A Structured Workflow for Mapping Human Sin3 Histone Deacetylase Complex Interactions Using Halo-MudPIT Affinity-Purification Mass Spectrometry*

Charles A. S. Banks‡, Janet L. Thornton‡, Cassandra G. Eubanks‡, Mark K. Adams‡, Sayem Miah‡, Gina Boanca‡, Xingyu Liu‡, Maria L. Katt‡, Tari J. Parmely‡, Laurence Florens‡, and Michael P. Washburnদ

Although a variety of affinity purification mass spectrometry (AP-MS) strategies have been used to investigate complex interactions, many of these are susceptible to artifacts because of substantial overexpression of the exogenously expressed bait protein. Here we present a logical and systematic workflow that uses the multifunctional Halo tag to assess the correct localization and behavior of tagged subunits of the Sin3 histone deacetylase complex prior to further AP-MS analysis. Using this workflow, we modified our tagging/expression strategy with 21.7% of the tagged bait proteins that we constructed, allowing us to quickly develop validated reagents. Specifically, we apply the workflow to map interactions between stably expressed versions of the Sin3 subunits SUDS3, SAP30, or SAP30L and other cellular proteins. Here we show that the SAP30 and SAP30L paralogues strongly associate with the core Sin3 complex, but SAP30L has unique associations with the proteasome and the myelin sheath. Next, we demonstrate an advancement of the complex NSAF (cNSAF) approach, in which normalization to the scaffold protein SIN3A accounts for variations in the proportion of each bait capturing Sin3 complexes and allows a comparison among different baits capturing the same protein complex. This analysis reveals that although the Sin3 subunit SUDS3 appears to be used in both SIN3A and SIN3B based complexes, the SAP30 subunit is not used in SIN3B based complexes. Intriguingly, we do not detect the Sin3 subunits SAP18 and SAP25 among the 128 high-confidence interactions identified, suggesting that these subunits may not be common to all versions of the Sin3 complex in human cells. This workflow provides the framework for building validated reagents to assemble quantitative interaction networks for chromatin remodeling complexes and provides novel insights into focused protein interaction networks. Molecular & Cellular Proteomics 17: 1432–1447, 2018. DOI: 10.1074/mcp.TIR118.000661.
ent subsets of centrosomal proteins (3). In another study, Banks et al. used a series of truncation mutants of the TNIP2 protein to identify two distinct regions of TNIP2 that interact either with NFκB1 or with the ESCRT1 complex (4). In a third study, Sardi et al. investigated the effect of the cancer drug SAHA on interactions centered around subunits of the Sin3 co-repressor complex (5). Such studies depend on efficient experimental workflows for developing and testing reagents that can be used to identify legitimate protein-protein interactions and reduce the likelihood of false positive results.

Here we present a medium-throughput strategy applied to investigating interactions among components of the Sin3 complex, an important regulator of global gene expression (6). Clarifying the precise role of Sin3 components in controlling critical cellular processes is important as abnormal Sin3 function is associated with carcinogenesis (7). Notably, Sin3A, the gene encoding the key scaffolding protein around which Sin3 complexes are built, was recently identified as one of the top 127 significantly mutated genes across common cancer types (8). We first describe how a structured decision-making process can be used to screen recombinant complex subunits and verify correct bait protein localization, before exploiting the validated reagents in comprehensive AP-MS analyses. We have used the workflow to develop 23 cell lines constitutively expressing modest levels of Halo-tagged Sin3 components. Here we present AP-MS analyses based on three of these cell lines, Halo-SUDS3, Halo-SAP30 and Halo-SAP30L, and identify 128 high-confidence bait-prey interactions. These interactions include 17 subunits previously identified as Sin3 subunits, several Sin3 associated transcription factors, and putative Sin3 associated proteins BAHCC1 and TNRC18. Surprisingly, we did not find Sin3 subunits SAP18 and SAP25 in our purifications, suggesting that these components are not tightly integrated into all versions of the Sin3 complex. Finally, adjusting the Halo-MudPIT protocol can result in notable changes to the interactions identified, highlighting the importance of interpreting data mindful of factors such as whether nucleases have been included in buffers. To summarize, workflows for developing protein interaction networks that address the suitability of the reagents and purification procedures are important for generating refined models of protein complex organization and function.

1 The abbreviations used are: AP-MS, affinity purification mass spectrometry; CMV, cytomegalovirus; cNSAF, complex normalized spectral abundance factor; DNA/cDNA, deoxyribonucleic acid/complementary DNA; dNSAF, distributed normalized spectral abundance factor; FDR, false discovery rate; FPKM, fragments per kilobase of transcript per million fragments mapped; GO, gene ontology; HEK293T, human embryonic kidney 293 cells expressing the SV40 T antigen; MudPIT, multidimensional protein identification technology; SAF, spectral abundance factor; SAN, salt active nuclease; TEV, tobacco etch virus; Tris, Tris(hydroxymethyl)aminomethane.

**EXPERIMENTAL PROCEDURES**

**Materials—** Rabbit anti-HaloTag® polyclonal antibody (G9281) and Magne® HaloTag® affinity beads were purchased from Promega (Madison, WI). Mouse anti-α-tubulin (T9026) mouse anti-FLAG (M2) (F3165) monoclonal antibodies, and PreScission Protease (GE27-0843-01) were obtained from Sigma-Aldrich (St. Louis, MO). Rabbit anti-SAP30 (ab125187), rabbit anti-SUDS3 (ab184556), and rabbit anti-SIN3A (ab3479) polyclonal antibodies were from Abcam (Cambridge, United Kingdom). Mouse monoclonal anti-SIN3B (sc13145) was from Santa Cruz Biotechnology (Dallas, TX). IRDye® 800CW labeled goat anti-Rabbit (926–3211) and IRDye® 680LT labeled goat anti-Mouse (926–68020) secondary antibodies were from LI-COR Biosciences (Lincoln, NE). SNAP-Capture magnetic beads (S9145S) were from NEB (Ipswich, MA). Salt Active Nuclease (SAN) was from ArcticZymes (Tromso, Norway).

**Cloning ORFs into Plasmid Vectors—** cDNA was prepared from human placental total RNA (Clontech, Palo Alto, CA) using the iScript RT kit (Bio-Rad, Hercules, CA). The coding sequences of SUDS3 (NP_071936), SAP30L (NP_078908) and SAP18 (NP_005861) were amplified from cDNA using the PCR primers listed in Supplementary Data and cloned into either pFN21A, pCF14A, or CMVd2 pcDNAs/FRT PacI PmeI (9) (supplementary Data). Halo-SIN3A in pFN21A was generated using site directed mutagenesis to modify clone #FHC11647 (Promega) to match the SIN3A NCBI sequence NP_001138830, and to insert a stop codon at the 3’ end of the ORF. Halo-SIN3B in pFN21A was generated by modifying clone #FHC01991 (Promega) to match the SIN3B NCBI sequence NP_056075, and to insert a stop codon at the 3’ end of the ORF. Halo-SAP25 was made by cloning a synthetic sequence (supplementary Data) coding for the SAP25 protein (UniProtKB - Q8TEE9) into pFN21A. SNAP-tagged versions of SUDS3, SAP30 and SAP30L were generated by subcloning the coding sequences between the Sgli and Pmel restriction sites in SNAP-FLAG pcDNA5 (supplementary Data).

**Stable Cell Line Construction—** Flip-In™-293 cell lines stably expressing Halo-SUDS3, Halo-SAP30 or Halo-SAP30L were generated using the Flip-In™ System (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. In brief, Flip-In™-293 host cells were cotransfected with pcDNAs/FRT constructs (coding for a Halo-Sin3 subunit) and pOG44 (coding for Flip recombinase) and cultured in media containing Hygromycin B. Twenty-four hygromycin resistant clones were isolated and expression of the Halo-Sin3 subunit was confirmed by Western blotting. Three stable transformants for each subunit were selected for expansion.

**Preparation of Whole Cell Extracts—** Extracts were prepared from Flip-In™-293 cell lines essentially as described previously (10) with minor modifications. Approximately 1×10⁶ cells were collected, washed twice with ice-cold PBS, incubated at –80 °C until frozen, thawed, and resuspended in ice-cold lysis buffer containing 20 mM HEPES (pH 7.5), 1.5 mM MgCl₂, 0.42 mM NaCl, 10 mM KCl, 0.2% Triton X-100, 0.5 mM DTT, 0.1 mM benzamidine HCL, 55 μM phenanthrolinol, 10 μM bestatin, 20 μM leupeptin, 5 μM pepstatin A, 1 mM PMSF, and 500 units SAN (Salt Active Nuclease). Following Dounce homogenization, extracts were incubated for 2 h at 4 °C and centrifuged 40,000 × g for 30 min at 4 °C. The salt concentration was adjusted to 0.3 mM NaCl by adding ice-cold buffer containing 10 mM HEPES (pH 7.5), 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT, 0.1 mM benzamidine HCL, 55 μM phenanthrolinol, 10 μM bestatin, 20 μM leupeptin, 5 μM pepstatin A, and 1 mM PMSF. Lysates were centrifuged at 40,000 × g for 30 min at 4 °C.

For transiently transfected HEK293T cells, lysates were prepared from ~1 × 10⁷ cells transiently transfected with pFN21A based...
plasmids expressing the Sin3 subunits indicated in the figure legends as described previously (11). In brief, cells were harvested ~ 48 h after transfection, washed twice in ice cold PBS and the resulting cell pellet was frozen and thawed. Cells were resuspended in lysis buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton® X-100, 0.1% sodium deoxycholate, 0.1 mM benzamidine HCl, 55 mM phenanthroline, 10 mM beutepain, 20 µM leupeptin, 5 µM pepstatin A and 1 mM PMSF. Extracts were passed through a 26-gauge needle 5 times, centrifuged at 21,000 x g for 30 min at 4 °C, and 300 µl of the resulting supernatant was diluted with 700 µl of TBS (25 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl) for further analysis.

Purification of Recombinant Proteins Complexes—For HaloTag purifications, lysates were incubated with Magne® HaloTag® beads prepared from either 0.1 ml of bead slurry (transiently transfected HEK293T cells) or 0.2 ml of bead slurry (Flip-in™-293 cell lines) overnight at 4 °C. Beads were washed a minimum of four times with buffer containing either 10 mM HEPES (pH 7.5), 1.5 mM MgCl2, 0.3 mM NaCl, or 0.2% Triton X-100 (Flip-in™-293 cell lines) or 50 mM Tris-HCl pH 7.4, 137 mM NaCl, 2.7 mM KCl and 0.05% Nonidet® P40 (transiently transfected cells). Bound proteins were eluted by incubating the beads with buffer containing 50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 1 mM DTT, and 2 Units AcTEV™ Protease (Invitrogen) for 2 h at 25 °C. Eluates were passed through a Micro Bio-Spin™ chromatography column (Bio-Rad) to remove residual traces of beads. For SNAP purifications, lysates were incubated overnight at 4 °C with SNAP-Capture magnetic beads prepared from 0.08 ml bead slurry. Beads were washed with the wash buffer described for Halo purifications above. Bound proteins were eluted by incubating the beads with buffer containing 50 mM Tris-HCl pH 8.0, 0.5 mM EDTA, 1 mM DTT, and 1 Unit PreScission Protease at 4 °C overnight.

Mass Spectrometry—HaloTag® protein complexes were TCA precipitated and digested with Trypsin as described by Swanson et al. (12). Digested samples were first loaded onto microcapillary columns packed with three immobilized phases (reversed phase; strong cation exchange; reversed phase) and bound peptides were then eluted from the column using an Agilent 1100 series quaternary HPLC pump. Peptides were resolved using 10 multidimensional chromatography steps as previously described (13), and analyzed using a linear ion trap (LTQ) mass spectrometer (Thermo Fisher Scientific, Waltham, MA).

The resulting .RAW files were converted to .mz files using RAWDistiller v. 1.0 before matching MS/MS spectra to a database containing: 36628 human protein sequences (National Center for Biotechnology Information, June 2016 release) and 199 common contaminants using the ProLuCID algorithm version 1.3.5 (14). The database also included shuffled versions of all sequences to enable estimation of false discovery rates (FDRs). The total number of sequences searched was 73654. The database was searched for peptides with static modifications of +57 Daltons on cysteine residues (carboxamidomethylation) and with dynamic modifications of +16 Daltons on methionine residues (oxidation). A mass tolerance of 4.5 amu was used for precursor ions and 0 amu for fragment ions. Only fully tryptic peptide sequences were considered.

Inaccurate matches were filtered out using DTASelect (15) in combination with an in-house software algorithm, swallow. Swallow works with DTASelect to select peptide spectral matches at given specific spectrum, peptide, and protein FDRs. The spectral and protein FDRs for each run are presented in supplemental Table S1. The mean spectral FDR for the 33 MudPIT runs was 0.19% ± 0.15% (standard deviation), and the mean protein FDR was 0.92% ± 0.95% (standard deviation). A DTASelect filter also established a minimum peptide length of 7 amino acids, and proteins that were subsets of others were removed using the parsimony option in Contrast (15).

We have reported some of the mass spectrometry data used in this study previously. A summary of the mass spectrometry runs used in this study, including details of where they were first reported is presented in supplemental Table S1. Mass spectrometry data has also been deposited to the PeptideAtlas repository (16) (www.peptideatlas.org), or in the MassIVE repository (http://massive.ucsd.edu); the identifiers for each data set are listed in supplemental Table S1.

Fluorescence Microscopy—Microscopy was performed essentially as described previously (17). In brief, cells stably expressing either Halo-SUDS3 or Halo-SAP30 were plated at 40% confluence in MatTek glass bottom culture dishes. Cells were first cultured for ~ 24 h at 37 °C in 5% CO2. HaloTag® TMRDirect™ was then added to the culture medium at a concentration of 20 nM to label Halo tagged proteins, and cells were further cultured overnight at 37 °C in 5% CO2. Cells were stained with Hoechst dye for 1 h to label nuclei, washed twice with Opti-MEM® reduced serum medium, and imaged using either an UltraVIEW VoX spinning disk confocal microscope (Perkin Elmer) with a 40x plan-apochromat (NA 1.2) objective or a LSM (Zeiss, Thornwood, NY) microscope with a 40x C-Apochromat (NA 1.2) objective. We excited TMRDirect and Hoescht with a 561 and 405 nm laser, respectively. All experiments implemented a multiband dichroic filter (Vox 405/488/561/640 nm or LSM 405/458/561) and a multiband emission filter (Vox 415–475/580–650 nm or LSM 426–583/583–696 nm).

Experimental Design and Statistical Rationale—AP-MS analyses used a minimum of three biological replicates (supplemental Table S1). The number of biological replicates for analyses using transiently transfected cells was: 9 (Halo-control); 3 (Halo-SUDS3); 3 (Halo-SAP30). For analyses using stable cell lines, the number of biological replicates used was: 5 (controls); 4 (Halo-SUDS3); 5 (Halo-SAP30); 4 (Halo-SAP30L). Search data was analyzed using Contrast and NSAF7 software (15) and lists of proteins with a high probability of being in the experimental samples relative to the control samples were made using parameters generated by the QSPEC algorithm (18). A high-confidence interaction was established when a protein was both identified in > 50% experimental replicates, and when QSPEC interaction parameters met the conditions QSPEC log2FC > 2, and QSPEC FDR < 0.05. Four samples were excluded from analysis for the following reasons: MudPIT run failure (1 sample: Halo-control #1, transient transfection experiment), sample cross contamination during purification (1 sample: Halo-control #1, stable cell line experiment), fewer than 200 total proteins detected (1 sample: Halo-SAP30L #2, stable cell line experiment), impure sample of TEV protease used for elution (1 sample: Halo-SUDS3 #1, stable cell line experiment).

RESULTS

Establishing a Structured Workflow for Mapping Sin3 Complex Interactions—Previously, we had identified a set of more than 20 proteins consistently copurifying with several FLAG-tagged subunits of the human Sin3 complex (5). To map the network of interactions among these proteins more precisely, as well as to investigate the interactions between the Sin3 complex and other protein complexes more broadly, we developed a structured approach for building a Sin3 protein interaction network that combines innovative features of the Halo affinity tag with Multidimensional Protein Identification Technology (MudPIT) (19).

The decision-making process that we used to construct and test our Halo- tagged reagents is outlined in Fig. 1A using
A Structured Workflow for AP-MS

**Fig. 1.** Establishing a structured workflow for mapping Sin3 complex interactions. A, Sin3 subunit ORFs are first cloned into pFN21A for transient CMV driven expression of the N-terminally Halo tagged subunit in 293T cells ➀. The pFN21A construct is used in trial AP-MS experiments to assess whether the recombinant subunit can capture other known Sin3 subunits ➁. If the Halo- tagged subunit copurifies with Sin3 complex components, the subunit ORF is cloned into pcDNA5/FRT for building Flp-In™-293 cell lines ➂; otherwise, the bait protein is reengineered by changing the location of the affinity tag or it’s expression is increased by adjusting the strength of the promoter ➃. The pcDNA5/FRT construct is tested for bait expression in 293T cells prior to generating cell lines expressing the Halo-tagged subunit using the weak CMVd2 promoter. Clonal isolates are screened by SDS-PAGE followed by Western blotting and three isolates expressing each Halo-tagged subunit are screened for correct localization of the subunit to the nucleus using fluorescently-tagged cell-permeable Halo ligands ➄. Finally, complexes are purified from 1 x 10⁹ Flp-In™-293 cells stably expressing the Halo tagged subunit and analyzed by MudPIT ➅.

B, Application of the workflow to developing 23 Sin3 related stable cell lines for AP-MS analysis. Of the 5 cases where we modified the bait expression strategy, we used a higher strength promoter for detectable expression in 3 cases and switched tag location in another 2 cases (the tag caused both improper association with Sin3 complexes and abnormal localization). C, An example of affinity tag location influencing bait localization in HEK293T cells expressing either Halo-HDAC2 or HDAC2-Halo. Halo tag localization in live cells is with HaloTag® TMRDirect™ fluorescent ligand (red); nuclei are stained with Hoechst dye (blue). D, Temporal (green) and economic (blue) impact of testing transiently transfected bait constructs prior to investigating cell lines stably expressing baits.

---

**Table B**

| Category                        | Number | Percentage |
|---------------------------------|--------|------------|
| Unique ORFs tested              | 23     | 100.0      |
| Switched tag location           | 2/23   | 8.7        |
| CMV substituted for CMVd2 promoter | 3/23   | 13.0       |
| strategy modified               | 5/23   | 21.7       |

---

**Table D**

| Reagent                      | Costs          | Step 1 | Scale Up Purifications 2 x Roller Bottles |
|------------------------------|----------------|--------|-------------------------------------------|
| Costs to generate stable cell line | $169.95       |        |                                           |
| Time to generate stable cell line | 24 days       |        |                                           |
| Media                        | $3.87          | $38.87 |                                           |
| Culture containers           | $1.65          | $15.98 |                                           |
| Time for cell propagation    | 3 days         | 12 days |                                           |
| Transfection Reagent         | $23            |        |                                           |
| Protease Inhibitors          | $0.44          | $25.90 |                                           |
| Magnetic Beads               | $11.50         | $23.00 |                                           |
| Time for purification        | 0.1 day        | 0.5 days |                                           |
| Total costs                  | $40.46         | $273.70 |                                           |
| Total costs                  | 3.1 days       | 36.5 days |                                           |
standard flowchart symbols (20). First, we subcloned coding sequences for Sin3 subunits into pFN21A for expression of the subunit with an N-terminal Halo tag in human cells. pFN21A drives robust expression of the recombinant protein using the CMV promoter allowing us to test the behavior of the protein in small scale experiments (Fig. 1A). To ask whether the tagged subunit could associate with other Sin3 subunits, we Halo-purified complexes from lysates of cells transfected with the pFN21A (CMV) construct and analyzed the proteins copurifying with the Halo-tagged bait by MudPIT. If the expected Sin3 subunits failed to copurify with the bait, we then moved the Halo tag to the C terminus of the protein and retested the ability of the bait to connect with other Sin3 complex components.

Once we had confirmed that other Sin3 subunits were present in purified samples generated using a transiently transfected bait, we made Flp-In™-293 cells stably expressing CMVd2 driven Halo-tagged Sin3 subunits. We had previously found that the weaker CMVd2 promoter allows us to express the tagged recombinant proteins at levels that more closely reflect the expression levels of the endogenous version of the protein (4, 9). Consistent with these previous findings, we found that although CMV driven expression of the tagged Sin3 subunit Halo-SAP30 yields greater amounts of the recombinant compared with the endogenous protein (supplemental Fig. S1), CMVd2 driven expression of Halo-SAP30 yields less recombinant than endogenous SAP30 protein. In addition to enabling lower bait expression levels, the pcDNA5/FRT vector that we used contains an FRT recombination site, which allows systematic integration of the Sin3 subunit into a distinct site in human Flp-In™-293 cells. After confirming expression of the CMVd2 expressed recombinant protein in transiently transfected HEK293T cells, we isolated hygromycin B resistant stable transformants expressing Halo-tagged Sin3 subunits in Flp-In™-293 cells as described under “Experimental Procedures.” Having confirmed stable expression of the protein by Western blotting, we imaged these cells to assess proper subunit localization. We found that proper nuclear localization of the recombinant protein could be discerned more clearly in the stable cell lines, which had modest isogenic protein expression, than in the transiently transfected cells, which were susceptible to mislocalization of the overexpressed recombinant protein (21). If the recombinant protein was not localized to the nucleus, we again adapted our tagging strategy. For recombinant proteins that both copurified with other Sin3 subunits in trial AP-MS experiments and localized to the nucleus when stably expressed, we used the stable cell lines for large scale AP-MS analyses. Using a larger number of stably expressing cells (10⁶) than we had with transiently transfected cells (10⁷) allowed us to isolate sufficient amounts of the CMVd2 (weakly expressed) Halo-tagged bait protein for further analysis.

We have applied this workflow for the development of 23 cell lines stably expressing affinity tagged Sin3 related proteins (Fig. 1B). We modified our strategy in 5 cases (21.7%). In two of these cases, we moved the tag to the C terminus of the protein (this resulted both in a more robust association with Sin3 subunits and in proper nuclear localization (Fig. 1C)). The economic cost of identifying problems such as these quickly by performing trial mass spectrometry analyses ($40.46 × 23 = $930 for a single analysis of each bait (Fig. 1D)) was mitigated by economic savings in not using suboptimal reagents for AP-MS analyses ($273 per biological replicate per cell line). Importantly, we achieved significant time savings by identifying suboptimal reagents early, for example we could purify a sample from transiently transfected cells for AP-MS analysis within 3.1 days, whereas generation of stable cell lines prior to purification extended this time to 36.5 days (Fig. 1D).

This structured workflow (Fig. 1) not only establishes a basis for building a detailed Sin3 complex interaction network in Flp-In™-293 cells, but also lays the groundwork for investigating Sin3 interaction networks in other cell types and under different cellular conditions. For example, although there is evidence that the Sin3 interactome may be modulated in response to drug treatment with HDAC inhibitors in HEK293 cells (5), such effects have yet to be studied in clinically relevant cancer cell lines. Two key Sin3 subunits, SUDS3 and SAP30, have well characterized direct interactions with the core Sin3 scaffolding protein SIN3A (22–25), and so we initially characterized these two proteins using our Halo-MudPIT workflow; the results of these analyses are presented below to illustrate our approach.

**Transiently Transfected Halo-SUD33 and Halo-SAP30 Associate with Other Sin3 Complex Components**—The Sin3 subunits SUDS3 and SAP30 are tightly integrated with SIN3A. The interaction between SUDS3 and SIN3A has been mapped by NMR and involves an alpha helix within SUDS3 residues 205–228 binding the SIN3A HDAC interaction domain (SIN3A[608–729]) (25). The interface between SAP30 and the Sin3 complex has also been well characterized and involves contacts between a SIN3A paired amphipathic helix domain (PAH3) and a SAP30 domain that resembles the SAP DNA binding motif (23). Consistent with these observations, we also found that overexpressed SIN3A copurified with affinity tagged versions of SUDS3 and SAP30, as well as with a paralog of SAP30, SAP30L (supplemental Fig. S2). To test whether N-terminally Halo-tagged versions of human SUDS3 and SAP30 might associate with the full complement of known Sin3 subunits, we purified either Halo-SUDS3 or Halo-SAP30 complexes from transiently transfected HEK293T cells and identified proteins in the resulting eluates using MudPIT (Fig. 2A). Cells for control samples were transfected with plasmids expressing the Halo tag alone. Peptides shared by multiple proteins (for example peptides present in different isoforms or paralogs) were assigned to proteins as described by Zhang et al. (26). We next used the QSPEC algorithm (18) to define sets of proteins significantly enriched in experimen-
Fig. 2. Transiently transfected Halo-SUDS3 and Halo-SAP30 associate with other Sin3 complex components. A, Halo-MudPIT analysis of transiently expressed Halo-SUDS3. Protein complexes were Halo affinity purified from HEK293T cells transiently transfected with Halo-SUDS3 in pFN21A (3 replicates) or Halo tag alone controls (9 replicates). Proteins in the eluates were identified by digesting samples with trypsin prior to separating tryptic peptides using multidimensional chromatography for mass spectrometry analysis (MudPIT). Proteins identified in each replicate were quantitated by spectral counting; proteins significantly enriched in Halo-SUDS3 compared with control samples were identified using QSPE (18) (QSPE log2FC > 2, QSPE FDR \( < 0.05 \), detected in > 50% replicates; supplemental Table S2). B, Gene ontology terms (cellular component) enriched in Halo-SUDS3 or Halo-SAP30 complexes. Groups of proteins significantly enriched with either Halo-SUDS3 or Halo-SAP30 in pFN21A transiently expressed in HEK293T cells were analyzed for GO term enrichment (cellular component) using the DAVID annotation tool (27) (supplemental Table S2, “DAVID analysis results”); Significantly enriched GO terms (\( p_{\text{adj}} < 0.05 \)) were summarized using REViGO (57) (supplemental Table S2, “Revigo analysis results”); treemap segment areas are proportional to \(-\log_{10} p_{\text{adj}}\). C, Sin3 complex subunits copurify with both Halo-SUDS3 and Halo-SAP30. dNSAF values for Sin3 subunits copurifying with Halo-SUDS3 or Halo-SAP30 visualized using Circos (58). Values indicated are dNSAF av × 10000. Bait dNSAF values have been set to 0 for visualization. D, Expression of SAP25 and SAP18 mRNA in HEK293T cells. FPKM values (Fragments Per Kilobase of transcript per Million fragments mapped) for the indicated Sin3 subunit transcripts are taken from our previously published RNA-seq dataset (4) available from the NCBI GEO (Gene Expression Omnibus) repository under the accession number: GSE79656. Values are the mean of 3 biological replicates. Error bars represent standard deviation. E, SUDS3 and SAP30 copurify with Halo-SAP25, but not with Halo-SAP18-Halo. Halo purified samples from HEK293T cell lysates transfected with either Halo-SAP25, Halo-SAP18, or a control plasmid expressing the Halo tag alone were resolved by SDS-PAGE. Copurifying proteins were either detected by Coomassie staining or by Western blotting using the antibodies indicated. The Coomassie stained bands corresponding to the bait protein after Halo tag cleavage are indicated by yellow asterisks. F, Halo-SUDS3 and Halo-SAP30 stably expressed in Flp-In™-293 cells localize to the nucleus. Flp-In™-293 cells were cotransfected with pOG44 and Halo-SUDS3 or Halo-SAP30 (in CMVd2 pcDNA5/FRT PacI Pmel) and colonies of hygromycin resistant cells isolated and screened for recombinant protein expression by SDS page and Western blotting. Western blots were probed with rabbit anti-Halo polyclonal antibodies and IRDye® 800CW labeled anti-Rabbit secondary antibodies together with anti-tubulin mouse monoclonal antibodies and IRDye®
tal but not control replicates (log₂FC > 2, FDR < 0.05), limiting the sets to proteins detected in > 50% of experimental replicates (Fig. 2A). The 146 SUDS3-associated and 100 SAP30-associated proteins are listed in supplemental Table S2. To assess whether these sets of proteins contained known Sin3 subunits, we first used the DAVID bioinformatics resource (27) to survey these lists for enriched gene ontology (GO) terms in the cellular component category (Fig. 2B and supplemental Table S2). GO terms associated with both sets of proteins were enriched for the descriptor “Sin3 complex” (GO: 0016580), with $p_{adj} = 2.5 \times 10^{-12}$ (SUDS3) and $p_{adj} = 3.3 \times 10^{-10}$ (SAP30). In addition to examining proteins for enriched GO terms, we asked whether our sets of Halo-SUDS3 and Halo-SAP30 associated proteins contained 22 Sin3 subunits and associated proteins previously described by our group and others (5, 6, 28), only 7 of which have been annotated with the term GO:0016580 (Sin3 complex). We have grouped these 22 proteins as either “core Sin3 subunits” (a descriptor used by Silverstein et al. (29)/Grzenda et al. (6)), “Sin3 other associated proteins” (a descriptor used by Grzenda et al. (6)), or “Sin3 transcription factors” (a table outlining these descriptors is presented in supplemental Table S2). Both Halo-SUDS3 and Halo-SAP30 purifications contained 20 of these proteins (Fig. 2C and supplemental Table S2). Intriguingly neither SUDS3 nor SAP30 associated with SAP18 or SAP25. Although SAP18 had initially been described as a subunit of the Sin3 complex (30), SAP18 functions in the absence of other Sin3 subunits as part of the ASAP complex to regulate RNA splicing (31, 32). Murine SAP25 had previously been identified as a core component of the mSin3 complex (28) and exogenously expressed mSAP25 also associates with human Sin3 components in HeLa cell lysates (28). To ascertain whether SAP25 and SAP18 are available for integration into the Sin3 complex in HEK293T cells, we analyzed RNA-seq data that we had generated in a previous study using RNA microarray analysis of the eluates from these purifications by SDS-PAGE and silver staining indicated putative endogenous Sin3 subunits in the experimental purifications but not in controls (a representative analysis is shown in Fig. 3A). MudPIT mass spectrometry analysis of the eluates confirmed the presence of core Sin3 subunits, other Sin3 associated proteins and Sin3 associated transcription factors (Fig. 3B). Consistent with our previous observations, we failed to detect significant amounts of the Sin3 subunits SAP18 or SAP25. In addition to monitoring the purifications for the presence of Sin3 subunits, we also asked whether the three bait proteins were associating primarily with the Sin3 complex, or whether they additionally copurified with other groups of proteins. There were 21 proteins copurifying with all three bait proteins (Fig. 3C). We used the ClueGO/CluePedia algorithm (33, 34) to ask if this subset of proteins shared GO terms and found that 16/21 of the proteins shared the associated enriched GO terms “nuclear transcriptional repressor complex” and/or “histone deacetylase complex.” The other 5 proteins include: the previously characterized Sin3 associated proteins ARID4A, ARID4B and ING1; FASN (fatty acid synthase); and LGALS3BP, a glycoprotein elevated in human colorectal carcinoma (34) that might function as an immune stimulator after oncogenic transformation (34). In addition to these 21 shared proteins, a second set of 38 proteins consistently copurified with Halo-SAP30L, but not with Halo-SAP30 or with Halo-

680LT labeled goat anti-Mouse secondary antibodies (loading control). Western blots were imaged using a Li-Cor infra-red imaging system. Subcellular localization of Halo-SUDS3 and Halo-SAP30 proteins was assessed by labeling live cells with HaloTag® TMRDirect™ fluorescent ligand (red) and staining DNA (nuclei) with Hoechst dye (blue).
SUDS3. ClueGO analysis revealed that these included several tubulin proteins as well as components of the proteasome. Although we and others (35) found Halo-SAP30L or endogenous SAP30L mainly in the nucleus and the majority of cellular tubulin is detected in the cytoplasm (supplemental Fig. S3), SAP30L could potentially associate with the fraction of cellular tubulin associated with chromatin in the nucleus (36, 37).

**Halo-MudPIT Analysis Extends Network of STRING Annotated Sin3 Subunit Interactions**—We next wanted to compare interactions captured using our Halo-MudPIT approach to interactions that have been previously reported. We initially compared our data to experimental evidence for protein-protein interactions curated in the STRING database (38). To ask whether STRING-identified Sin3 subunit interactors were also identified using our Halo-MudPIT approach, we used the DyNet analysis tool (39) to compare our network of high-confidence Halo-MudPIT identified interactions (log\(_{2}\)FC > 2, FDR < 0.05) with medium confidence STRING annotated interactions (STRING score > 0.4) for the same bait proteins (Fig 4A).

Although most bona fide Sin3 subunits were identified both in our purifications and by STRING by at least one bait (Fig. 4A gray lines), neither FAM60A nor SAP30L were identified by STRING. In contrast, both FAM60A and SAP30L were identified in all of our Halo-MudPIT analyzed samples. In addition, whereas none of the Sin3 core subunits/other associated proteins were detected by STRING alone, we detected many interactions between either Halo-SUDS3, Halo-SAP30 or Halo-SAP30L.

---

**Fig. 3. Evaluating complexes purified from Sin3 stable cell lines.** A, Complexes purified from Flp-In™-293 cells expressing CMVd2 driven Halo-SUDS3. Complexes were purified from five replicates of ~1 \(\times\) 10\(^{6}\) Flp-In™-293 cells stably expressing Halo-SUDS3 using the CMVd2 promoter. Proteins present in the eluates were resolved by SDS-PAGE using 4–15% gradient gels and visualized by silver staining (replicate 1, clone #14 is shown as a representative example). Three biologically distinct clonal Halo-SUDS3 cell lines were used (clone #2, replicate 2; clone #14, replicates 2 and 4; clone #24, replicates 3 and 5) and replicate purifications were performed by two people (replicates 1 and 3 by CAB and replicates 2, 4 and 5 by JLT) on five separate dates to avoid batch effects. B, Sin3 subunits identified in complexes purified from Flp-In™-293 cell lines stably expressing Halo-SUDS3, Halo-SAP30 and Halo-SAP30L. Proteins from replicate purifications were digested with trypsin and identified using MudPIT (Fig. 2A and supplemental Tables S1 and S3). Average dNSAF values for components of the Sin3 complex (green and blue) and Sin3 associated transcription factors (pink) have been visualized using Genesis (39). C, Gene ontology terms (cellular component) enriched in complexes purified from cells stably expressing Halo-SUDS3, Halo-SAP30 or Halo-SAP30L. Groups of proteins significantly enriched (QSPEC log\(_{2}\)FC > 2, QSPEC FDR\(_{UP} < 0.05\), detected in > 50% replicates) with Halo-SUDS3, Halo-SAP30 or Halo-SAP30L were identified (supplemental Table S3); the overlap among these groups was visualized using Venny (60). Gene ontology terms (cellular component) enriched in the indicated subgroups (p\(_{adj}\) < 0.05) were determined using the ClueGO/CluePedia tool (ClueGO v2.3.3 and CluePedia v1.3.3) with default settings (33)(34). Protein connections to enriched GO terms are indicated (red).
Fig. 4. Halo-MudPIT analysis extends network of STRING annotated Sin3 subunit interactions. A, Complexes were purified from Flp-In™-293 cells stably expressing Halo-SUDS3, Halo-SAP30 or Halo-SAP30L (bait proteins - green hexagons). Proteins significantly enriched in these bait purifications were identified by Halo-MudPIT AP-MS (QSPEC log2FC > 2, QSPEC FDR < 0.05, detected in > 50% replicates; supplemental Table S3). Interactions derived from published experimental evidence for proteins interacting with human SUDS3, SAP30 and SAP30L were retrieved from the STRING database (38) (medium confidence interactions (STRING score >0.4). Halo-bait captured and STRING annotated interactions were compared using the DyNet analysis tool (39). Interactions detected by Halo-bait only (green edges), STRING only (red edges) or by both Halo-bait and STRING (gray edges) are indicated. Nodes representing core Sin3 subunits (green
lo-SAP30L and core Sin3 subunits/other associated proteins that were not detected by STRING (Fig. 4A, green lines connecting the Halo-tagged baits to other Sin3 subunits/associated proteins).

Several proteins were annotated in the STRING network and not captured in our purifications. First CTPS1 and CTPS2 were shown as interacting with both SAP30 and SAP30L in the STRING network but were not in our purifications. Peculiarly, this STRING annotation is based on an unrefenced interaction observed among putative homologs of these proteins in other species. Similarly, most of the SUD3 interactions detected by STRING only (red lines connecting SUD3 to MORF4L1, MORF4L2, AIRE, MSL3, TAF3, PHF12, and ING3/4/5) are not in fact because of experimentally evidenced interactions among the human proteins but are based on interactions among putative homologs in other species. The other SUD3 interactors detected by STRING only, LAC31 (C13orf31) and MXD1 were reported in a high throughput screen for interactors in HEK293 cells (2). In this screen, both proteins used as baits detected SUD3 as prey, whereas LAC31 and MXD1 were not detected in SUD3 purifications, suggesting that overexpression of LAC31 and MXD1 might be needed to capture these interactions. Retinol Binding Protein 1 (RBP1), recently reported in association with SAP30 in HeLa cell purifications (1), was also not detected in our Flp-In™-293 Halo-SAP30 purifications.

In addition to comparing the components captured by our Halo-tagged baits with interactions listed in the STRING database, we asked whether our baits captured proteins previously detected in FLAG-SAP30 and FLAG-SAP30L purifications by Sardiu et al. (5) (Fig. 4B). Except for the transcription factor BBX, we detected all proteins that had been previously captured by both FLAG-SAP30 and FLAG-SAP30L in all our Halo purifications (Fig. 4B). Curiously, we did not detect many of the proteins that were previously only found in FLAG-SAP30L (but not in FLAG-SAP30) purifications (Fig. 4B). As we had used a different approach to purifying complexes using Halo-tagged baits, the absence of these proteins might result from the loss of weakly associated interactors under our modified purification conditions.

Our Halo-MudPIT approach for capturing high-confidence Sin3 complex interactions recapitulates most information annotated in the STRING database (Fig. 4A) as well as most previously reported interactions defined using FLAG-tagged Sin3 subunits (5). Taken together, the analyses of Fig. 4 support that our Halo AP-MS procedure isolates Sin3 complexes containing most Sin3 components isolated previously in other studies.

Quantifying Relative Sin3 Subunit Composition for Different Bait Purifications with the Halo-MudPIT Approach—Previously, we found evidence that populations of Sin3 complexes appear to be heterogeneous, with some subunits not appearing together in the same complex. For example, although complexes purified through the SAP30 subunit contain most other Sin3 subunits, they do not contain the homologous subunit SAP30L; similarly, SAP30L complexes do not contain SAP30 (5). When we compare Sin3 complexes purified using different baits, we would like to ask whether each bait associates with a similar set of Sin3 subunits, or whether there is variation in subunit preference among baits. To estimate the relative levels of different subunits within specific complexes, some previous studies have used complex normalized spectral abundance factor (cNSAF) values instead of dNSAF values (40, 41). When we calculated cNSAF values for Sin3 complexes purified from cells stably expressing Halo-SUD3, Halo-SAP30, and Halo-SAP30L, we noticed that Halo-SUD3 purifications contained lower cNSAF values than the other baits for several subunits including SIN3A, HDAC1, ING1/2, and FAM60A (Fig. 5A (upper bar graph)). We thought that such differences might result from technical differences among purifications using different bait proteins. Although our AP-MS strategy uses the CMVd2 promoter to limit the effects of overexpressing bait proteins, we still observe differences in expression among different baits. Some baits might capture comparatively more of the endogenous Sin3 complexes than others, either because they are expressed at a higher level or because they are more efficiently incorporated into Sin3 complexes (Fig. 5B). Alternatively, some bait proteins (SUD3) might dissociate faster from the Sin3 complex than other baits (SAP30 or SAP30L), resulting in proportionally more bait than prey in these purifications. Therefore, if we compared the cNSAF values for a given prey captured by three different baits, it would be difficult to assess whether a particular low (or high) cNSAF value for a given bait prey complex reflected abnormal bait expression/incorporation into Sin3 complexes, or alternatively reflected a genuine antagonism for this combination of subunits in cells (for example, SAP30 and SAP30L not coexisting in the same Sin3 complex).

With these two possibilities in mind, we decided to calculate the ratios of the cNSAF value of each bait to the cNSAF value of a key complex subunit (in this case we chose SIN3A as the key subunit) (Fig. 5A (lower bar graph) and Fig. 5B). If different baits were to capture heterogeneous populations of SIN3A containing complexes without preference for subunit combinations, we would expect ratios of cNSAF_{PREY} / cNSAF_{SIN3A} to be the same for different baits (Fig. 5B).
We calculated the ratios cNSAF_{PREY}/cNSAF_{SIN3A} for the baits Halo-SUDS3, Halo-SAP30 and Halo-SAP30L (Fig. 5A (lower bar graph)) and found that the cNSAF_{PREY}/cNSAF_{SIN3A} ratios for the preys HDAC1, ING1/2, and FAM60A varied less than the original cNSAF values had done between the three baits (compare Fig. 5A (upper bar graph) with Fig. 5A (lower bar graph)). This is consistent with a model where the lower cNSAF values for some Halo-SUDS3 preys are caused by Halo-SUDS3 capturing proportionally lower amounts of SIN3A containing complex, rather than by a preference of Halo-SUDS3 for SIN3A containing complexes lacking these subunits. In contrast, this analysis supported that Halo-SAP30 captured minimal amounts of endogenous SAP30L, and that Halo-SAP30L captured minimal amounts of endogenous SAP30 (Fig. 5A (lower bar graph)), corroborating our previous findings that these subunits likely do not coexist in the same Sin3 complexes (5). Although we chose the cNSAF approach here, repeating the analysis using dNSAF rather than cNSAF values gives similar results (supplemental Fig. S4).

Intriguingly, this analysis suggested that SIN3B associated with Halo-SUDS3 but not Halo-SAP30 or Halo-SAP30L. To ask whether SIN3B does indeed associate with SUDS3, but...
not with SAP30, we expressed either Halo-SIN3A, Halo-SIN3B or Halo tag alone in HEK293T cells and isolated complexes using Halo affinity purification (Fig. 5C). Consistent with the results of Fig. 5A, Halo-SIN3A captured both endogenous SUDS3 and SAP30. In contrast, Halo-SIN3B captured SUDS3 but significantly less SAP30 (Fig. 5C).

In summary, by normalizing to a key complex subunit cNSAF value (cNSAFKS), and then comparing the prey cNSAFPREY/cNSAFKS values across different baits, we can address whether different baits associate with Sin3 complexes containing distinct combinations of subunits.

**Differential Network Analysis of “Transient Transfection” and “Stable Cell Line” Networks**—Having used two different methods to analyze Sin3 associated proteins, first using transiently transfected cells overexpressing Sin3 subunits, and second using larger numbers of cells stably expressing Sin3 subunits at lower levels, we decided to compare the resulting networks of Sin3 subunit interactions mapped using each method. To do this we first generated two networks made using either transiently transfected Halo-SUDS3 and Halo-SAP30 (transient network), or with stably expressing Halo-SUDS3 or Halo-SAP30 cells (stable network). We used the QSPEC parameter log2FC (42) to quantitate prey proteins enriched with the bait proteins in these networks. We next compared these two networks using the DyNet tool (39) to generate the differential network shown in Fig. 6. This differential network indicates the differences in log2FC values (Δ log2FC) between the transient and stable networks by edge color and width; node colors indicate whether prey proteins are present in the transient network or stable network only, or in both networks.

Subunits of the Sin3 complex (diamond nodes) and several previously identified Sin3 associated proteins (hexagonal nodes) including the transcription factors FOXK1 and FOXK2, and the trinucleotide repeat containing protein TNRC18 (43, 44), were present in both transient and stable networks. In addition, the previously identified Sin3 associated protein BAHCC1 (44), a TNRC18 paralog, was identified only in the stable network. Of these 21 Sin3 subunits/associated proteins, most (17) only showed greater QSPEC log2FC values in the stable network (pink edges, Fig. 6), suggesting a greater enrichment of bona fide Sin3 complex components in the stable cell line purifications. One hundred fifty-three proteins were present in the transient network only, or in both networks.

In developing the workflow that we used to build suitable reagents for AP-MS analyses, we anticipated, and tried to avoid, several possible sources of systematic error. First, although we used a system designed to integrate our recombinant bait at a predetermined genomic locus to create isogenic cell lines, we still considered the possibility of variation among different clonal isolates. For example, Fukushige and Sauer (50) had previously used targeted DNA integration to make β-galactosidase expressing cell lines and noted higher β-galactosidase expression in some anomalous transfor-
mants. They identified chromosomal duplication in cells from one such transformant as a possible source for the difference in phenotype. To reduce the impact of possible artifacts resulting from abnormal cell lines, we initially purified complexes from three distinct clonal isolates for each Sin3 subunit analyzed (in some cases only two clonal isolates were represented in final AP-MS analyses; for example, one Halo-SAP30L replicate was excluded from our final analysis—see Experimental Procedures). A comparison of the three clonal isolates used for 5 replicate Halo-SAP30 AP-MS analyses (supplemental Fig. S6) did not suggest consistent differences in prey dNSAF values that depended on the clonal isolate used. In addition to taking measures to mitigate biological artifacts from abnormal cell lines, we also considered the impact of systematic technical artifacts that might be introduced during purification. Indeed, Leek et al. had previously stressed the negative impact of such batch effects for studies using high throughput technologies (51). As small amounts of buffer contaminants could result in spurious protein identifications with the sensitive protein identification technology that we were using, we avoided situations where all replicates corresponding to a specific cell line were analyzed in parallel.
For example, four people performed the 18 purifications used for the stable cell line network on 14 separate dates with no similar samples processed together by the same person.

Having acted to reduce systematic errors in analyses using cells stably expressing a bait, we noticed reproducible differences between these analyses and similar analyses using transiently transfected cells and an alternative purification protocol. Although we have not systematically investigated the causes for these differences, we believe there are several factors that might explain them. First, the Halo purification protocol for affinity purifying complexes from transiently transfected cells did not use a nuclease in the lysis buffer. Additional proteins could reflect genuine specific nucleic acid mediated interactions that are captured in the purifications using the transiently transfected lysates. This could explain the presence of RNA binding proteins in these purifications that were not detected in the equivalent purifications using stable cell line lysates. Nucleic acids might also regulate protein–protein interactions by binding to proteins and changing their secondary structure. For example, the TIPS subunit of the chromatin remodeling complex NoRC, which targets HDAC1 containing complexes to ribosomal gene promoters (52), binds to pRNA, promoting a more open configuration exposing new interaction surfaces (53). Other differences among the purification protocols might also explain why some interactions are observed only in the transient transfection experiments. For example, capture of some interactors may depend on the salt concentration in the lysis buffer (here the buffer used with transiently transfected cells contained 150 mM NaCl compared with 0.42 M NaCl in the buffer used to lyse cells stably expressing the bait). Finally, the differences between the two analyses methods might simply result from substantial overexpression of the bait in the transiently transfected cells. Gibson et al. have suggested that transiently overexpressed baits can misfold (21), and misfolded or partially folded proteins can result in aberrant bait–prey interactions (54). To summarize, our observations suggest that the set of proteins identified in AP-MS experiments can change appreciably with different purification protocols; a larger systematic study could determine how different factors influence such changes.

As we focused specifically on components of the Sin3 complex enriched in the stable cell line purifications, we observed a “bait effect” with marked differences in absolute values of cNSAF (or dNSAF) for Sin3 subunits captured using different baits (Fig. 5). Normalizing cNSAF values to the cNSAF value of a key complex subunit can moderate these effects and can help assess whether certain combinations of subunits are favored. Curiously, this analysis indicates that although Halo-SUDS3 appears to assemble into both SIN3A and SIN3B based complexes, Halo-SAP30 and Halo-SAP30L appear to be preferentially used with SIN3A based complexes (Fig. 5A (lower bar graph) and Fig. 5C). None of our purifications contained the subunits SAP18 and SAP25. Although SAP18 was initially identified through its association with Sin3/HDAC complexes (30), Zhang et al. also observed that the interaction depended on exogenous overexpression of SAP18 in HEK293T cells. Schwerk et al. later found that the majority of SAP18 is associated with the RNA processing complex ASAP, and suggested that only a small proportion of endogenous SAP18 might associate with the Sin3 complex (55). We did not detect Sin3 components by Western blotting in either Halo-SAP18 or SAP18-Halo purifications using HEK293T cells (Fig. 2E). In contrast to SAP18, SAP25 appears to be a definitive Sin3/HDAC component but may only be expressed and incorporated into the Sin3 complex in response to distinct cellular signals. Although SAP25 binds to PAH1 domain at the N terminus of the SIN3A scaffolding protein (56), it was found to be mainly cytoplasmic in HeLa cells (28), and its nuclear involvement with Sin3/HDAC complexes may depend on Ras-regulated signal transduction pathways. The absence of SAP25 in our purifications was likely because of low expression of endogenous SAP25 in HEK293T cells (Fig. 2D)—indeed we were able to capture Sin3 components with exogenously expressed Halo-SAP25 (Fig. 2E).

To summarize, we report a robust Halo-MudPIT based workflow for mapping protein–protein interactions among components of large protein complexes and show how the workflow can be applied to identify high-confidence interactions among components of protein complexes such as the Sin3/HDAC complex. Determining the architecture of such complexes is challenging, as cells assemble bespoke versions of the Sin3/HDAC complex with distinct combinations of paralogous subunits for different situations (29). The Halo-MudPIT approach provides a platform for addressing these challenges and developing more refined understanding of Sin3/HDAC subunit organization and usage.

Original data underlying this manuscript can be accessed from the Stowers Original Data Repository at http://www.stowers.org/research/publications/LIBPB-1214.

Acknowledgments—We thank the Stowers Tissue Culture core facility for technical support during this project.

DATA AVAILABILITY

Mass spectrometry datasets are available from the PeptideAtlas repository (www.peptideatlas.org) using identifier PASS00598, and from the MassIVE repository (http://massive.ucsd.edu) using identifiers MSV000081349, MSV000081352, MSV000081355, MSV000081359, MSV000081360, and MSV000081361 (supplemental Table S1).

* This work was supported by the Stowers Institute for Medical Research and the National Institute of General Medical Sciences of the National Institutes of Health under Award Number RO1GM112639 to MPW and F32GM122215 to MKA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[5] This article contains supplemental material.
REFERENCES

1. Hein, M. Y., Hubner, N. C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A., and Mann, M. (2015) A human interactome in three quantitative dimensions organized by stoichiometries and abundances. Cell 163, 712–723.

2. Huttlin, E. L., Ting, L., Bruckner, R. J., Gebreab, F., Gygi, M. P., Szpyt, J., Tam, S., Zarraga, G., Colby, G., Baltier, K., Dong, R., Guanari, V., Valtes, L. P., Ordureau, A., Rad, R., Erickson, B. K., Wühr, M., Chik, J., Zhai, B., Kollipakkam, D., Mintsers, J., Obar, R. A., Harris, T., Artavanis-Tsakonas, S., Sowa, M. E., De Camilli, P., Julio, A. P., Harper, J. W., and Gygi, S. P. (2015) The BioPlex network: a systematic exploration of the human interactome. Cell 162, 425–440.

3. Hubner, N. C., Bird, A. W., Cox, J., Splettstoss, B., Bandilla, P., Poser, I., Hyman, A., and Mann, M. (2010) Quantitative proteomics combined with BAC Transgeneomics reveals in vivo protein interactions. J. Cell Biol. 189, 739–754.

4. Banks, C. A. S., Boanca, G., Lee, Z. T., Eubanks, C. G., Hattem, G. L., Peak, A., Weens, L. E., Conkright, J. J., Floros, L., and Washburn, M. P. (2016) TNIP2 is a hub protein in the Nf-κB network with both protein and RNA mediated interactions. Mol. Cell. Proteomics 15, 3435–3449.

5. Sardiu, M. E., Smith, K. T., Groppe, B. D., Gilmore, J. M., Saraf, A., Egidy, I. A., Weisswange, I., Mansfeld, J., Buchholz, F., Hyman, A. A., and Mann, M. (2014) Controlling for gene expression changes in transcription factor protein networks. Mol. Cell. Proteomics 13, 3114–3125.

6. Grzenda, A., Lomberk, G., Zhang, J.-s., and Urrutia, R. (2009) Sin3: Master Scaffold and Transcriptional Corepressor. Biochim. Biophys. Acta 1789, 443–450.

7. Bansal, N., David, G., Farias, E., and Waxman, S. (2016) in Adv. Cancer Res., eds Croce CM, Fisher PB (Elsevier Inc.), pp 113–135. 1stEd.

8. Kandath, C., Mcellan, D. M., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J. F., Wyczalkowski, M. A., Leiserson, M. D. M., Miller, C. A., Welch, J. S., Walter, M. J., Wendl, M. C., Ley, T. J., Wilson, R. K., Raphael, B. J., and Ding, L. (2013) Mutational landscape and significance across 12 major cancer types. Nature 502, 333–339.

9. Banks, C. A. S., Boanca, G., Lee, Z. T., Floros, L., and Washburn, M. P. (2015) Proteins interacting with cloning scars: a source of false positive protein–protein interactions. Sci. Rep. 5, 8530.

10. Mahour, N., Redwine, W. B., Floros, L., Swanson, S. K., Martin-Brown, S., Bradford, W. D., Staehling-Hampton, K., Washburn, M. P., Conaway, R. C., and Conaway, J. W. (2008) Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. J. Biol. Chem. 283, 8005–8013.

11. Banks, C. A. S., Lee, Z. T., Boanca, G., Lakshminaranisamy, M., Gropp, B. D., Wen, Z., Hattem, G. L., Seidel, C. W., Floros, L., and Washburn, M. P. (2014) Controlling for gene expression changes in transcription factor protein networks. Mol. Cell. Proteomics 13, 1510–1522.

12. Swanson, S., and Floros, L. (2009) Generation and analysis of multidimensional protein identification technology datasets. Methods Mol. Biol. 492, 1–20.

13. Banks, C., Kong, S., and Washburn, M. (2012) Affinity purification of protein complexes for analysis by multidimensional protein identification technology. Mol. Cell. Proteomics 11, 105–115.

14. Xu, T., Park, S. K., Venable, J. D., Wohlschlegel, J. A., Diedrich, J. K., Cociorva, D., Lu, B., Liao, L., Hewel, J., Han, X., Wong, C. C. L., Fonslow, B., Delahunt, C., Gao, Y., Shah, H., and Yates, J. R. (2015) ProLuCID: an improved SEQUEST-like algorithm with enhanced sensitivity and specificity. J. Proteomics 129, 16–24.

15. Tabb, D. L., McDonald, W. H., and Yates III, J. R. (2002) DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome Res. 1, 21–26.

16. Desiere, F., Deutsch, E. W., King, N. L., Nesvizhskii, A. I., Mallick, P., Eng, J., Chen, S., Edels, J., Loevenich, S. N., and Aebersold, R. (2006) The PeptideAtlas project. Nucleic Acids Res. 34, D685–D688.

17. Gilmore, J. M., Sardiu, M. E., Gropp, B. D., Thornton, J. L., Liu, X., Dayebgadot, G., Banks, C. A., Slaughter, B. D., Unruh, J. R., Workman, J. L., Floros, L., and Washburn, M. P. (2016) WDR76 Co-Localizes with heterochromatin related proteins and rapidly responds to DNA damage. pLoS ONE 11, 1–16.

18. Choi, H., Kim, S., Fermin, D., Tsou, C.-C., and Nesvizhskii, A. I. (2015) QPROT: statistical method for testing differential expression using protein-level intensity data in label-free quantitative proteomics. J. Proteome Res. 14, 1291–126.
41. Mosley, A. L., Sardiu, M. E., Pattenden, S. G., Workman, J. L., Florens, L.

43. Yang, Q., Kong, Y., Rothermel, B., Garry, D. J., Bassel-Duby, R., and

38. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-

37. Mithieux, G., Alquier, C., Roux, B., and Rousset, B. (1984) Interaction of

tubulin with chromatin proteins. H1 and core histones. J. Biol. Chem. 259,

15523–15531

36. Menko, A. S., and Tan, K. B. (1980) Nuclear Tubulin of Tissue Culture Cells.

Biochim. Biophys. Acts 629, 359–370

49. Carey, M. F., Peterson, C. L., and Smale, S. T. (2009) Dignam and Roeder

nuclear extract preparation. Cold Spring Harb. Protoc. 4, pdb.prot5330