SCIENCE CHINA
Life Sciences

•RESEARCH PAPER•

https://doi.org/10.1007/s11427-023-2316-2

Specific pupylation as IDEntity reporter (SPIDER) for the identification of protein-biomolecule interactions

He-Wei Jiang[1†], Hong Chen[1†], Yun-Xiao Zheng[1†], Xue-Ning Wang[1†], Qingfeng Meng[1†], Jin Xie[2†], Jiong Zhang[3,4†], ChangSheng Zhang[5], Zhao-Wei Xu[6], Zi-Qing Chen[7], Lei Wang[1], Wei-Sha Kong[1], Kuan Zhou[1], Ming-Liang Ma[1], Hai-Nan Zhang[1], Shu-Juan Guo[1], Jun-Biao Xue[1], Jing-Li Hou[8], Zhe-Yi Liu[9], Wen-Xue Niu[9], Fang-Jun Wang[9], Tao Wang[10], Wei Li[11], Rui-Na Wang[12], Yong-Jun Dang[13], Daniel M. Czajkowsky[14], JianFeng Pei[2*], Jia-Jia Dong[4*] & Sheng-Ce Tao[1*]

[1]*Shanghai Center for Systems Biomedicine, Key Laboratory of Systems Biomedicine (Ministry of Education), Shanghai Jiao Tong University, Shanghai 200240, China;*
[2]*Center for Quantitative Biology, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing 100871, China;*
[3]*Inflammation and Immune Mediated Diseases Laboratory of Anhui Province, School of Pharmacy, Anhui Medical University, Hefei 230032, China;*
[4]*Key Laboratory of Organofluorine Chemistry, Center for Excellence in Molecular Synthesis, Shanghai Institute of Organic Chemistry, University of the Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200240, China;*
[5]*College of Chemistry and Molecular Engineering, Peking University, Beijing 100871, China;*
[6]*Key Laboratory of Gastrointestinal Cancer (Fujian Medical University), Ministry of Education, Fuzhou 350122, China;*
[7]*Department of Molecular Biology, Princeton University, New Jersey 08540, USA;*
[8]*Instrumental Analysis Center, Shanghai Jiao Tong University, Shanghai 200240, China;*
[9]*CAS Key Laboratory of Separation Sciences for Analytical Chemistry, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian 116023, China;*
[10]*Institute of Systems Biology, Shenzhen Bay Laboratory, Peking University Shenzhen Graduate School, Shenzhen 518055, China;*
[11]*Department of Medicinal Chemistry, School of Pharmacy, China Pharmaceutical University, Nanjing 211198, China;*
[12]*Key Laboratory of Metabolism and Molecular Medicine, the Ministry of Education, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai 200240, China;*
[13]*Center for Novel Target and Therapeutic Intervention, Chongqing Medical University, Chongqing 400016, China;*
[14]*School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*

Received January 20, 2023; accepted March 6, 2023; published online April 6, 2023

Protein-biomolecule interactions play pivotal roles in almost all biological processes. For a biomolecule of interest, the identification of the interacting protein(s) is essential. For this need, although many assays are available, highly robust and reliable methods are always desired. By combining a substrate-based proximity labeling activity from the pupylation pathway of *Mycobacterium tuberculosis* and the streptavidin (SA)-biotin system, we developed the Specific Pupylation as IDEntity Reporter (SPIDER) method for identifying protein-biomolecule interactions. Using SPIDER, we validated the interactions between the known binding proteins of protein, DNA, RNA, and small molecule. We successfully applied SPIDER to construct the global protein interactome for m^6A and mRNA, identified a variety of uncharacterized m^6A binding proteins, and validated SRSF7 as a potential m^6A reader. We globally identified the binding proteins for lenalidomide and CobB. Moreover, we identified SARS-

†Contributed equally to this work
*Corresponding authors (Sheng-Ce Tao, email: taosc@sjtu.edu.cn; Jia-Jia Dong, email: jiajia@sioc.ac.cn; JianFeng Pei, email: jfpei@pku.edu.cn)

CoV-2-specific receptors on the cell membrane. Overall, SPIDER is powerful and highly accessible for the study of protein-biomolecule interactions.

proximity labeling, protein-biomolecule interaction, proteomics, pupylation

Citation: Jiang, H.W., Chen, H., Zheng, Y.X., Wang, X.N., Meng, Q., Xie, J., Zhang, J., Zhang, C.S., Xu, Z.W., Chen, Z.Q., et al. (2023). Specific pupylation as IDEntity reporter (SPIDER) for the identification of protein-biomolecule interactions. Sci China Life Sci 66, https://doi.org/10.1007/s11427-023-2316-2

INTRODUCTION

Biomolecules perform their biological functions mainly by interacting with proteins. There is always high interest and necessity to identify protein-biomolecule interactions globally, such as protein-protein interactions (PPIs), protein-nucleic acid interactions (PNIs) and protein-small molecule interactions (PSMIs).

Canonical methods for studying protein-biomolecule interactions include immunoprecipitation (IP) (Ho et al., 2002), GST-pull down (Vikis and Guan, 2004), protein microarray (Zhu et al., 2001), and affinity purification mass spectrometry (AP-MS) (Bürckstümmer et al., 2006; Butter et al., 2009; Edupuganti et al., 2017). Usually, the bindings for all these methods are noncovalent, thus, these methods are suitable for moderate but not stringent washing, easily leading to nonspecific results. In addition, proximity labeling (PL) methods, such as BioID (Roux et al., 2012), APEX (Rhee et al., 2013), and PUP-IT (Liu et al., 2018b), are good choices for discovering molecular interactions, such as PPIs (Mathew et al., 2022) and PNIs in living cells. With PL, the biomolecules of interest need to be fused to an engineered enzyme with proximity-tagging activity, which enables the identification of physiologically relevant molecular interactions, including weak binders. However, these PL methods are not suitable for discovering the interacting proteins of specifically modified biomolecules (e.g., N^6-methyladenosine (m^6A)-modified RNA and posttranslationally modified peptides) and small molecules. In addition, because of the necessity to express the protein of interest in cells, these PL methods are usually challenging for nontransfectable cells, nonmodel organisms and clinical samples. In these scenarios, more applicable choices are *in vitro* assays.

To better capture the dynamic and transient protein-biomolecule interactions *in vitro*, many methods have been developed, e.g., ultraviolet (UV) crosslinking for PNI (Chodosh, 2001) or covalent affinity probes for PSMI (Backus et al., 2016). The addition of an ultraviolet (UV) crosslinking step for PNI improves binding efficiency and is beneficial for the identification of weak and transient binders. However, UV radiation often has very low crosslinking efficiency, necessitating large numbers of cells to be used as input in a single assay (McHugh et al., 2015). Covalent affinity probes (also known as trifunctional affinity probes) have been shown to effectively identify low-abundance targets through additional covalent capture for PSMI. However, small molecules must be equipped with additional photoreactive or electrophilic groups in a case-by-case manner through tedious *de novo* synthesis (Kulkarni et al., 2016; MacKinnon and Taunton, 2009). To address these challenges, there is an urgent need for a general method that can easily, efficiently and specifically convert noncovalent binding into covalent binding *in vitro* (MacKinnon and Taunton, 2009).

Herein, we serendipitously discovered a substrate-based proximity labeling activity from the pupylation pathway of *Mycobacterium tuberculosis*. Taking advantage of this activity, we developed the Specific Pupylation as IDEntity Reporter (SPIDER) as an efficient and widely applicable method for the identification of protein-biomolecule interactions, especially for specifically modified biomolecules and small molecules. We demonstrated the simplicity, feasibility and broad applications of SPIDER by characterizing a variety of challenging protein-biomolecule interactions, from PNI, PSMI, and PPI to ligand-receptor interactions on the cell surface.

RESULTS

Design and validation of the SPIDER proximity-tagging system

In *Mycobacterium tuberculosis*, PafA catalyzes the covalent linkage of the C-terminus of Pup^E to target proteins for proteasomal degradation (Pearce et al., 2008). We serendipitously discovered that PafA can catalyze the covalent binding of Pup^E to its proximal proteins when Pup^E is attached to a protein of interest via its N-terminus, i.e., GFP-Pup^E (Figure S1A in Supporting Information). Self-pupylation of GFP-Pup^E but not Pup^E-GFP was detected by macromolecular mass spectrometry and LC-MS/MS assays (Figure S1A–F in Supporting Information). Structural analysis showed that these pupylation sites are located on the surface of GFP (Figure S1G in Supporting Information). This proximal ligation, which involves the attachment of the substrate (Pup^E) to the protein of interest, is distinct from the known proximity ligation reaction using PafA fused protein (Liu et al., 2018b). Based on this discovery, we designed

SPIDER to identify or validate the protein-molecule interactions (Figure 1A). In a typical SPIDER assay, the interaction between the biotinylated bait and the prey protein is transformed into a covalent linkage between the prey protein and SA_m-Pup^E through enzymatic proximity pupylation. After incubation, without concern about the degradation of the bait and the destruction of protein-biomolecule interactions, especially ligand-receptor interactions on the cell surface, the reaction could be subjected to extremely stringent washes without losing the specific bindings. Because of the covalent linkage, the results of SPIDER could be easily visualized on a gel by simply monitoring the mobility shift of the SA-protein conjugate (Figure 1A).

In the SPIDER assay, Pup^E-fused streptavidin (SA-Pup^E) and PafA (Figure S1H in Supporting Information) are the two key components. To avoid the possible efficiency reduction caused by self-ligation of the SA-Pup^E tetramer or self-pupylation of PafA, the lysine residues on the surface of streptavidin (fused with Pup^E) were mutated to arginine (referred hereafter as SA_m-Pup^E) (Figure S1I in Supporting Information), and all seven lysine residues on the surface of PafA were mutated to arginine (referred hereafter as

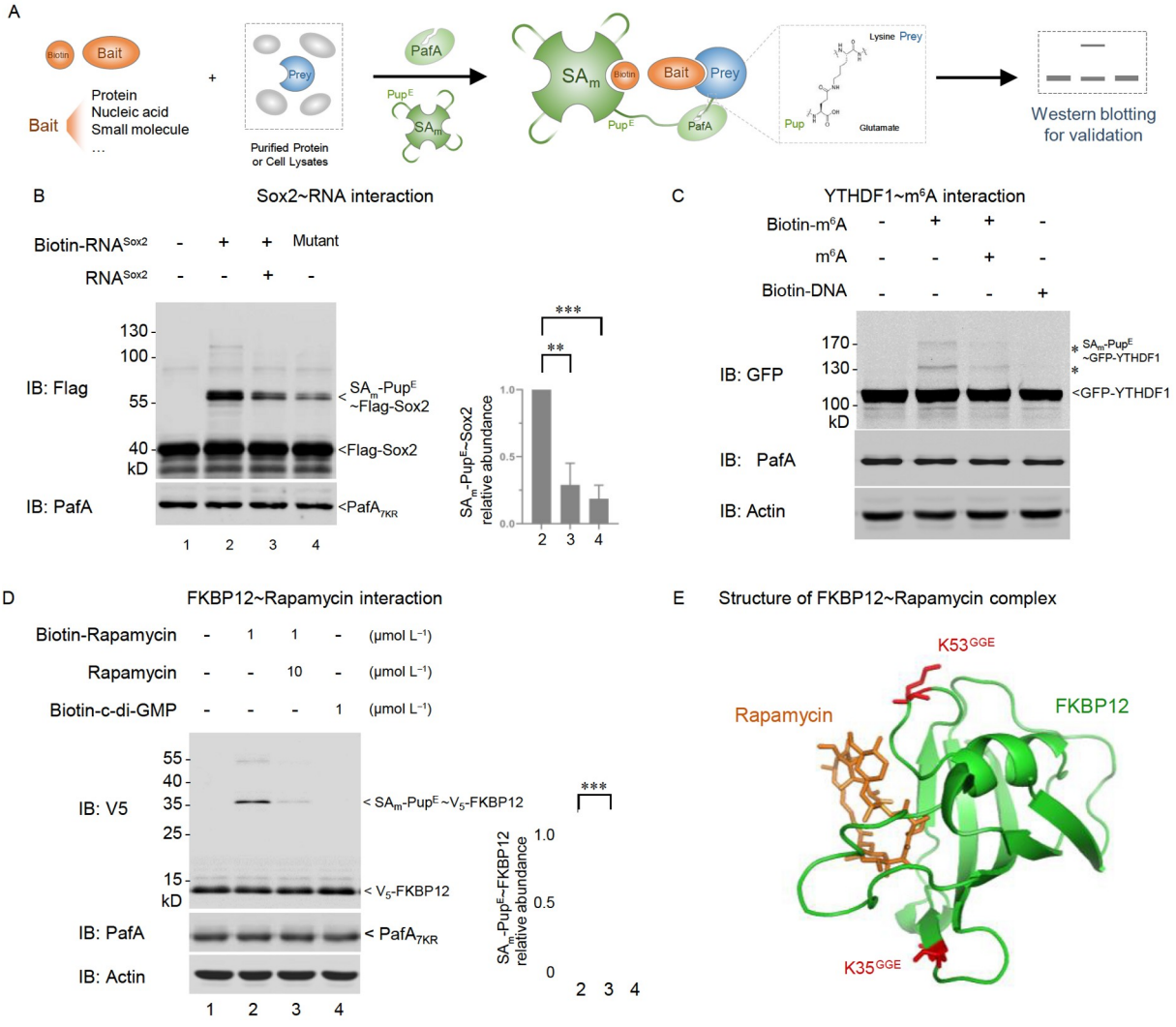

Figure 1 Design and validation of the SPIDER proximity-tagging system. A, Schematic design of the SPIDER. Pup^E fused streptavidin (SA_m-Pup^E) and PafA are the key components of SPIDER. When a biotinylated bait including protein, nucleic acid and small molecules binds to a prey protein, it brings Pup^E to the proximity of the bait through the binding of biotin and SA_m-Pup^E. PafA is recruited and covalently links Pup^E to the adjacent lysine(s) on the prey protein, thus transforming the binding of the biotinylated bait and the prey protein into a covalent linkage between SA_m-Pup^E and the prey protein. The SA_m-Pup^E-prey protein complex is enriched by biotin agarose through stringent washing, e.g., 8 mol L^{-1} urea. The binding of the biotinylated bait and the prey protein is visualized through gel electrophoresis or identified by mass spectrometry. B, SPIDER assay with Flag-tagged Sox2 and specific biotin-RNA. Mutant: Sox2 binding RNA with multiple mutated sites. C, SPIDER assay with the cell lysate of YTHDF1-overexpressing HEK293T cells and biotin-m^6A. D, SPIDER assay with Biotin-Rapamycin and lysate of HEK293T cells. V5-tagged FKBP12 was overexpressed in these cells. B and D, Quantitation of the relative abundance of the major band of SA_m-Pup^E~Prey based on Western blotting is shown in the right panel. Data are the mean±SEM. and representative of three independent experiments, ** P<0.01 and *** P<0.001 (two-tailed unpaired t-test). E, Pupylation sites were mapped on the crystal structure of FKBP12 (PDB: 1FKL).

$PafA_{7KR}$) (Figure S1L in Supporting Information). As expected, SA_m-Pup^E retains comparable biotin-binding activity to that of wild-type (WT) SA and similar stability to WT even in 8 mol L^{-1} urea (Figure S1J and K in Supporting Information). Self-pupylation was sharply reduced for $PafA_{7KR}$ (Figure S1M in Supporting Information), while $PafA_{7KR}$ maintained pupylation activity comparable to that of WT PafA (Figure S1L and N in Supporting Information).

SPIDER proximity-labeling system for validating the protein-biomolecule interactions

Next, we examined whether SPIDER could efficiently validate the protein-biomolecule interactions. We chose the well-studied CheAs and CheZ interaction as a PPI model system (affinity constant (K_d) ~350 nmol L^{-1}) (Figure S2A and B in Supporting Information) (Wang and Matsumura, 1996). We biotinylated CheZ, and as expected, CheAs could indeed be linked to SA_m-Pup^E through binding with biotin-CheZ (Figure S2C in Supporting Information). Mass spectrometry analysis confirmed the success of the SPIDER reaction and showed that SA_m-Pup^E was linked to CheAs at several lysine residues (Figure S2D and E in Supporting Information). It is known that specific mutations of CheAs (L123A, L126A) reduce its affinity with CheZ (Cantwell and Manson, 2009), which we also confirmed (Figure S2B and C in Supporting Information).

To determine whether SPIDER could effectively detect PNIs, we incubated Sox2 (a transcriptional regulator) with its DNA or RNA (Hou et al., 2020) binders and observed the production of SA_m-Pup^E~Sox2 upon the addition of the specific biotin-RNA/DNA (Figure 1B; Figure S2F in Supporting Information). A similar mobility shift was also observed with the SARS-CoV-2 N protein and its interacting RNA (Dinesh et al., 2020) (Figure S2G in Supporting Information). We further found that all the pupylated lysines of both the Sox2 and N proteins were on their surfaces (Figure S2H–J in Supporting Information). Interestingly, the pupylated lysines from Sox2~DNA (Figure S2H in Supporting Information) and Sox2~RNA (Figure S2I in Supporting Information) are different, and all of the modified lysines of the N protein (Figure S2J in Supporting Information) are at its C-terminus. We expect that the pupylated sites on the proteins are dependent on their spatial relation to the nucleic acid binding region on the proteins. Thus, our results suggest that Sox2 has different binding regions for DNA and RNA, which is consistent with a previous study (Hou et al., 2020). Next, we asked whether the SPIDER could identify specific posttranscriptionally modified nucleic acid-binding proteins within a complex environment, such as the cellular milieu. We carried out the SPIDER assay by incubating biotinylated m^6A-modified ssRNAs with the total lysate of YTHDF1-overexpressing HEK293T cells under mild lysis conditions. As expected, gel-shift electrophoresis assays confirmed the formation of SA_m-Pup^E~YTHDF1 directly from the reaction mixture without enrichment, and the binding could be reduced by excessive nonbiotinylated m^6A (Figure 1C).

To determine whether SPIDER could effectively detect PSMI, we detected the interaction between rapamycin and FK-binding protein 12 (FKBP12) by incubating biotin-Rapamycin with the total lysate of FKBP12-overexpressing HEK293T cells. The binding of rapamycin to FKBP12 is a well-known mechanism to suppress the immune system (Itoh and Navia, 1995). As expected, a significant production of SA_m-Pup^E~FKBP12 was observed directly from the reaction mixture without enrichment, while the addition of excessive rapamycin effectively competed off the production of SAm-Pup^E~FKBP12 (Figure 1D). Structural analysis of the FKBP12-rapamycin complex shows that pupylated lysines (K35 and K53) are located on the surface of the complex adjacent to the rapamycin binding site, consistent with our observations of the spatial accessibility of SPIDER-mediated pupylation on the target proteins (Figure 1E).

To further characterize the SPIDER, it is necessary to compare the SPIDER with related technologies, especially technologies based on proximity labeling/capturing through enzymatic reactions or UV crosslinking. These technologies include AP-MS, tri-functional affinity probes, BioID and PUP-IT. We put SPIDER together with these technologies, took the identification of CheZ interacting proteins as an example, performing side-by-side comparisons. We prepared varied versions of the bait protein, i.e., CheZ-birA for BioID, CheZ-PafA for PUP-IT, and Biotin-CheZ for SPIDER, AP-MS and Tri-functional affinity probe. All assays were performed *in vitro* with *E. coli* lysate (Figure S3A and data set S1 in Supporting Information). Judging by the identification of CheA, the known binder of CheZ, higher specificity and better enrichment were observed for SPIDER than for AP-MS, the tri-functional affinity probe and BioID, While SPIDER demonstrated comparable ligation capacity to PUP-IT (Figure S3B–F in Supporting Information). In addition, SPIDER identified 3–4-fold more Pupylation sites than PUP-IT, which indicates that SPIDER has more robust activity for the identification of Pupylation modification sites than PUP-IT; interestingly, 3–4 correlates to 4, the number of independent Pup^E arms on the SA_m-Pup^E tetramer (Figure S3G in Supporting Information).

Taken together, these results demonstrate that SPIDER is indeed highly effective for studying protein-biomolecule interactions.

SPIDER-PNI to identify specific posttranscriptional modification m^6A-binding proteins

As the most prevalent modification in eukaryotic mRNA, m^6A modification plays vital roles in cell differentiation and

tumorigenesis and is interpreted by its readers, such as proteins containing the YTH domain, to regulate the fate of mRNA (Huang et al., 2018; Wang et al., 2014). To identify m^{6}A-binding proteins, we performed SPIDER-PNI and biotinylated RNA probe assays by coincubating biotinylated m^{6}A-modified single-stranded RNA (m^{6}A-ssRNA, with the consensus sequence GG(m^{6}A)CU) with the total lysate of YTHDF family-overexpressing HEK293T cells. Biotinylated ssRNA (A-ssRNA) was used as the control (Figure 2A). MS-based protein identification showed that 20 m^{6}A-interacting proteins were significantly enriched over the control (with a cutoff of >1.5-fold) by SPIDER (Figure 2B and data set S2 in Supporting Information). Among these enriched proteins, five interacting proteins have been reported as known m^{6}A readers, including YTDHF1, YTHDF2, and YTHDF3 scored in the top 3, and IGF2BP2 (Huang et al., 2018) and YTHDC1 (Xiao et al., 2016) ranked in the top list (Figure 2B; data set S2 in Supporting Information). YTHDC1 was not found by the biotinylated RNA probe assay (Figure S4A and data set S2 in Supporting Information). Gene Ontology (GO) analysis showed that more than half of the identified proteins are well connected to the m^{6}A reader signaling pathway and clustered into two categories: regulation of mRNA metabolic process and fidelity maintenance of DNA replication (Figure 2C; Figure S4B in Supporting Information).

In addition to known m^{6}A readers, SPIDER also identified some new binding proteins. Among these proteins, we focused on SRSF7 for further validation. SRSF7 is a known splicing factor that modulates the mRNA export of oncogenes in many cancers and regulates m^{6}A modification on mRNAs related to cell proliferation and migration with an unclear mechanism (Cun et al., 2021). To validate the direct binding between m^{6}A-ssRNA and SRSF7, we performed surface plasmon resonance (SPR) assays. SRSF7 is selectively bound to m^{6}A-ssRNA with a 2-fold higher affinity than the control (Figure 2D; Figure S4C in Supporting Information). To evaluate binding in cells, we overexpressed Flag-SRSF7 in HEK293T cells and performed RNA immunoprecipitation and sequencing (RIP-Seq). A significant enrichment of m^{6}A modification in SRSF7-bound RNA was observed (Figure S4D in Supporting Information). A set of 2,558 transcripts were identified from the Flag-SRSF7-RIP assay, and 3,851 were identified from the m^{6}A-RIP assay by an m^{6}A antibody, of which 584 transcripts overlapped as high-confidence targets of SRSF7 (Figure 2E). GO analysis showed that overlapping SRSF7 target transcripts were enriched in the regulation of the microtubule cytoskeleton, microtubule binding and cytoplasmic microtubule organization, which are critically involved in cell stress responses (Figure S4E in Supporting Information). Further analysis of sequence consistency showed that SRSF7 preferentially binds to the AGGAG(C)AAG consensus sequence in RIP-Seq, similar to the UGGAC sequence in m^{6}A-Seq, which contains the GGAC m^{6}A core motif (Figure 2F; Figure S4F in Supporting Information). Moreover, most of the SRSF7 binding sites are highly enriched in the coding sequence (CDS) and intron, consistent with the known m^{6}A distribution (PMID: 28759256), i.e., near the stop codon and in the 3′ untranslated region (UTR), and 81.38% of SRSF7 binding sites are located in protein-coding transcripts (Figure 2G–I).

Together, these results provide a comprehensive m^{6}A binding map including the potential m^{6}A reader SRSF7, suggesting that SPIDER-PNI is an efficient tool for the identification of binding proteins for modified nucleic acids.

SPIDER-PNI to construct the mRNA-protein interactome of THP-1 cells

mRNA-binding proteins play essential roles in mRNA maturation, localization, degradation and translational regulation (Baltz et al., 2012; Castello et al., 2012). Traditional approaches for systematically characterizing the mRNA binding proteome (mRNA interactome) are mainly based on pull-down assays using oligo(dT)-coated beads, which are highly dependent on low-efficiency cross-linking by UV or chemical reagents; thus, approximately 30 petri dishes of cells are required for a single assay (Baltz et al., 2012; Castello et al., 2012). Here, we investigated whether SPIDER could reveal the mRNA interactome. We used THP-1 cells with phorbol-12-myristate-13-acetate (PMA)-induced differentiation (Schwende et al., 1996) as an example. We carried out the SPIDER assay by incubating biotin-oligo(dT) directly with cell lysates of THP-1 cells before or after PMA induction, prepared under mild lysis conditions (Figure 3A; Figure S5A and B in Supporting Information). A total of 279 mRNA interacting proteins were identified with high reproducibility: 121 proteins for THP-1 cells (PMA-) and 214 proteins for PMA-induced THP-1 cells (PMA+) (Figure 3B; Figure S5C and data set S3 in Supporting Information). To validate the mRNA-binding activities of the identified proteins, we applied a biotinylated oligo(dT) probe assay (PMID: 22681889). We tested the RNA-binding activity of 3 candidates and found that 2 of them, namely, MTA1 and UBE4A, showed clear signals in immunoprecipitations of HEK 293T cells (Figure S5D in Supporting Information), indicating that these proteins are likely in close contact or directly binding to mRNA. MTA1 is a transcriptional coregulator that can act as both a transcriptional corepressor and coactivator (Gururaj et al., 2006), and UBE4A is a ubiquitin-protein ligase that probably functions as an E3 ligase in conjunction with specific E1 and E2 ligases (Hatakeyama et al., 2001). The comparison of our mRNA interactome with two other interactomes generated by traditional oligo(dT) pull-down (Baltz et al., 2012; Castello et al., 2012) assays showed that a variety of new proteins were identified (Figure

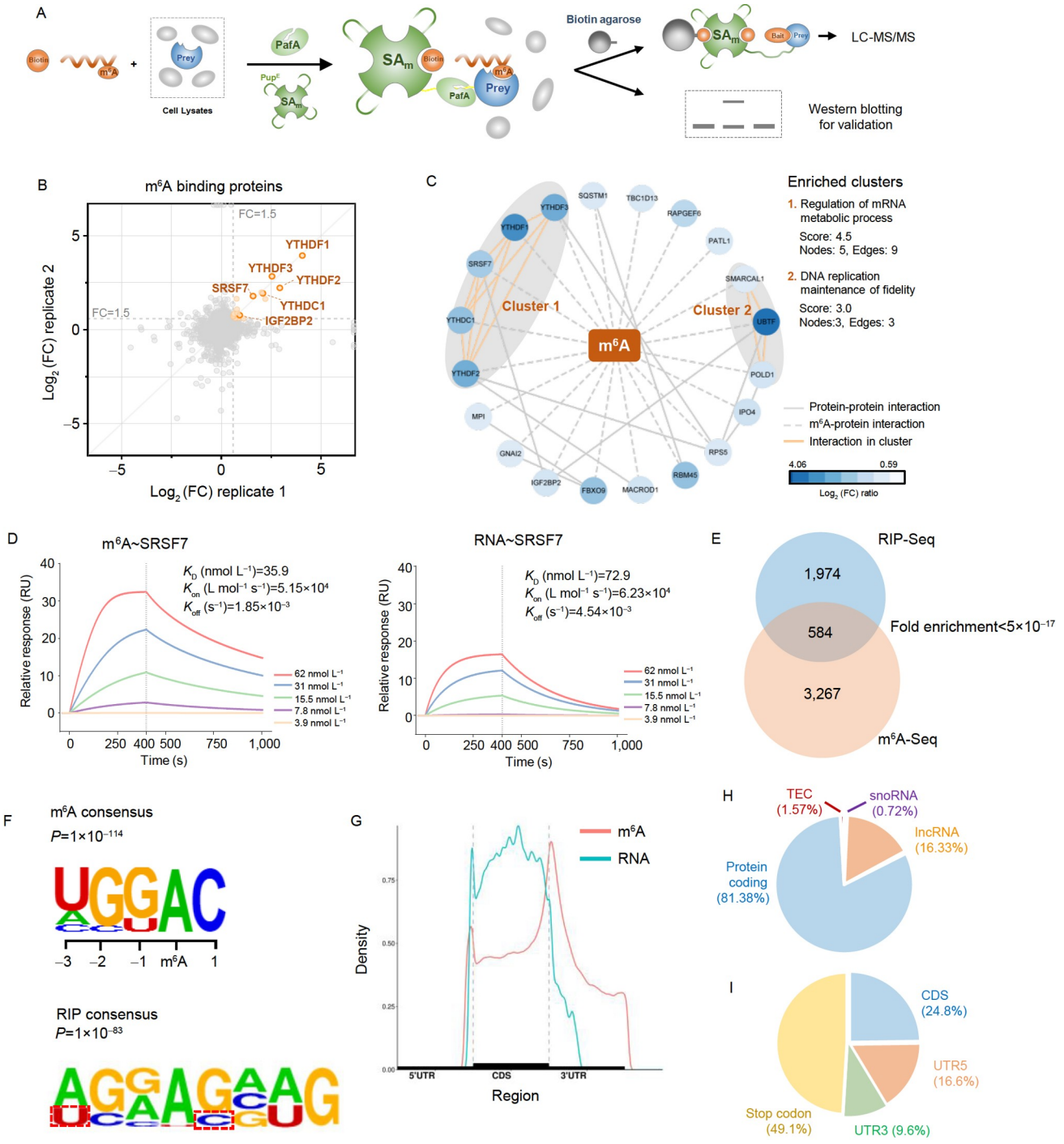

Figure 2 SPIDER-PNI to identify m^{6}A-binding proteins. A, Schematic diagram of SPIDER in capturing m^{6}A-interacting proteins. B, Scatter plots show a comparison of the fold change of two replicate experiments of SPIDER assays with the cell lysates of HEK293T cells overexpressing YTHDF1, YTHDF2 and YTHDF3 and biotin-m^{6}A. The *y*-axis represents the log$_2$-fold change of label-free protein quantification (SPIDER~Biotin-m^{6}A/SPIDER~Biotin-ssRNA). Significantly enriched proteins are shown as orange dots. FC: fold change. C, The interaction network of the m^{6}A-interacting proteins obtained by STRING with a confidence score>0.4 by using the 20 interacting proteins identified by SPIDER as input. The top two and tightly connected network clusters obtained with MCODE are gray color coded. D, The K_ds for m^{6}A-SRSF7 and RNA-SRSF7 were detected by SPR. E, Overlap of SRSF7 target genes identified by RIP-seq and m^{6}A-modified genes identified by m^{6}A-seq. F, Top consensus generated according to the m^{6}A-seq and RIP-seq data. G, Metagene profiles of enrichment of SRSF7-binding sites and m^{6}A modifications across the mRNA transcriptome. H, Percentages of various RNA species bound by SRSF7. lncRNA, long noncoding RNA; snoRNA, small nucleolar RNA; TEC, to be experimentally confirmed. I, The distribution (top) of SFSR7-binding peaks within different gene regions.

3B; data set S4 in Supporting Information). Molecular function analysis showed that SPIDER has a higher enrichment capability with that of other methods when identifying poly(A) and mRNA 3′-UTR binding proteins but not for 5′-

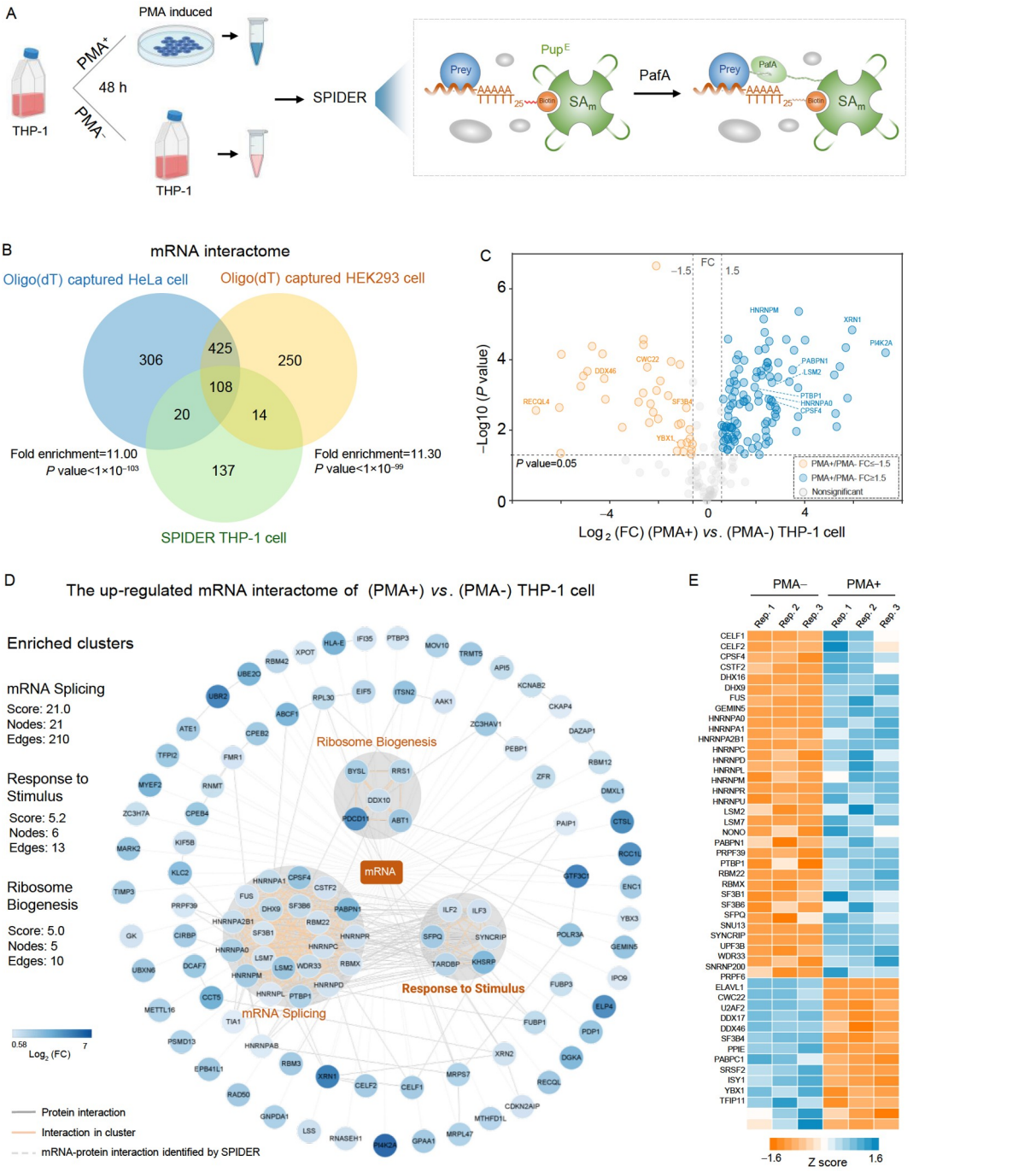

Figure 3 SPIDER-PNI to construct the mRNA-protein interactome of THP-1 cells. A, SPIDER assay for the identification of the mRNA-protein interactome of THP-1 cells before and after PMA-induced differentiation. B, The mRNA-protein interactome of THP-1 cells identified by SPIDER was compared with that of two Oligo(dT) capturing-based studies, i.e., HEK293 (Baltz et al., 2012) and HeLa (Castello et al., 2012). Proteins with FC (fold change) of protein label-free quantification intensity (biotin-oligo(dT)/biotin-control)≥1.5 were defined as (PMA− or PMA+) mRNA interactomes. C, Volcano plot of the mRNA-protein interactome of THP-1 cells before PMA (PMA−) and after PMA-induced differentiation (PMA+). The logarithmic FC of protein label-free quantification intensity (PMA+) *vs*. (PMA−) was plotted against negative logarithmic *P* values (Wilcoxon rank-sum test, False Discovery Rate corrected $P<0.05$) (data set S4 in Supporting Information). Proteins with FC≥1.5 are shown in blue, and those with FC≤−1.5 are shown in orange (three independent experiments). D, The 108 upregulated mRNA interacting proteins of THP-1 cells upon PMA treatment, generated through STRING (confidence score>0.7). The three top-ranked and tightly connected network clusters obtained with MCODE are gray color coded. E, Heatmap of the protein abundance of mRNA splicing factors identified by SPIDER. The *Z* score is measured in terms of standard deviations from the mean. The *Z* score represents the protein difference in abundance, with blue representing a positive score (0 to 1.6) and upregulated, and orange representing a negative score (−1.6 to 0) and downregulated.

UTR binding proteins, which could be explained by the spatial range of SPIDER (Figure S5E and data set S4 in Supporting Information). In PMA-treated THP-1 cells, we found that 108 proteins were significantly upregulated, while

39 proteins were significantly downregulated (Figure 3C; data set S4 in Supporting Information). Most of the proteins identified in THP-1 cells were highly related (Figure 3D; Figure S5F and G in Supporting Information). The upregulated mRNA-protein interactome of PMA-treated cells can be roughly divided into three groups: mRNA splicing, ribosome biogenesis and response to stimulus (Figure 3D). Notably, consistent with a THP-1 differentiation model (Liu et al., 2018a), a total of 46 splicing factors were identified in the THP-1 mRNA interactome, revealing a close connection between dynamic spliceosome function and THP-1 differentiation (Figure 3E; Figure S5G in Supporting Information). Among them, PTBP1 has been proven to play critical roles in mediating alternative splicing during neuronal differentiation (Linares et al., 2015), suggesting that it may also play an essential role in the differentiation of THP-1 cells (Figure 3D and E; Figure S5G in Supporting Information). Thus, these results show that SPIDER could be applied for the construction of mRNA-protein interactomes in a highly efficient manner.

SPIDER-PSMI to capture the interacting proteins of lenalidomide

Lenalidomide is well known for its effectiveness in treating hematological malignancies via its interaction with the E3 ligase DDB1-CRBN complex (Fischer et al., 2014). We carried out a SPIDER-PSMI assay by incubating biotin-lenalidomide with HEK293T cell lysate, followed by immunoprecipitation and identification using biotin agarose coupled with MS, with biotin as a control (Figure 4A; Figure S6A in Supporting Information). As expected, the DDB1-CRBN complex was highly enriched by biotin-lenalidomide (Figure 4B; data set S5 in Supporting Information). We found that the pupylated lysines were all located on the protein surface within close proximity to the binding interface, reflecting sites of greatest access to the SPIDER reaction (Figure 4C and D; Figure S6B and C in Supporting Information). In addition, besides CRBN and DDB1, SPIDER also identified other interacting proteins for lenalidomide (Figure 4E; data set S5 in Supporting Information). GO analysis showed that more than half of the identified proteins were well connected to the CRBN-DDB1 signaling pathway and clustered into the following categories: nucleotide-excision repair, mitochondrial translation and Golgi vesicle transport (Figure 4E; Figure S6D in Supporting Information). These binding proteins could serve as a resource to explore other therapeutic applications of lenalidomide or to explain the known side effects (Chen et al., 2017). Collectively, these results show that the SPIDER can identify target proteins for small molecules or drugs within a complex environment, such as the cellular milieu.

SPIDER-PPI to identify enzyme-substrate interactions

To expand the application of the SPIDER in detecting transient interactions within a cellular milieu, such as PPI, we identified the interactome of the *E. coli* protein deacetylase CobB (Starai et al., 2002). CobB, as the only member of the Sir2 family of deacetylases in *E. coli*, is known to play a role in many different pathways (Castaño-Cerezo et al., 2014; Weinert et al., 2013), although its complete range of interactors is still incompletely known. The SPIDER assay was performed by incubating biotin-CobB with stable isotope-labeled *E. coli* lysate prepared under mild lysis conditions. A total of 84 interacting proteins were identified (Figure S7A, B and data set S6 in Supporting Information). Notably, while half of the interacting proteins in our dataset were reported to be potential protein substrates of CobB (Castaño-Cerezo et al., 2014; Weinert et al., 2013), we identified a variety of novel interactors (Figure S7A, B and data set S6 in Supporting Information). These interacting proteins are involved in 22 pathways, including gene expression that has also been found in previous studies (Castaño-Cerezo et al., 2014; Weinert et al., 2013) (Figure S7C in Supporting Information). To verify these findings, we randomly selected 8 proteins for bio-layer interferometry (BLI) assays, and 5 of the 8 proteins were confirmed to interact with CobB, namely, RoxA, RraA, GldA, DksA and VacB (Figure S7D–I in Supporting Information). Next, deacetylation assays clearly showed that both VacB and DksA could be acetylated in *E. coli* and that acetylation could be efficiently removed by CobB (Figure S7J and K in Supporting Information). VacB (RNase R) is a 3′–5′ exoribonuclease whose activity is known to be regulated by acetylation (Cheng and Deutscher, 2002; Liang et al., 2011), whereas DksA is a critical component of the transcription initiation machinery (Paul et al., 2004). Our results thus imply that CobB plays a previously unappreciated role in regulating these processes. Collectively, these results show that SPIDER is indeed capable of identifying enzyme-substrate interactions within a complex environment, such as cell lysate.

SPIDER to identify ligand-receptor interactions

In comparison to protein-biomolecule interactions among soluble proteins, the study of protein-biomolecule interactions that involve membrane proteins such as ligand-receptor interactions is considerably more challenging due to the difficulty of isolating membrane proteins while maintaining their conformations and activities.

To test the capability of the SPIDER to identify cell-surface receptors (Figure 5A), we first examined the interaction between the SARS-CoV-2 S1 protein and its well-known receptor, ACE2. SPIDER reactions were carried out directly on cell culture in petri dishes. We observed no change in cell

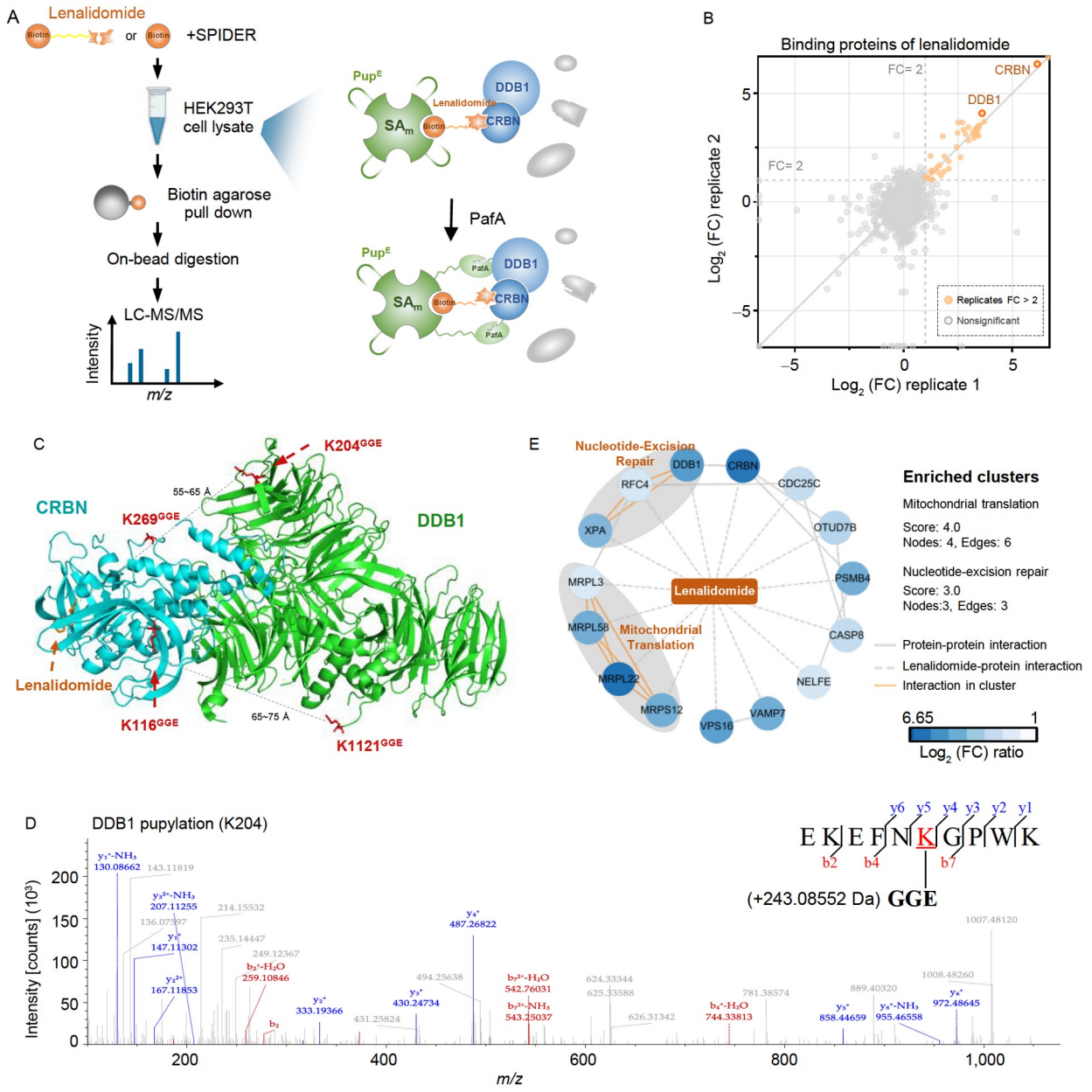

Figure 4 SPIDER-PSMI to capture the interacting proteins of lenalidomide. A, SPIDER assays for capturing the binding proteins of biotinylated lenalidomide. B, Scatter plots show comparison of fold changes (FC) of two replicate experiments of SPIDER~biotin-lenalidomide. The *y*-axis represents the $\log_2$ (FC) of SPIDER~Biotin-Lenalidomide/SPIDER~Biotin. Significantly enriched proteins are shown as orange dots. C, Pupylation sites were mapped on the structure of the lenalidomide~CRBN-DDB1 complex (PDB: 4CI2). Pupylated lysines are highlighted in red. D, The pupylation site (K204) on DDB1 after SPIDER assay identified by LC-MS/MS analysis. E, The interaction network of the lenalidomide-interacting proteins obtained using STRING (confidence score>0.4) by using the 15 interacting proteins identified by SPIDER as input. The two top-ranked and tightly connected network clusters obtained with MCODE are gray color coded.

mortality before and after the SPIDER reaction (Figure S8A in Supporting Information). The binding between SARS-CoV-2 S1 and ACE2 was confirmed by showing the mobility shift of ACE2-SA_m-Pup^E (Figure 5B).

Next, we coupled SPIDER with MS to identify the SARS-CoV-2 RBD binding proteins on the surface of H1299, Calu-3 and Vero E6 cells (Figure 5C; Figure S8B in Supporting Information). H1299 and Calu-3 cells represent low and high expression of ACE2, respectively. We identified 55 interactors for SARS-CoV-2 RBD (Figure 5D; data set S7 in Supporting Information), including several of the known receptor/coreceptors, e.g., ACE2 and Vimentin (Suprewicz et al., 2022). To further confirm ligand-receptor binding, based on the MS data, pupylated lysine sites were identified (Figure S8C and D in Supporting Information). The pupylated site on the surface of ACE2 is accessible to the SPIDER

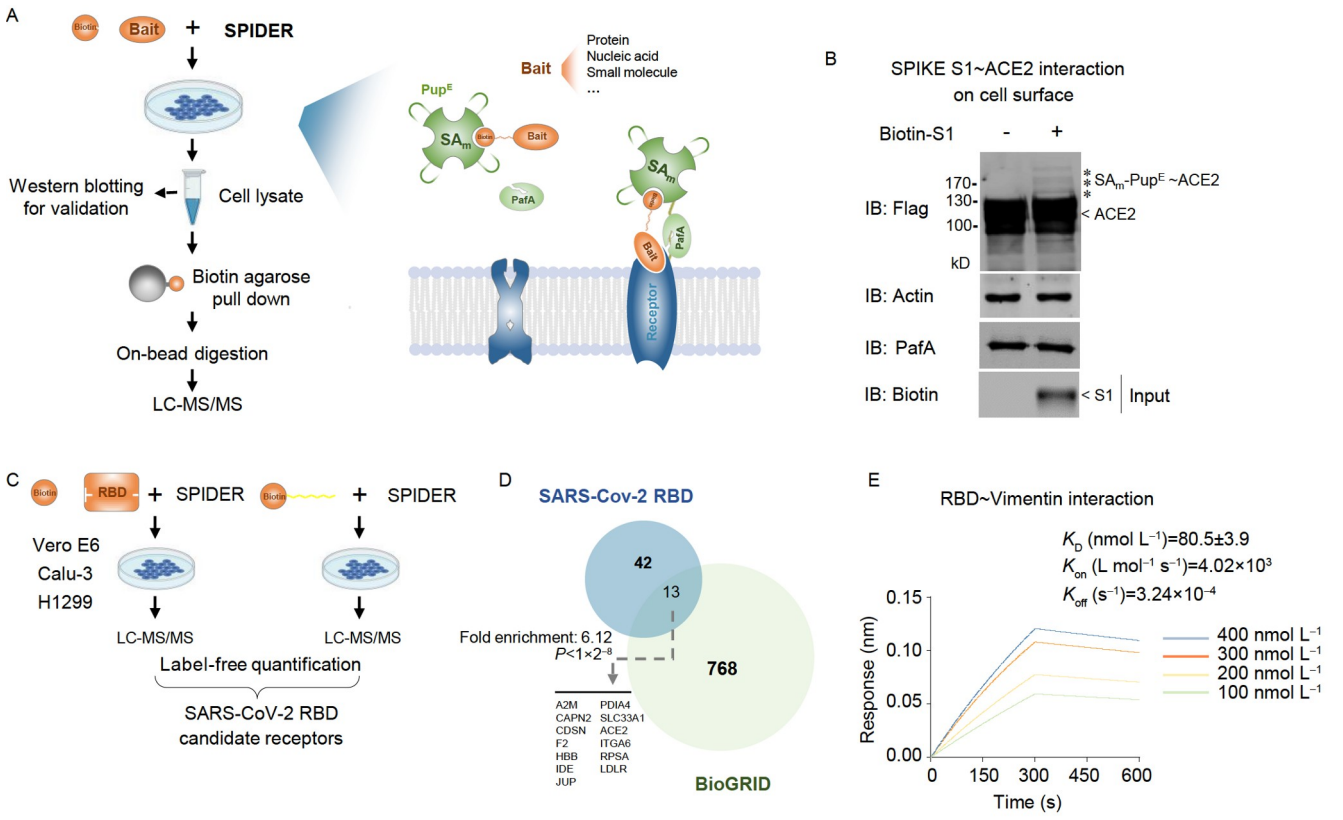

Figure 5 SPIDER to identify the SARS-CoV-2 receptor on the cell surface. A, SPIDER assays for identifying the binding proteins for biotinylated bait on the cell surface. B, SPIDER assay with biotin-S1 and lysate of HEK293T cells. Flag-tagged ACE2 was overexpressed in these cells. The asterisks represent the bands of SA_m-Pup^E~ACE2. C, SPIDER assays for the identification of biotinylated RBD binding proteins on the cell surface. The prey protein(s) was identified by mass spectrometry. D, The SARS-CoV-2 RBD interactors identified by SPIDER were compared with known interactors in the BioGRID database. E, Binding affinity of the SARS-CoV-2 RBD~Vimentin.

reaction (Figure S8E in Supporting Information). We compared these identified RBD interactors and the interactors in the BioGRID database and found that 13 interactors, e.g., ACE2, ITGA6, RPSA and LDLR, overlapped (Figure 5D; data set S7 in Supporting Information).

To further explore the functional roles of the potential receptors, we analyzed the high confidence interactors of SARS-CoV-2 RBD in regard to their protein interaction network (Figure S9A in Supporting Information). We identified 3 major cell processes of the interactors of the SARS-CoV-2 RBD, e.g., focal adhesion (Figure S9B in Supporting Information). Among these proteins, Vimentin is highly conserved within species (Figure S9B in Supporting Information). It is a major component of class III intermediate filaments and is also known to localize on the extracellular surface of the plasma membrane as a coreceptor participating in SARS-CoV-2 infection (Suprewicz et al., 2022). We performed BLI assays and confirmed the binding of the SARS-CoV-2 RBD and SARS-CoV-2 S1 to Vimentin (Figure 5E; Figure S9C in Supporting Information).

Taken together, these results show that SPIDER can efficiently identify membrane-localized receptors.

DISCUSSION

Although there are many methods for studying protein-biomolecule interactions, a method of high generality, robustness, reliability and accessibility is still needed.

Based on a mechanism whereby the covalent linkage occurs when the substrate (Pup^E) is proximal to protein(s), SPIDER has been developed to discover binding proteins of a variety of biomolecules, for example, RNA with specific modifications, e.g., m^6A. To our knowledge, SPIDER is the only substrate-based proximity labeling system, whose proximal covalent ligation is distinct from the proximity ligation reaction using PafA fused protein, i.e., PUP-IT. The distinguishing proximity labeling feature of SPIDER mainly relies on the disordered structure of Pup in the absence of cofactor binding (Liao et al., 2009), which provides a large degree of flexibility of potential linking configurations to the prey protein, thus holding powerful scalability for the validation and identification of a wide range of protein-biomolecule interactions coupled with the streptavidin-biotin system.

It is necessary to assess the specificity of the SPIDER. However, because the major application of SPIDER is for discovery-related studies, it is difficult to obtain an ideal

example (with a set of well-validated binders and non-binders) to accurately assess the specificity of the SPIDER assay. A more practically applicable way is to estimate the specificity by looking for how many known binders were successfully "reidentified". Bore this in mind, we have demonstrated the specificity by several examples, e.g., the enrichment of rapamycin binding protein FKBP12 from the cell lysate. In addition, for the identification of the interactome of the *E. coli* protein deacetylase CobB, 5 of the 8 proteins could be validated. In another assay, we tested the RNA-binding activity of 3 candidates and found that 2 of them could be validated by an independent assay. Together, these results indicate that the specificity of SPIDER is acceptable.

Additionally, SPIDER shows several other notable advantages. First, SPIDER is generally applicable for biomolecules, requiring only the biotinylation of the molecule of interest, which could be easily achieved by synthesis or labeling *in vitro* or *in vivo* (Figure S10A in Supporting Information). Due to the small size of the biotin molecule, compared with other methods, such as trifunctional affinity probes or enzyme-fused bait proximity labeling, biotin labeling in SPIDER may have much less interference. Second, SPIDER shows the same characteristics of simplicity and easy operation as AP-MS but has stronger robustness (Figure S10A and B in Supporting Information). It is not necessary to worry about some common limitations as frequently observed in other approaches, such as the degradation of the bait (especially RNA) and the easy disruption of protein-biomolecule interactions. After incubation, the samples could be subjected to extremely stringent washes without losing the specific binding. In addition, because of the covalent linkage, the results of SPIDER could be readily visualized on a gel by monitoring the mobility shift of the SA-protein conjugate. This feature is very suitable for fast *in vitro* validation of protein-biomolecule interactions. Last, because the two key reagents of SPIDER (PafA and SA_m-Pup^E) could be standardized and premade in large quantities and lyophilized (Figures S10A and S11 in Supporting Information), when applying SPIDER, the researcher can simply prepare the biotinylated bait and the intact cells/cell lysate, when necessary, even the intact cells/cell lysate could be pre-made (Figure S10A in Supporting Information). Thus, high-throughput SPIDER assays are possible to investigate the interaction map of a large number of different baits simultaneously.

Nonetheless, there are a few limitations of the SPIDER. First, nonspecific binding is also inevitable. This limitation can be largely overcome by adopting the SILAC (stable isotope labeling of amino acids) strategy and further validation *in vivo*. Second, in its current form, SPIDER is for *in vitro* analysis and is not applicable to the discovery of spatial interactions *in vivo*. Third, for SPIDER~PSMI, a major limitation is to prepare biotinylated small molecules and retain their designated activity: typically, a small molecule is biotinylated through a preferred chemical group, often, the chemistry is challenging, or the activity is impaired after biotinylation. This is indeed a general challenge for small molecule related studies (but not for protein, DNA, RNA and peptide), and not just for SPIDER. To overcome this challenge, one possible solution is to adopt a conjugate chemistry which is active for a variety of chemical groups (Gomes and Gozzo, 2010). In this way, a small molecule could be "shotgun" biotinylated at multiple sites, and so by controlling the molar ratio of conjugate reagent to the small molecule, it would be possible to link one biotin per small molecule on varied sites. By doing this, we may lower the difficulty of the chemistry and retain the designated activity for at least a portion of the biotinylated small molecules.

To conclude, SPIDER is a substrate-based proximity labeling system. Because of the enzymatic catalytic nature of converting noncovalent to covalent interactions, it enables efficient and specific identification and validation of biomolecule-protein interactions, especially weak, transient and membrane-localized interactions, as long as the biomolecule can be biotinylated.

MATERIALS AND METHODS

Plasmid construction and cloning

The protein sequences were downloaded from GenBank. The corresponding DNA sequences were codon-optimized and synthesized by Sangon Biotech (Shanghai, China). The synthesized genes were cloned into pET28a or pGEX-4T-1 for prokaryotic expression or pcDNA3.1 for eukaryotic expression. $PafA_{7KR}$ was constructed based on wild-type PafA using a QuickChange® Site-Directed Mutagenesis Kit (Agilent Technologies, USA) and cloned into pTrc99a. The DNA sequences of streptavidin and Pup were fused together through PCR. The PCR product (SA-Pup) was cloned into pET28a. Proteins were expressed in *E. coli* BL21(DE3). N- or C-terminal His and Avi (GLNDIFEAQKIEWHE)-tagged protein of interest (POIs) were cloned into pET32a, and the pET32a plasmids were cotransformed with pET28a carrying BirA into *E. coli* BL21(DE3). All biotin-DNA was synthesized by Sangon Biotech, and all biotin-RNA was synthesized by GenScript (Nanjing, China). Nucleic acid sequences and detailed information (protein tags, protein expression vectors, etc.) of the clones constructed in this study can be found in data set S8 in Supporting Information.

Protein expression and purification

E. coli BL21, carrying the expression plasmid, was cultured in LB medium at 37°C until the A_{600} reached 0.6. Protein

expression was induced with 0.5 mmol L^{-1} isopropyl-β-D-thiogalactoside (IPTG) at 18°C overnight or 37°C for 4 h. For the purification of 6×His-tagged proteins, cell pellets were resuspended in lysis buffer containing 50 mmol L^{-1} Tris-HCl pH 8.0, 500 mmol L^{-1} NaCl, and 20 mmol L^{-1} imidazole (pH 8.0) and then lysed by a high-pressure cell cracker (Union-Biotech, Shanghai, China). Cell lysates were centrifuged at 10,000×*g* for 20 min at 4°C. Supernatants were collected and purified with Ni^{2+} Sepharose beads (Smart-lifesciences, SA003100) or Ni-NTA Agarose (QIAGEN, USA), washed with lysis buffer (for the purification of biotinylated protein, 1 mmol L^{-1} biotin was added to the wash buffer) and eluted with buffer containing 50 mmol L^{-1} Tris-HCl pH 8.0, 500 mmol L^{-1} NaCl and 300 mmol L^{-1} imidazole (pH 8.0). For the purification of GST-tagged proteins, cells were harvested and lysed by a high-pressure cell cracker in lysis buffer containing 50 mmol L^{-1} Tris-HCl, pH 8.0, 500 mmol L^{-1} NaCl, and 1 mmol L^{-1} DTT. After centrifugation, the supernatant was incubated with GST-Sepharose beads (Smart-Lifesciences, Changzhou, China) or Glutathione Sepharose 4 Fast Flow (Cytiva, Sweden). The target proteins were washed with lysis buffer twice and eluted with 50 mmol L^{-1} Tris-HCl, pH 8.0, 500 mmol L^{-1} NaCl, 1 mmol L^{-1} DTT, and 40 mmol L^{-1} glutathione.

To recover soluble SA_m-Pup^E for inclusion bodies, the insoluble fraction cultured from 200 mL LB medium was washed four times with 50 mmol L^{-1} Tris HCl, pH 8.0, 50 mmol L^{-1} EDTA, 100 mmol L^{-1} NaCl, 1 mol L^{-1} urea, and 1% Triton X-100. The inclusion body was dissolved in 10 mL of 50 mmol L^{-1} Tris-HCl, pH 8.0, 8 mol L^{-1} urea, 15 mmol L^{-1} DTT, 1 mmol L^{-1} EDTA and dialyzed against 2 L of 20 mmol L^{-1} Tris-HCl, 0.8 mol L^{-1} urea, 2 mmol L^{-1} L-arginine, pH 10, for 24-48 h and then dialyzed against 2 L of 150 mmol L^{-1} NaCl, 50 mmol L^{-1} Tris-HCl, pH 7.5, and 20 mmol L^{-1} $MgCl_2$ for another 24 h. The dialysate was centrifuged and concentrated in a stirred ultrafiltration cell (Amicon, USA).

Cell culture and transfection

HEK293T cells, Calu-3 cells, Vero E6 cells, and H1299 cells were cultured in DMEM (Corning, USA) by adding 10% fetal bovine serum (FBS) (Invitrogen, USA) and 1% penicillin-streptomycin (Invitrogen) at 37°C in a 5% CO_2 incubator. During passaging, cells were digested with 0.25% trypsin. Before transfection, cells were passaged by 1:3–1:5 and cultured for 12–16 h. The cells were transfected using Lipofectamine™ 2000 transfection reagent (Invitrogen) according to the manufacturer's instructions. Briefly, in a 10-cm petri dish, 10 μg plasmid and 25 μL Lipofectamine™ 2000 were added. After transfection, the cells were cultured for 36–48 h.

THP-1 cells (human acute monocytic leukemia cells) were purchased from Cellbank (Shanghai, China). THP-1 cells were incubated at 37°C with 5% CO_2 and cultured in RPMI 1640 (Corning) supplemented with 0.05 mmol L^{-1} β-mercaptoethanol and 10% FBS. For differentiation to a macrophage phenotype, RPMI 1640 medium supplemented with 0.3% BSA was used, and THP-1 cells were treated with 100 ng mL^{-1} PMA (phorbol-12-myristate-13-acetate) (Sigma-Aldrich, USA) for 24 h. Then, PMA-containing media was replaced with fresh media, and nonadherent cells were removed. Then, the cells were rested in culture for another 24 h. To obtain the differentiated cell lysate, the medium was removed and washed once in PBS buffer, and M-PER Reagent (Thermo Fisher Scientific, USA) was added, followed by 5 min of gentle shaking. The cell images were taken by light microscopy to record the differentiation of monocytes to macrophages.

Cell lysate preparation

For adherent cells, cells were washed once in PBS buffer. Then, 1 mL M-PER Reagent (Thermo Fisher Scientific) with 0.5 mmol L^{-1} PMSF was added to a 10 cm petri dish, followed by 5 min of gentle shaking to obtain the cell lysate. For suspended cells, cells were counted using a cell counting chamber. The cell suspension was centrifuged at 800 r min^{-1} for 3 min. The cell pellet was washed with PBS and centrifuged at 800 r min^{-1} for 3 min. The supernatant was carefully removed. M-PER Reagent (1 mL) with 0.5 mmol L^{-1} PMSF was added to 1 mL/10^7 cells and vortexed briefly to obtain a homogeneous cell suspension. The cell suspension was incubated for 40 min at 4°C with a 10-level speed (the highest speed) and vortexed every 10 min. The lysate was collected and transferred to a microcentrifuge tube and centrifuged at 16,000×*g* for 5–10 min to pellet the cell debris. The supernatant was transferred to a new tube and stored at −80°C for future use.

SPIDER assay for PPI validation

The reaction was composed of 1 μmol L^{-1} PafA or $PafA_{7KR}$, 0.1–2.5 μmol L^{-1} biotin-POI, 1–5 μmol L^{-1} purified prey protein, 10 μmol L^{-1} SA_m-Pup^E and 10 mmol L^{-1} ATP in reaction buffer. The reaction buffer included 150 mmol L^{-1} NaCl, 50 mmol L^{-1} Tris-HCl, pH 7.5, 20 mmol L^{-1} $MgCl_2$, and 10% glycerol. After gentle mixing, the reaction was carried out at 30°C for 4 h. Tips: it will be hard to differentiate on the gel when the MWs are similar between the bait and the prey. This problem could be solved by fusing a tag to the bait protein.

SPIDER assay for PNI validation

The reaction was composed of 1 μmol L^{-1} $PafA_{7KR}$,

0.25–2.5 μmol L^{-1} biotin-nucleic acid, 1–5 μmol L^{-1} purified prey protein, 0.05 μg μL^{-1} poly d(I-C), 10 μmol L^{-1} SA_m-Pup^E and 10 mmol L^{-1} ATP in reaction buffer. Ten-fold unlabeled nucleic acid as a competitive probe was added to the reaction mixture when necessary. The reaction buffer included 150 mmol L^{-1} NaCl, 50 mmol L^{-1} Tris-HCl, pH 7.5, 20 mmol L^{-1} $MgCl_2$, and 10% glycerol. After gentle mixing, the reaction was carried out at 30°C for 4 h.

SPIDER assay for PPI screening in cell lysates

For the SPIDER reaction, there was 3 mg of cell lysate, 0.1–2.5 μmol L^{-1} biotinylated POI or biotin control and 0.5 mmol L^{-1} PMSF, and the reaction buffer was composed of 50 mmol L^{-1} Tris-HCl pH 7.5, 150 mmol L^{-1} NaCl, 20 mmol L^{-1} $MgCl_2$, and 10% (v/v) glycerol. The reaction was carried out on a rotating wheel at room temperature for 15 min, followed by the addition of 10 mmol L^{-1} ATP, 10 μmol L^{-1} SA-Pup and 1 μmol L^{-1} PafA and incubation on a rotating wheel at 30°C for 4–6 h. A 10–30 μL sample was saved for Western blot analysis. Urea powder was added to the sample to a final concentration of 8 mol L^{-1} and incubated at 30°C for 2 min. A total of 150 μL biotin-agarose (Sigma-Aldrich) was added to the reaction and incubated at 4°C overnight or room temperature for 2 h to capture the pupylation-linked interacting proteins. The biotin-agarose was collected and stored at −80°C for mass spectrometry analysis.

For SILAC cell lysate, SPIDER assays were performed as described above with the "heavy" (heavy isotope labeled) reaction containing 0.1–2.5 μmol L^{-1} biotinylated POI and the "light" reaction containing biotin control. After incubation, the lysate was denatured by the addition of urea to a final concentration of 8 mol L^{-1}. The "heavy" and "light" reactions were then combined, and the mixture was incubated with 300 μL biotin-agarose (Sigma-Aldrich) at 4°C overnight or room temperature for 2 h to capture the pupylation-linked interacting proteins. The biotin-agarose was collected and stored at −80°C for mass spectrometry analysis.

Affinity purification for protein-protein interactions

For affinity purification, the reaction was composed of 1 μmol L^{-1} biotin-POI and 3 mg cell lysate in PBS buffer. After gentle mixing, the reaction was carried out at room temperature for 1 h. Streptavidin agarose (Sangon Biotech) was added to the cell lysate and incubated overnight to capture the biotinylated protein interactors. After washing three times with PBS, streptavidin agarose was collected and stored at −80°C for mass spectrometry analysis.

Tri-functional affinity probe

To label the bait protein with Sulfo-SBED (Pierce, USA), Sulfo-SBED dissolved in DMSO was added to desalting bait protein at a 1:1 molar ratio in Label Transfer buffer containing 50 mmol L^{-1} HEPES and 150 mmol L^{-1} NaCl (pH 7.5). The label reaction was carried out in the dark at room temperature for 30 min. To remove the nonreacted cross-linker, the labeling reaction was desalted or dialyzed. To quantify biotinylation, the Biotin Quantitation kit (Thermo Fisher Scientific) can be used or dot blot can be performed using IRDye 800CW (Li-COR, USA) as the probe.

For the cross-linking reaction, the reaction was composed of 1 μmol L^{-1} Sulfo-SBED labeled bait protein and 3 mg cell lysis in PBS buffer. After incubation at room temperature for 1 h, the reaction mixture was photoactivated for 15 min with UV light at 300–370 nm on ice. Streptavidin agarose (Sangon Biotech) was added to the reaction and incubated overnight to capture the crosslinked bait-interactors. After washing three times with PBS, the cross-linked bait-interactors were broken with 100 mmol L^{-1} DTT at room temperature for 1 h. Streptavidin agarose was collected and stored at −80°C for mass spectrometry analysis.

BioID and Pup-IT reaction

For the BioID reaction (Roux et al., 2012), the reaction was composed of cell lysate, 1 μmol L^{-1} BirA(R118G) or POI-BirA(R118G), 50 μmol L^{-1} biotin, and 5 mmol L^{-1} ATP in reaction buffer. For the PUP-IT reaction (Liu et al., 2018b), the reaction was composed of cell lysate, 1 μmol L^{-1} PafA or POI-PafA, 10 μmol L^{-1} biotin-Pup^E, and 5 mmol L^{-1} ATP in reaction buffer. The reaction buffer included 150 mmol L^{-1} NaCl, 50 mmol L^{-1} Tris-HCl, pH 7.5, and 20 mmol L^{-1} $MgCl_2$. After gentle mixing, the reaction was carried out at 37°C for 1 h. Then, urea powder was added to the reaction to a final concentration of 8 mol L^{-1}. The lysate was incubated on a rotating wheel at room temperature until urea dissolved. After centrifugation of the lysate, Streptavidin agarose (Sangon Biotech) was added to the supernatant and incubated overnight to capture the biotinylated proteins. Following washing three times with 8 mol L^{-1} urea, streptavidin agarose was collected and stored at −80°C for mass spectrometry analysis.

SPIDER assay for PNIs screening in cell lysate

For the SPIDER reaction, there was 3 mg cell lysate (approximately 2×10^7–3×10^7 cells), 1 μmol L^{-1} m^6A ssRNA or control for the m^6A binding protein assay or 2 μmol L^{-1} Oligo(dT) for the mRNA interactome assay, 0.05 μg μL^{-1} poly d(I-C) and 0.5 mmol L^{-1} PMSF, and the reaction buffer was composed of 50 mmol L^{-1} Tris-HCl pH 7.5, 150 mmol L^{-1} NaCl, 20 mmol L^{-1} $MgCl_2$, and 10% (v/v) glycerol. For the identification of m^6A binding proteins and THP-1-cell mRNA interactome assays, additional RNase

inhibitors were added to the cell lysate to a final concentration of 10 U mL^{-1}. The reaction was carried out on a rotating wheel at room temperature for 15 min, followed by the addition of 10 mmol L^{-1} ATP, 10 μmol L^{-1} SA_m-Pup^E and 1 μmol L^{-1} PafA and incubation on a rotating wheel at 30°C for 4–6 h. A 10–30 μL sample was saved for Western blot analysis. Urea powder was added to the sample to a final concentration of 8 mol L^{-1} and incubated at 30°C for 2 min. A total of 150 μL biotin-agarose (Sigma-Aldrich) was added to the reaction and incubated at 4°C overnight or room temperature for 2 h to capture the pupylation-linked interacting proteins. The biotin-agarose was collected and stored at −80°C for mass spectrometry analysis.

Biotinylated nucleic acid probe assay for protein-nucleic acid interaction

For affinity purification to capture protein-nucleic acid interactions, the cells were cross-linked at 254 nm UV with 0.15 J cm^{-2} for 30 min on ice followed by lysis with NP-40 buffer containing RNase inhibitor (0.5 U mL^{-1}) (Yeasen, Shanghai, China). For the identification of the mRNA interactome, the reaction was composed of 5 μmol L^{-1} biotin-oligo d(T) and 6 mg cell lysate in lysis buffer. For the identification of the m^6A interactome, 2 μmol L^{-1} biotin-m^6A or biotin-RNA and 6 mg of cell lysate were added to lysis buffer. After gentle mixing, the reaction was carried out at room temperature for 1 h. Streptavidin agarose (Sangon Biotech) was added to the cell lysate and incubated overnight at 4°C for m^6A or 1 h at room temperature for oligo d(T) to capture the protein-interactors. Streptavidin agarose was washed two times in PBS buffer. For mass spectrometry analysis, nucleic acid was digested by incubation with nuclease P1 (10 U μg^{-1} RNA) (NEB, USA) for 30 min at 37°C in PBS buffer. For Western blot analysis, protein-mRNA complexes were heat eluted from beads in 10 mmol L^{-1} Tris HCl (pH 7.5) for 2 min at 80°C followed by nucleic acid digestion as before.

SPIDER assay for PSMI screening in cell lysate

In a 3 mL reaction, there was 3 mg cell lysate, 1–2 μmol L^{-1} biotin-small molecule or biotin control, and 0.5 mmol L^{-1} PMSF. The reaction buffer was composed of 50 mmol L^{-1} Tris-HCl pH 7.5, 150 mmol L^{-1} NaCl, 20 mmol L^{-1} $MgCl_2$, and 10% (v/v) glycerol. The reaction was carried out on a rotating wheel at room temperature for 15 min, followed by the addition of 10 mmol L^{-1} ATP, 10 μmol L^{-1} SA_m-Pup^E and 1 μmol L^{-1} PafA and incubation on a rotating wheel at 30°C for 4–6 h. A 10–30 μL sample was saved for Western blotting analysis. Urea powder was added to the sample to a final concentration of 8 mol L^{-1} and incubated at 30°C for 2 min. A total of 150 μL biotin-agarose (Sigma-Aldrich) was added to the reaction and incubated at 4°C overnight or room temperature for 2 h to capture the pupylation-linked interacting proteins. The biotin-agarose was collected and stored at −80°C for mass spectrometry analysis.

SPIDER reaction on the cell surface

The adherent cells (HEK293T, Vero E6, H1299, Calu-3) were cultured in petri dishes (15 cm in diameter). After the removal of the medium, the cells were washed gently with the reaction buffer. Biotinylated POI (0.2–1 μmol L^{-1}), SPIKE RBD (ACROBiosystems, Beijing, China), nucleic acid and 10 μmol L^{-1} biotin-hexapeptide (synthesized by Sangon Biotech), or the control (free biotin only) were incubated with the cells at room temperature for 15 min in the reaction buffer. The reaction buffer contained 50 mmol L^{-1} Tris-HCl pH 7.5, 150 mmol L^{-1} NaCl, 20 mmol L^{-1} $MgCl_2$, and 10% (v/v) glycerol. Then, 10 μmol L^{-1} SA_m-PupE, 10 mmol L^{-1} ATP and 1 μmol L^{-1} PafA were added to the reaction system and incubated at 30°C for 2 h. The reaction system was removed, and the cells were lysed by radioimmunoprecipitation assay buffer (RIPA) (Beyotime Biotechnology, Shanghai, China) for 40 min with a 10-level speed (the highest speed) vortex every 10 min. A total of 10–30 μL of sample was saved for Western blot analysis. Then, urea powder was added to the reaction to a final concentration of 8 mol L^{-1}. The lysate was incubated on a rotating wheel at room temperature until urea dissolved. After centrifugation of the cell lysate, 150 μL of biotin agarose (Sigma-Aldrich) was added to the supernatant and incubated at 4°C overnight or room temperature for 2 h to capture the pupylation-linked interacting proteins on the cell surface. The biotin-agarose was collected and stored at −80°C for mass spectrometry analysis.

Mass spectrometry analysis of intact protein

The mass of GFP-Pup proteins from the experimental and control groups was determined by Thermo Exactive Plus EMR mass spectrometry coupled to an Agilent 1100 HPLC system. The proteins were first buffer exchanged into 0.1% formic acid with ultrafiltration. Then, 0.5 μg of protein was loaded onto a 5 cm×200 μm i.d. trap column (C5, 5 μm, 300 Å, Phenomenex, USA) and separated by a 15 cm×150 μm i.d. analytical column (C5, 5 μm, 300 Å, Phenomenex) with isocratic elution (50% acetonitrile 0.1% formic acid). The mass spectra were collected with a resolution of 17,500 in an Orbitrap analyzer. The EMR mode was enabled. The AGC (automatic gain control) was set at 1×10^6 with an injection time of 100 ms. Finally, the mass spectra were deconvoluted with Thermo Protein Deconvolution 4.0 to obtain the exact mass of the proteins.

Mass spectrometry and data analysis

Proteins enriched by biotin-agarose were reduced by adding DTT to a final concentration of 10 mmol L^{-1} and incubated at 37°C for 1 h. Subsequently, alkylation was performed by adding iodoacetamide to a final concentration of 25 mmol L^{-1} and incubated in the dark for 20 min. Proteins on the biotin-agarose were then digested with trypsin (1:30 protein-to-enzyme ratio) at 37°C overnight. The biotin-agarose was rinsed twice with 200 μL 50 mmol L^{-1} NH_4HCO_3. All the supernatant, including the resulting peptides, was collected and desalted using a MonoSpin C18 care desalting column (GL Science, Japan) according to the manufacturer's instructions.

The tryptic peptide digests of the proteins were analyzed with an EASY-nL 1200 system coupled online to a Q Exactive Plus Mass spectrometer (Thermo Fisher Scientific). The peptide sequences were determined by searching MS/MS spectra against the Protein database using the Proteome Discoverer (version 2.4) software suite with a precursor ion mass tolerance of 10 ppm and fragment ion mass tolerance of 0.02 Da. Carbamidomethyl (C) was set as the fixed modification, and oxidation (M), GGE (K) and deamidated (NQ) were set as the variable modification. The search results were automatically processed at an FDR of 1% for both the protein and peptide. The unique peptide included in the protein group was used for quantification. All proteins were identified with more than one unique peptide. Label-free quantification was used to quantify the difference in protein abundance between different samples (Cox and Mann, 2008).

For SILAC MS data analysis, raw MS spectra were processed by using Proteome Discoverer 2.4 software. The SILAC 2plex (Arg10Lys6) method was selected for quantification analysis. The following search parameters were employed: full tryptic specificity was needed, and two missed cleavages were allowed. Carbamidomethylation was set as a fixed modification, whereas oxidation (M), deamidation (NQ), acetylation (N-terminus) and GGE (K) were considered variable modifications. The precursor ion mass tolerances were 10 ppm for all MS spectra acquired, and the fragment ion mass tolerance was 0.02 Da for all MS2 spectra. The search results were automatically processed at an FDR of 1% at both the protein and peptide levels. The unique peptide included in the protein group was used for quantification. Proteins with a SILAC ratio greater than or equal to 2 are considered candidate interacting proteins.

Functional analysis of the proteins identified in SPIDER

To analyze protein-protein interactions, the protein list was uploaded to the STRING (Szklarczyk et al., 2019) database (https://string-db.org). The list of protein interactions was imported into Cytoscape (Shannon et al., 2003) (Version 3.8.2) for network presentation. The plugin of MCODE was used for network clustering, in which the node score cutoff was set by 0.2 and K-Core was set by 2. GO analysis was performed by PANTHER online tools (http://www.pantherdb.org, February 2021). GO term enrichment analysis was performed using Fisher's exact tests (false discovery rate corrected $P<0.05$, minimum two-fold enrichment) using the annotations of interactors. The statistical significance for the overrepresentation test was determined using Fisher's exact test. Volcano plots of proteins identified in this study were generated using MetaboAnalyst 5.0 (https://www.metaboanalyst.ca/) ($P<0.05$, fold change$\geqslant$1.5). The protein expression score of Omicron RBD interactors in human tissues was obtained from the Human Protein Atlas database (Uhlén et al., 2015) (www.proteinatlas.org). Protein expression of Omicron RBD interactors in cells from human tissues was analyzed using a single-cell mRNA sequencing dataset (Han et al., 2020) (http://bis.zju.edu.cn/HCL/).

Structure analysis and sequence alignment

The nucleic acid and protein structures were obtained from the PDB database (www.rcsb.org). Structural analysis was carried out by PyMOL (Version 2.5.0) at the default setting. The sequence of Vimentin in different species was obtained from UniProt (www.uniprot.org). Sequence alignment was performed by SnapGene (6.0).

Biolayer Interferometry

To measure the binding kinetics, biotinylated protein was loaded at 25 ng μL^{-1} in kinetics buffer containing 1×PBS with 0.1% BSA and 0.02% Tween 20 onto streptavidin biosensors (ForteBio, USA). The association of candidate proteins was tested in kinetics buffer at gradient concentrations for 3–5 min. Dissociation in kinetics buffer was measured for 3–5 min. BLI assays were carried out in 96-well black plates and analyzed on OctetRed96 (Fortebio) equipment. Mean K_{on}, K_{off}, and K_d values were calculated by a 1:1 global fit model using the Fortebio Data Analysis software 8.0. The curves were processed using Prism software (GraphPad Prism 8.0.0).

Surface plasmon resonance

The K_ds for m6A-SRSF7 and RNA-SRSF7 were detected by SPR experiments using a Biacore 8K instrument (GE Healthcare, USA). Biotin-m6A, biotin-RNA and SRSF7 protein were diluted in HBS-P buffer, which was also used as a running and analyte-binding buffer. HBS-P buffer was composed of HBS and 0.3% v/v surfactant P20. The ligand and analytes were injected at a flow rate of 30 μL min^{-1}.

Immobilization involved binding of the ligands, 5 nmol L^{-1} biotin-m^6A or RNA oligo, in the SA sensor chip (Cytiva) after activation of NaCl/NaOH and blockage of isopropyl alcohol plus active buffer. During analysis, the association and dissociation times were set at 400 and 600 s, respectively. SRSF7 protein (analyte) at 3–62.5 nmol L^{-1} flowed over the immobilized-ligand surface, and the binding response was recorded. The chip surface was regenerated by the removal of analyte with a regeneration buffer glycine (pH 3.0). The results were analyzed by Biacore Insight Evaluation Software (v3.0.12). The multicycle kinetics evaluation method and the kinetics fit model of 1:1 binding were applied to fit the curve.

Immunofluorescence staining

HEK293T cells were seeded on poly-L-lysine (Beyotime Biotechnology)-coated coverslips and transfected with ACE2-Flag, PD-1-Flag, Flag-CCDC25 and Flag-tagged empty vector individually for 48 h. The cells were incubated with a wheat germ agglutinin (WGA) conjugate (Thermo Fisher Scientific) at 5 μg mL^{-1} for 15 min away from light at room temperature. After washing with PBS, the cells were fixed with 4% paraformaldehyde for 15 min and blocked with 5% BSA in PBS for 30 min at room temperature. Cells were immunostained with an anti-Flag antibody (Sigma-Aldrich) (1:100) overnight at 4°C, followed by incubation with Alexa-488 conjugated anti-mouse antibody (Thermo Fisher Scientific) (1:1,000) for 1 h. The coverslips were mounted on microscope slides using Antifade Mounting Medium (Beyotime Biotechnology). Images were acquired with a confocal microscope (Leica SP8 STED).

Flow cytometry

Cell suspensions were incubated in PBS with 2% FBS. Cells (1×10^6) were stained with PE anti-human CD11b antibodies (20 μL per test, BD Pharmingen, USA) and APC-Cy7 anti-human CD14 antibodies (5 μL per test, BD Pharmingen) for 30 min on ice. After staining with antibodies, the cells were washed with PBS (2% FBS) twice by centrifugation at 600 r min^{-1} for 1 min. A total of ~1×10^5 events were collected for each sample with a BD LSRFortessa system (BD Biosciences, USA), and FlowJo software 7.6.1 was used for data processing.

RIP-Seq

RIP assays were performed on Huh7 cells by LC-Bio Technology Co., Ltd. (Hangzhou, China). Briefly, Huh7 cells were lysed with cell lysis buffer. The 10% lysis sample was used for input, and 90% was used for immunoprecipitation with anti-SRSF7 antibody (Abcam, USA). TRIzol reagent was used for the extraction of RNA. The stranded RNA sequencing library was generated by a Stranded mRNA Library Prep Kit for Illumina based on the manufacturer's instructions. The duplication bias in PCR and sequencing steps is eliminated by using a unique molecular identifier (UMI) of 8 random bases to label the preamplified cDNA molecules. The library products corresponding to 200–500 bp were enriched, quantified and finally sequenced on an Illumina Novaseq™ 6000 with the PE150 model.

RIP-Seq data analysis

Data analysis was conducted by LC-Bio Technology Co., Ltd. Raw sequencing data were first filtered by Trimmomatic (version 0.36), low-quality reads were discarded, and the reads contaminated with adaptor sequences were trimmed. Clean reads were further treated with in-house scripts to eliminate duplication bias introduced in library preparation and sequencing. In brief, clean reads are first clustered according to the UMI sequences, in which reads with the same UMI sequence are grouped into the same cluster. Reads in the same cluster are compared to each other by pairwise alignment, and then reads with sequence identity over 95% are extracted to a new subcluster. After all subclusters are generated, multiple sequence alignment is performed to obtain one consensus sequence for each subcluster. After these steps, any errors and biases introduced by PCR amplification or sequencing are eliminated. The deduplicated consensus sequences were used for protein binding site analysis. They were mapped to the reference genome of *Mus musculus* (GRCm38) using STAR software (version 2.5.3a) with default parameters. RSeQC (version 2.6) was used for read distribution analysis. ExomePeak (Version 3.8) software was used for peak calling. Peaks were annotated using bedtools (Version 2.25.0). DeepTools (version 2.4.1) was used for peak distribution analysis. The differentially binding peaks were identified by a python script using Fisher's test. Sequence motifs enriched in peak regions were identified using Homer (version 4.10). GO analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for annotated genes were both implemented by KOBAS software (version: 2.1.1) with a corrected *P*-value cutoff of 0.05 to judge statistically significant enrichment.

m^6A-seq

Total RNA was isolated and purified using TRIzol reagent (Invitrogen). Approximately more than 25 μg of total RNA representing a specific adipose type was used to deplete ribosomal RNA with an Epicenter Ribo-Zero Gold Kit (Illumina, USA). Following purification, the ribosomal-depleted RNA was fragmented into small pieces using a Magnesium RNA Fragmentation Module (NEB) at 86°C for 7 min. Then,

the cleaved RNA fragments were incubated for 2 h at 4°C with m^6A-specific antibody (Synaptic Systems, Germany) in IP buffer (50 mmol L^{-1} Tris-HCl, 750 mmol L^{-1} NaCl and 0.5% Igepal CA-630). Then, the IP RNA was reverse-transcribed to create cDNA by SuperScript™ II Reverse Transcriptase (Invitrogen), which was used to synthesize U-labeled second-stranded DNAs with *E. coli* DNA polymerase I (NEB), RNase H (NEB) and dUTP Solution (Thermo Fisher Scientific). An A-base is then added to the blunt ends of each strand, preparing them for ligation to the indexed adapters. Each adapter contains a T-base overhang for ligating the adapter to the A-tailed fragmented DNA. Single- or dual-index adapters were ligated to the fragments, and size selection was performed with AMPureXP beads. After heat-labile UDG enzyme (NEB) treatment of the U-labeled second-stranded DNAs, the ligated products were amplified with PCR by the following conditions: initial denaturation at 95°C for 3 min; 8 cycles of denaturation at 98°C for 15 s, annealing at 60°C for 15 s, and extension at 72°C for 30 s; and then final extension at 72°C for 5 min. The average insert size for the final cDNA library was (300±50) bp. Finally, we performed 2×150 bp paired-end sequencing (PE150) on an Illumina NovaSeq™ 6000 (LC-Bio Technology Co., Ltd.).

m^6A-Seq data analysis

fastp software (https://github.com/OpenGene/fastp) was used to remove the reads that contained adaptor contamination, low-quality bases and undetermined bases with default parameters. The sequence quality of the IP and input samples was verified using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and RseQC (http://rseqc.sourceforge.net/). Then, we used HISAT2 (http://daehwankimlab.github.io/hisat2) to map reads to the reference genome *Homo sapiens* (Version: v101). Peak calling and diff peak analysis were performed by the R package exomePeak2 (https://bioconductor.org/packages/release/bioc/html/exomePeak2.html), and peaks were annotated by intersection with gene architecture using the R package ANNOVAR (http://www.openbioinformatics.org/annovar/). MEME (http://meme-suite.org) and HOMER (http://homer.ucsd.edu/homer/motif) were used for de novo and known motif finding followed by localization of the motif with respect to the peak summit. StringTie (https://ccb.jhu.edu/software/stringtie) was used to determine the expression levels of all transcripts and genes from the input libraries by calculating FPKM (total exon fragments/mapped reads (millions)×exon length (kb)). The differentially expressed transcripts and genes were selected with $\log_2$(fold change)$\geqslant$1 or $\log_2$(fold change)$\leqslant$−1 and P value<0.05 by the R package edgeR (https://bioconductor.org/packages/edgeR).

RIP

SRSF7 was transiently overexpressed in three 15-cm dishes of confluent HEK293T cells. Cells grown in 15-cm dishes were crosslinked by UV on ice and lysed with RIP buffer (150 mmol L^{-1} KCl, 10 mmol L^{-1} HEPES (pH 7.4), 5 mmol L^{-1} EDTA, 0.5 mmol L^{-1} DTT, 0.5% NP40, 1× protease inhibitor, 400 U mL^{-1} RNase inhibitor). Cell lysis was precleared with magnetic beads for 2 h at 4°C. One microgram of SRSF7 antibody (Abcam) or control antibody rabbit IgG (Millipore, USA) was conjugated to protein A/G magnetic beads (Thermo Fisher Scientific) by incubation for 4 h at 4°C, followed by washing three times with ish buffer (200 mmol L^{-1} NaCl, 50 mmol L^{-1} HEPES (pH 7.4), 2 mmol L^{-1} EDTA, 0.5 mmol L^{-1} DTT, 0.05% NP40, 200 U mL^{-1} RNase inhibitor). Conjugated beads were incubated with precleared cell lysis in RIP buffer at 4°C overnight. After washing with buffer twice, the beads were resuspended in 800 μL PBS. The beads were digested by DNase RQ1 (Promega, USA) at 37°C for 30 min and incubated with proteinase K (NEB) at 37°C for 30 min. The input and coimmunoprecipitated RNAs were recovered by TRIzol (Invitrogen) extraction.

RIP-MS

Input, flow-through and SRSF7-bound RNA were purified with TRIzol reagent (Invitrogen). mRNA from the three portions was further purified by a Dynabeads mRNA Purification Kit (Thermo Fisher Scientific). A total of 50–100 ng of mRNA was digested by nuclease P1 (NEB) and alkaline phosphatase CIP (NEB) at 37°C for 30 min. The sample was diluted to 50 μL and filtered (0.22 μm pore size, 4 mm diameter, Millipore), and 5 μL of the solution was injected into the LC-MS/MS instrument. Nucleosides were separated by reversed-phase ultra-performance liquid chromatography and mass spectrometry detection using an UHPLC 30 A/AB Sciex Quadrupole 5500 LC mass spectrometer in positive electrospray ionization mode. The nucleosides were quantified by using retention time and nucleoside-to-base ion mass transitions of 282.1 to 150.1 (m^6A) and 268 to 136 (A) (Huang et al., 2018). Quantification was performed in comparison with the standard curve of pure nucleoside standards (m^6A (Selleck, USA), adenosine (Selleck)) at 0.05 to 200 ng mL^{-1}. The ratio of m^6 A to A is calculated based on the calibrated concentrations.

Lyophilization (freeze-drying) of SPIDER

The SPIDER sample was lyophilized using trehalose as a protective agent. Each 1.5 mL screw tube was composed of 0.825 mg PafA, 2.16 mg SA_m-Pup^E and 165.3 mg ATP·Na2 in reaction buffer. Trehalose (30%) in water (wt/v) was

mixed into tubes to a final concentration of 50 mg mL^{-1}. Lyophilization was performed on these tubes using the SCIENTZ-30F Vacuum Function Big LCD Display Heating Function Freeze Dryer (Ningbo Scientz Biotechnology, Ningbo, China). Samples were frozen at −40°C for 6 h, followed by a drying process under vacuum according to the following cycle: −30°C for 10 h, −20°C for 10 h, −10°C for 10 h, 0°C for 8 h, 10°C for 6 h, and 20°C for 5 h. After that, the lyophilized samples were stored at −20°C until use. Before use, the lyophilized SPIDER sample was resuspended in 3 mL reaction buffer containing 0.1–2.5 μmol L^{-1} biotin-POI and 1–5 μmol L^{-1} purified prey protein. The following step is as described previously.

Data and materials availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the iProX (Ma et al., 2019) partner repositor with dataset identifiers as follows:
PXD026509: Pupylation sites on CheA identified by SPIDER assay;
PXD026527: Intramolecular pupylation of GFP-Pup^E identified by mass spectrometry;
PXD026478: Pupylation sites on protein GFP-Pup^E;
PXD026511: Pupylation sites on protein FKBP12;
PXD037952: Side-by-side comparison of AP-MS (PXD037953), Tri-functional affinity probe (PXD037952), BioID (PXD037952), PUP-IT (PXD037952) and SPIDER (PXD037952);
PXD026514: CobB interacting proteins identified by SPIDER assay;
PXD026517: Pupylation sites on protein Sox2;
PXD026519: m^6A binding proteins identified by the SPIDER assay;
PXD026521: mRNA-protein interactome of THP-1 cells;
PXD037881: m^6A binding proteins identified by biotinylated m^6A probe assay;
PXD026518: Pupylation sites on SARS-CoV-2 N;
PXD031035: SARS-CoV-2 receptor on the cell surface identified by SPIDER assay;
PXD026523: Lenalidomide binding proteins identified by SPIDER assay.

Additional data related to this paper may be requested from the authors.

Compliance and ethics *The author(s) declare that they have no conflict of interest.*

Acknowledgements *This work was supported by the National Key Research and Development Program of China (2020YFE0202200), the National Natural Science Foundation of China (31900112, 21907065, 31970130 and 31670831). We thank Dr. Heng Zhu for his long-term guidance. We thank Prof. Pilong Li of School of Life Sciences of Tsinghua University for kindly providing the YTHDF family plasmids. We thank Prof. Xichen Bao of Guangzhou Institutes of Biomedicine and Health of Chinese Academy of Sciences for kindly providing the Sox2 plasmid. We thank Prof. Xiaodong Zhao of Shanghai Jiao Tong University and Prof. Jian Yang of Tongji University for critical comments.*

References

Backus, K.M., Correia, B.E., Lum, K.M., Forli, S., Horning, B.D., González-Páez, G.E., Chatterjee, S., Lanning, B.R., Teijaro, J.R., Olson, A.J., et al. (2016). Proteome-wide covalent ligand discovery in native biological systems. Nature 534, 570–574.
Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., et al. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell 46, 674–690.
Bürckstümmer, T., Bennett, K.L., Preradovic, A., Schütze, G., Hantschel, O., Superti-Furga, G., and Bauch, A. (2006). An efficient tandem affinity purification procedure for interaction proteomics in mammalian cells. Nat Methods 3, 1013–1019.
Butter, F., Scheibe, M., Mörl, M., and Mann, M. (2009). Unbiased RNA-protein interaction screen by quantitative proteomics. Proc Natl Acad Sci USA 106, 10626–10631.
Cantwell, B.J., and Manson, M.D. (2009). Protein domains and residues involved in the CheZ/$CheA_S$ interaction. J Bacteriol 191, 5838–5841.
Castaño-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sánchez-Díaz, N.C., Sauer, U., Heck, A.J., Altelaar, A.M., and Cánovas, M. (2014). Protein acetylation affects acetate metabolism, motility and acid stress response in *Escherichia coli*. Mol Syst Biol 10, 762.
Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–1406.
Chen, N., Zhou, S., and Palmisano, M. (2017). Clinical pharmacokinetics and pharmacodynamics of lenalidomide. Clin Pharmacokinet 56, 139–152.
Cheng, Z.F., and Deutscher, M.P. (2002). Purification and characterization of the *Escherichia coli* exoribonuclease RNase R—comparison with RNase II. J Biol Chem 277, 21624–21629.
Chodosh, L.A. (2001). UV crosslinking of proteins to nucleic acids. Curr Protoc Mol Biol 36.
Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26, 1367–1372.
Cun, Y., An, S., Zheng, H., Lan, J., Chen, W., Luo, W., Yao, C., Li, X., Huang, X., Sun, X., et al. (2021). Specific regulation of m^6A by SRSF7 promotes the progression of glioblastoma. Genomics Proteomics Bioinformatics doi: 10.1016/j.gpb.2021.11.001.
Dinesh, D.C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., and Boura, E. (2020). Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. PloS Pathog 16, e100910.
Edupuganti, R.R., Geiger, S., Lindeboom, R.G.H., Shi, H., Hsu, P.J., Lu, Z., Wang, S.Y., Baltissen, M.P.A., Jansen, P.W.T.C., Rossa, M., et al. (2017). N^6-methyladenosine (m^6A) recruits and repels proteins to regulate mRNA homeostasis. Nat Struct Mol Biol 24, 870–878.
Fischer, E.S., Bohm, K., and Thoma, N.H. (2014). Structure of the DDB1-CRBN E3 ubiquitin ligase in complex with thalidomide: insights into CRL4 inhibition by small molecules. FEBS J 281, 248–249.
Gomes, A.F., and Gozzo, F.C. (2010). Chemical cross-linking with a diazirine photoactivatable cross-linker investigated by MALDI- and ESI-MS/MS. J Mass Spectrom 45, 892–899.
Gururaj, A.E., Singh, R.R., Rayala, S.K., Holm, C., den Hollander, P., Zhang, H., Balasenthil, S., Talukder, A.H., Landberg, G., and Kumar, R. (2006). MTA1, a transcriptional activator of breast cancer amplified sequence 3. Proc Natl Acad Sci USA 103, 6670–6675.
Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J.,

Tang, H., Ge, W., et al. (2020). Construction of a human cell landscape at single-cell level. Nature 581, 303–309.
Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N., and Nakayama, K.I. (2001). U box proteins as a new family of ubiquitin-protein ligases. J Biol Chem 276, 33111–33120.
Ho, Y., Gruhler, A., Heilbut, A., Bader, G.D., Moore, L., Adams, S.L., Millar, A., Taylor, P., Bennett, K., Boutilier, K., et al. (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. Nature 415, 180–183.
Hou, L., Wei, Y., Lin, Y., Wang, X., Lai, Y., Yin, M., Chen, Y., Guo, X., Wu, S., Zhu, Y., et al. (2020). Concurrent binding to DNA and RNA facilitates the pluripotency reprogramming activity of Sox2. Nucleic Acids Res 48, 3869–3887.
Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol 20, 285–295.
Itoh, S., and Navia, M.A. (1995). Structure comparison of native and mutant human recombinant FKBP12 complexes with the immunosuppressant drug FK506 (tacrolimus). Protein Sci 4, 2261–2268.
Kulkarni, P.M., Kulkarni, A.R., Korde, A., Tichkule, R.B., Laprairie, R.B., Denovan-Wright, E.M., Zhou, H., Janero, D.R., Zvonok, N., Makriyannis, A., et al. (2016). Novel electrophilic and photoaffinity covalent probes for mapping the cannabinoid 1 receptor allosteric site (s). J Med Chem 59, 44–60.
Liang, W., Malhotra, A., and Deutscher, M.P. (2011). Acetylation regulates the stability of a bacterial protein: growth stage-dependent modification of RNase R. Mol Cell 44, 160–166.
Liao, S., Shang, Q., Zhang, X., Zhang, J., Xu, C., and Tu, X. (2009). Pup, a prokaryotic ubiquitin-like protein, is an intrinsically disordered protein. Biochem J 422, 207–215.
Linares, A.J., Lin, C.H., Damianov, A., Adams, K.L., Novitch, B.G., and Black, D.L. (2015). The splicing regulator PTBP1 controls the activity of the transcription factor Pbx1 during neuronal differentiation. eLife 4, e09268.
Liu, H., Lorenzini, P.A., Zhang, F., Xu, S., Wong, M.S.M., Zheng, J., and Roca, X. (2018a). Alternative splicing analysis in human monocytes and macrophages reveals MBNL1 as major regulator. Nucleic Acids Res 46, 6069–6086.
Liu, Q., Zheng, J., Sun, W., Huo, Y., Zhang, L., Hao, P., Wang, H., and Zhuang, M. (2018b). A proximity-tagging system to identify membrane protein-protein interactions. Nat Methods 15, 715–722.
Ma, J., Chen, T., Wu, S., Yang, C., Bai, M., Shu, K., Li, K., Zhang, G., Jin, Z., He, F., et al. (2019). iProX: an integrated proteome resource. Nucleic Acids Res 47, D1211–D1217.
MacKinnon, A.L., and Taunton, J. (2009). Target identification by diazirine photo-cross-linking and click chemistry. Curr Protoc Chem Biol 1, 55–73.
Mathew, B., Bathla, S., Williams, K.R., and Nairn, A.C. (2022). Deciphering spatial protein-protein interactions in brain using proximity labeling. Mol Cell Proteomics 21, 100422.
McHugh, C.A., Chen, C.K., Chow, A., Surka, C.F., Tran, C., McDonel, P., Pandya-Jones, A., Blanco, M., Burghard, C., Moradian, A., et al. (2015). The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. Nature 521, 232–236.
Paul, B.J., Barker, M.M., Ross, W., Schneider, D.A., Webb, C., Foster, J. W., and Gourse, R.L. (2004). DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. Cell 118, 311–322.
Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., and Darwin, K.H. (2008). Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. Science 322, 1104–1107.
Rhee, H.W., Zou, P., Udeshi, N.D., Martell, J.D., Mootha, V.K., Carr, S.A., and Ting, A.Y. (2013). Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. Science 339, 1328–1331.
Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. J Cell Biol 196, 801–810.
Schwende, H., Fitzke, E., Ambs, P., and Dieter, P. (1996). Differences in the state of differentiation of THP-1 cells induced by phorbol ester and 1,25-dihydroxyvitamin D3. J Leukoc Biol 59, 555–561.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498–2504.
Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D., and Escalante-Semerena, J.C. (2002). Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science 298, 2390–2392.
Suprewicz, Ł., Swoger, M., Gupta, S., Piktel, E., Byfield, F.J., Iwamoto, D. V., Germann, D., Reszeć, J., Marcińczyk, N., Carroll, R.J., et al. (2022). Extracellular vimentin as a target against SARS-CoV-2 host cell invasion. Small 18, 2105640.
Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 47, D607–D613.
Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Tissue-based map of the human proteome. Science 347, 1260419.
Vikis, H.G., and Guan, K.L. (2004). Glutathione-*S*-transferase-fusion based assays for studying protein-protein interactions. In: Fu, H., ed. Protein-Protein Interactions. Methods in Molecular Biology. New York: Humana Press. 175–186.
Wang, H., and Matsumura, P. (1996). Characterization of the CheAS/CheZ complex: a specific interaction resulting in enhanced dephosphorylating activity on CheY-phosphate. Mol Microbiol 19, 695–703.
Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117–120.
Weinert, B.T., Iesmantavicius, V., Wagner, S.A., Schölz, C., Gummesson, B., Beli, P., Nyström, T., and Choudhary, C. (2013). Acetyl-phosphate is a critical determinant of lysine acetylation in *E. coli*. Mol Cell 51, 265–272.
Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H. Y., Li, A., Ping, X.L., Lai, W.Y., et al. (2016). Nuclear m^6A reader YTHDC1 regulates mRNA splicing. Mol Cell 61, 507–519.
Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., et al. (2001). Global analysis of protein activities using proteome chips. Science 293, 2101–2105.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-023-2316-2. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.