.e892 (2018).
55. X. Liu *et al*, Heterotrimer model with the highest cross-link score resulting from the HADDOCK2.2 docking of 1x SPIN1 and 2x SPINDOC chains with DSSO derived distance constraints. PDB-Dev. https://pdb-dev.wwpdb.org/entry.html?PDBDEV_00000061. Deposited 9 September 2020.
56. X. Liu *et al*, Original data for AP-MS and SCAP-MS Analyses of the SPIN1:SPINDOC complex by liquid chromatography mass spectrometry, SDS-PAGE, Western Blotting, FCCS, FRET, DisVis, I-Tasser, and HADDOCK. SIMR-ODR. ftp://odr.stowers.org/LIBPB-1496. Deposited 10 November 2020.