Photocatalysis Hot Paper

How to cite: *Angew. Chem. Int. Ed.* **2022,** *61,* e202202008
International Edition: doi.org/10.1002/anie.202202008
German Edition: doi.org/10.1002/ange.202202008

Photocatalytic Chemical Crosslinking for Profiling RNA–Protein Interactions in Living Cells

Huixin Luo[+], Wei Tang[+], Hongyu Liu, Xiangmei Zeng, William Shu Ching Ngai, Rui Gao, Heyun Li, Ran Li, Huangtao Zheng, Jianting Guo, Fangfei Qin, Gang Wang, Kexin Li, Xinyuan Fan, Peng Zou,* and Peng R. Chen**

Abstract: The dynamic interactions between RNAs and proteins play crucial roles in regulating diverse cellular processes. Proteome-wide characterization of these interactions in their native cellular context remains desirable but challenging. Herein, we developed a photocatalytic crosslinking (PhotoCAX) strategy coupled with mass spectrometry (PhotoCAX-MS) and RNA sequencing (PhotoCAX-seq) for the study of the composition and dynamics of protein-RNA interactions. By integrating the blue light-triggered photocatalyst with a dual-functional RNA–protein crosslinker (RP-linker) and the phase separation-based enrichment strategy, PhotoCAX-MS revealed a total of 2044 RBPs in human HEK293 cells. We further employed PhotoCAX to investigate the dynamic change of RBPome in macrophage cells upon LPS-stimulation, as well as the identification of RBPs interacting directly with the 5′ untranslated regions of SARS-CoV-2 RNA.

[*] H. Luo,[+] H. Liu, X. Zeng, W. S. C. Ngai, H. Li, H. Zheng, J. Guo, X. Fan, P. Zou, P. R. Chen
College of Chemistry and Molecular Engineering, Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University
Beijing, 100871 (China)
E-mail: xinyuanfan@pku.edu.cn
zoupeng@pku.edu.cn
pengchen@pku.edu.cn

W. Tang,[+] R. Gao, R. Li, F. Qin, G. Wang, K. Li, P. Zou, P. R. Chen
Peking-Tsinghua Center for Life Sciences
Beijing, 100871 (China)

P. Zou
PKU-IDG/McGovern Institute for Brain Research
Beijing, 100871 (China)

P. Zou
Chinese Institute for Brain Research (CIBR)
Beijing 102206 (China)

H. Luo[+]
State Key Laboratory of Bioactive Substance and Function of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking UnionMedical College
Beijing 100050 (China)

[+] These authors contributed equally to this work.

Introduction

The RNA–protein interaction dynamics are critical for various cellular processes. In particular, RNA regulations are rigorously controlled by the RNA-binding proteins (RBPs), which are involved in the RNA splicing, localization, translation and degradation, etc.[1] Perturbation of these RNA–protein interactions can lead to cellular dysfunctions and severe diseases (e.g., neurodegeneration).[2] Therefore, a comprehensive understanding of the RNA–protein interaction dynamics is important for disclosing the fundamental mechanisms underlying these cellular processes.[3] As a result, methodology development for RBPs profiling has attracted great attentions in recent years. Conventional approaches require UV-based crosslinking to form covalent bonds between RBPs and their cognate RNAs, such as oligo(dT) bead capture,[4] metabolic RNA labeling,[5] and organic-aqueous phase separation.[6] Such crosslinking only occurs when there is a direct contact between an RNA and a protein, ensuring the labeling fidelity of the RNA–protein interactome. However, the unsatisfying labeling efficiency as well as the UV cytotoxicity hampered their further applications.[7] To overcome these problems, proximity labeling-based enzymatic approaches have been developed. For example, RNA–protein interaction detection (RaPID) or CRISPR-assisted method (CARPID) allowed the biotinylation of RBPs via reactive chemical species catalytically generated by genetically encoded enzymes.[8] However, such genetic approaches are only applicable to samples that are capable of gene transfection, rendering a large number of sample types inaccessible, including immune cells, primary cells and intact tissues. Furthermore, the diffusive nature of reactive species generated may lead to false labeling from the indirect RNA–protein interactions.

With our continued efforts on the development and exploration of photocatalysis mediated labelling strategy in living systems,[9] we envisioned that such strategy might be further adopted for the in situ crosslinking and comprehensive profiling of RNA–protein interactions with integrated advantages compared to the current existing methods, including: i) the crosslinking fidelity of direct interactions between RNAs and RBPs, ii) catalytic generation of the reactive species for efficient crosslinking, iii) non-genetic approach renders wide applicability to hard-to-transfect samples, and iv) temporal resolution provided by an external photo-trigger.

Herein, we report a photocatalysis-enabled crosslinking strategy for efficient labelling and profiling of the direct RNA–protein interactome in living cells by external light (PhotoCAX) (Figure 1a). Eosin was employed as the photocatalyst to generate singlet oxygen ($^1O_2$) upon blue light irradiation, which activates a dual-functional RNA–protein

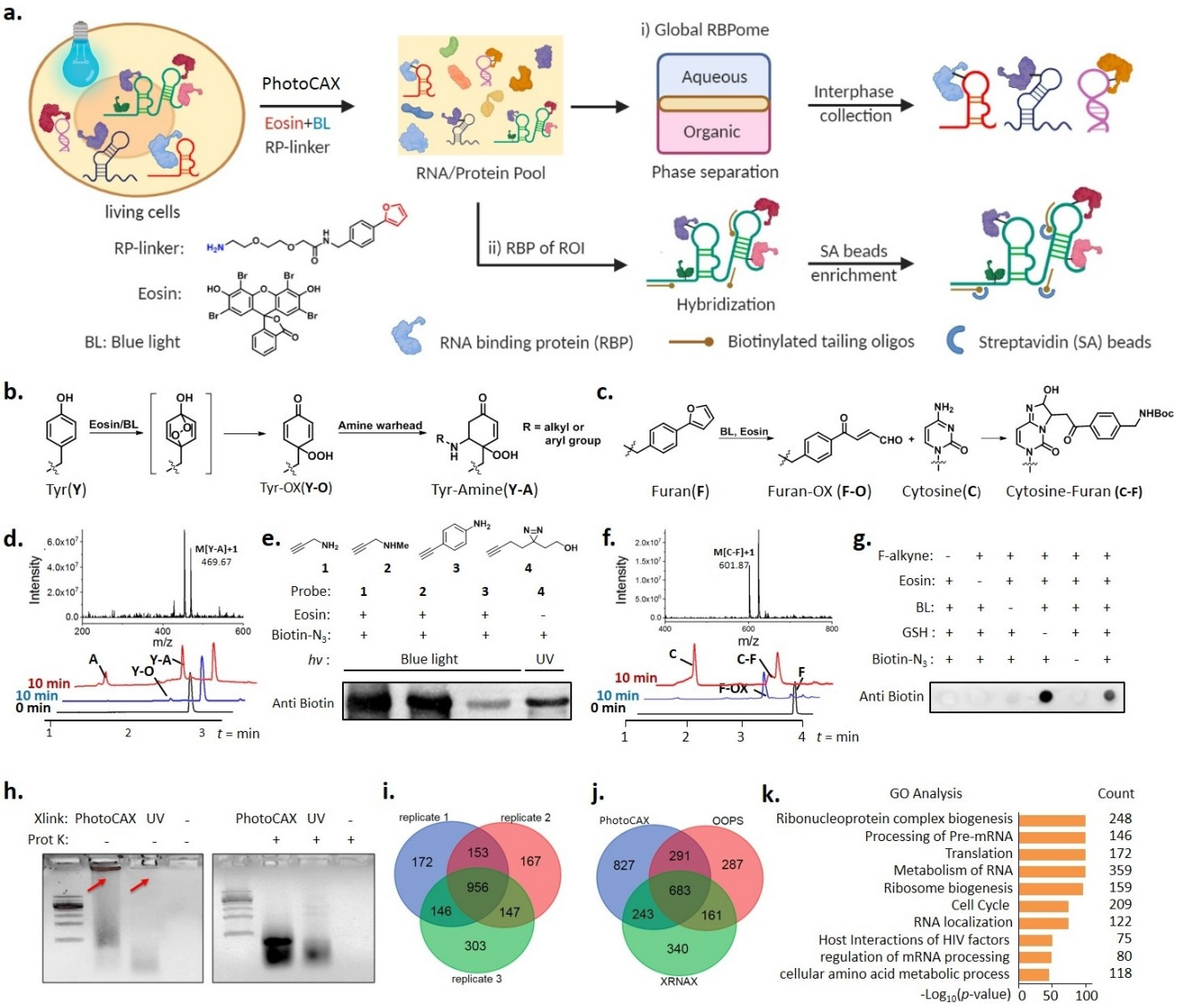

Figure 1. The development of photocatalytic RNA–protein crosslinking (PhotoCAX) strategy. a) Schematic illustration of the PhotoCAX strategy. A dual-functional RNA–protein crosslinker (RP-linker) containing the amine and furan moiety at each terminus was designed to capture direct RNA–protein interactions in living cells. b) Proposed mechanism of photocatalytic oxidation of amino acid side-chains such as tyrosine residues by photo-irradiation of eosin, which can subsequently react with the amine nucleophile on the dual-functional RP-linker. c) Proposed mechanism of photocatalytic conjugation reaction between cytosine and the furan moiety from the dual-functional RP-linker. d) LC-MS detection of eosin-mediated tyrosine photo-oxidation product (**Y-O**) and a tyrosine-amine adduct (**Y-A**). A mixture of eosin (0.05 mM), benzylamine (0.5 mM) and Cbz-Tyr-OMe (0.2 mM) in 0.1 mL ACN/H_2O (1:1) was irradiated under a mild blue light (BL, 465–475 nm) at 24 mW cm^{-2} for 10 min at room temperature, and subjected to LC-MS analysis without further treatment. The mass of **Y-A** adduct with the predicted structure shown was detected. $[M+H]^+$: 435.21. e) Conjugation of the amine warhead (**1**, **2**, **3**) from the crosslinker with model protein bovine serum albumin (BSA) comparing with the UV-triggered diazirine crosslinker (**4**). A mixture of BSA (30 ng) and crosslinkers (0.5 mM) was illuminated with BL or UV (200 mJ cm^{-2}, 254 nm), followed by CuAAC-mediated biotinylation using biotin-azide and subsequent immunoblotting analysis. f) LC-MS detection of eosin-mediated furan photo-oxidation product (**F-OX**) and a furan-cytosine adduct (**F-C**). The mass of **F-C** adduct with the predicted structure shown was detected. $[M+H]^+$: 601.24. g) Conjugation of furan warhead from the crosslinker with total RNA extracted from HEK293 cells. Furan-alkyne was used to label RNAs, and the adducts were then biotinylated and analyzed by streptavidin dot blotting. h) Comparison of crosslinking efficiency between PhotoCAX and UV-based crosslinking (200 mJ cm^{-2}, 254 nm). HEK293 cells were subjected to UV-based crosslinking or PhotoCAX. RNA–protein complexes were extracted with phase separation protocol. Enriched RNAs were analyzed by 2% agarose gel electrophoresis after removal of RBPs with Prot K (Proteinase K). i) Venn diagram of the RBPome identified in three independent PhotoCAX experiments in HEK293 cells. j) Venn diagram comparing three methods for pulling down RBPs in HEK293 cells; PhotoCAX, OOPS and XRNAX. k) GO analysis of the RBPs identified in PhotoCAX strategy. The count of proteins for each individual term is also shown.

crosslinker (RP-linker) to capture direct interactions between RNA and RBPs. Subsequently, crosslinked RNA–protein complexes could be conveniently enriched via phase separation-based workflow for further analysis (Figure 1a).[6a,b] By combining with quantitative mass spectrometry, we applied our PhotoCAX strategy to monitor the dynamic remodeling of RBPome in macrophages upon lipopolysaccharides (LPS)-stimulation, as well as to identify host proteins associated with 5′ untranslated regions (UTRs) of SARS-CoV-2 RNA.

Results and Discussion

We started by developing an efficient photocatalytic crosslinking chemistry to capture direct interactions between RNAs and RBPs. Considering the diversity in reactivity of these biomacromolecules, we designed a dual functional crosslinker (RP-linker), which consists of an amine group at one terminal and a furan group at the other terminal (Figure 1a). Upon singlet oxygen ($^1O_2$) generated by photocatalytic conditions, the electron-rich side chains (e.g., tyrosine, tryptophan and histidine) on RBPs are activated to conjugate with the amine group (Figure 1b).[10] Meanwhile, based on the furan-based DNA crosslinking chemistry reported previously,[11] the furan group could be transformed to a cytosine-reactive 1,4-enedione intermediate to facilitate RNA labelling (Figure 1c),[12] which results in the covalent linkage within RNA–protein complexes for subsequent analysis. Unlike the diffusion-based strategies, the crosslinking fidelity of such dual functional probe strategy relies on the length and flexibility of the linker, possessing further optimization space for reducing false positive crosslinking. Eosin was chosen as the photocatalyst to generate $^1O_2$ upon blue light irradiation due to its high efficiency as well as compatibility with living samples.[13]

The conjugation between amine and proteins triggered by eosin photocatalysis was first evaluated using benzylamine as a model. The reaction was carried out in aqueous acetonitrile solution ($ACN/H_2O = 1:1$, v/v) with mild blue light (465–475 nm) at 24 $mW\,cm^{-2}$ for 10 min at room temperature. Liquid chromatography-mass spectrometry (LC-MS) analysis revealed adduct formation between the benzylamine probe and tyrosine, histidine or tryptophan (Figure 1d and Supporting Information Figure 1). To test whether the reaction is applicable on protein labeling, we used bovine serum albumin (BSA) as a model and compared protein labeling efficiencies of alkyne-conjugated alkylamine (**1**, **2**), arylamine (**3**), and diazirine (**4**) probes in the presence of eosin and blue light (Figure 1e). The resulting alkyne adducts were derivatized with biotin-azide via the copper(I)-catalyzed alkyne-azide cycloaddition (CuAAC) and visualized through immunoblotting. As shown in figure 1e, alkylamine exhibited the highest crosslinking yield. Negative controls omitting either eosin, alkyne-amine, or blue light, yielded negligible biotin signal. We conclude from the above data that alkylamine is an appropriate functional group for eosin-mediated protein labeling.

We next evaluated the labelling efficiency and specificity of eosin-mediated furan probe on cytosine. The photo-oxidative conjugation product was confirmed by LC-MS analysis (Figure 1f and Supporting Information Figure 2). We further validated that the photo-oxidative product of furan could be used to label RNA molecules extracted from HEK293 cells (Figure 1g). Noticeably, glutathione (GSH) caused negligible perturbation to furan photo-oxidation, suggesting the high biocompatibility of eosin-mediated crosslinking in intracellular conditions. To evaluate the RNA-specificity of eosin-mediated labeling, we compared the labeling activity of furan probes on DNA, and RNA. Dot blot analysis showed that the furan-alkyne probe labels RNA only but not double-strand DNA (Supporting Information Figure 3). We reasoned that the existence of Watson–Crick base-pairing in double-stranded DNA blocks the labeling reaction.

After in vitro validation of PhotoCAX, we next evaluated its crosslinking efficiencies in living cells. We incorporated eosin and RP-linker to cultured HEK293 cells expressing the FLAG tagged SFPQ (splicing factor proline/glutamine-rich, known as an RBP) followed by blue light irradiation,[14] SFPQ-RNA complexes in the PhotoCAX extracts were then enriched by FLAG beads (Supporting Information Figure 4a). Enriched RNAs were analyzed by 2 % agarose gel electrophoresis after removal of RBPs using Proteinase K. As shown in Supporting Information Figure 4b and c, PhotoCAX exhibited higher crosslinking yield comparing to conventional UV-mediated crosslinking. Taken together, the above characterizations proved our PhotoCAX strategy as an effective and highly specific measure to conjugate RNAs and proteins under mild conditions in live cells.

We then applied our PhotoCAX technique for the global profiling of RNA–protein interactions in living cells. The crosslinked RNA–protein adducts were enriched and purified following the phase separation protocol (Supporting Information Figure 5a).[6a,b] Agarose gel electrophoretic mobility assay indicated the presence of RNA-RBP complexes in the extract (Figure 1h, left, Supporting Information Figure 5b). Proteinase K treatment restored in-gel RNA mobility, which suggested effective binding of RBPs and RNAs in the extract (Figure 1h, right). A small crosslinking background was observed in the absence of RP-linker which was at the similar level as UV-crosslinking (Supporting Information Figure 5c). This can be explained by the fact that residues from interface of RBPs may crosslink RNAs via singlet O_2-based reactions (e.g., the amine from lysine residues directly crosslink photo-oxidation products of guanosine). We pulled down crosslinked RNA–protein complexes from HEK293 cells and analyzed the sample with liquid chromatography-tandem mass spectrometry (LC-MS/MS). In this PhotoCAX-MS analysis, a total of 2044 proteins were identified in three independent experiments, with 1402 proteins from at least two replicates (Figure 1i). Among these, 827 proteins have not been previously reported by OOPS or XRNAX methods in HEK293 cells (Figure 1j),[6a,b] which demonstrated the higher efficiency of PhotoCAX-MS for in-depth RBPome profiling. Gene Ontology (GO)

analysis revealed an over-representation of RNA-related terms, including ribonucleoprotein complex biogenesis, processing of pre-mRNA, translation and metabolism of RNA (Figure 1k).

As a central component of the innate immune response, macrophage can sense pathogen components such as LPS, the essential constituent of the outer membrane of gram-negative bacteria.[15] While the LPS-stimulated transcriptome-wide changes in macrophage are well documented,[16] the effect on proteome changes are yet to be discovered. In addition, emerging evidence highlighted the impact of RBPs on the post-transcriptional control of immune responses.[17] To study the dynamic change of RBPs upon such stimuli, we used LPS to treat mouse-derived macrophage RAW 264.7 cells grown in medium with heavy isotope (Arg10-Lys8). Unstimulated cells grown in medium with light isotope (Arg0-Lys0) were used as the control. The PhotoCAX-MS strategy was applied to both samples and the SILAC ratio between these two samples was interpreted as the RNA-binding efficiency for a given protein (Figure 2a).

A total of 1926 proteins were identified by using MaxQuant, with 1299 of them observed in three independent replicates (Figure 2b). GO analysis indicated that RBPs were strongly enriched in PhotoCAX extracts (Figure 2c). To better describe the LPS induced changes in the RBPome, we applied a volcano plot analysis, in which only proteins with p-value <0.05 and fold-change >1.5 (either increase or decrease) were considered significant (Figure 2d). In particular, the abundance of 11 RNA-associated proteins were up-regulated and 12 were down-regulated, including SMU1, WDR77, PRC1 as well as NLRP3. Since NLRP3 is a key component of the inflammasome complex and plays essential roles in LPS-induced inflammasome priming and activation,[18] we were interested in the potential role of NLPR3-RNA interactions during inflammasome priming and activation.

To specifically analyze NLRP3-interacting RNAs, we treated RAW 264.7 cells with LPS and crosslinking was induced using the RP-linker. NLRP3-RNA complexes in the PhotoCAX extract were then enriched using anti-NLRP3 antibody. The enriched samples were analyzed by high-throughput sequencing (PhotoCAX-seq, Figure 3a). Two biological replicates exhibited high reproducibility (Pearson's correlation coefficient >0.98, Supporting Information Figure 6). For both basal and LPS stimulation conditions, we performed differential expression analysis between samples post- versus pre-enrichment with the anti-NLRP3 antibody. The enrichment cutoff was defined as log2 fold-change >0 and p-value <0.01, which yielded respectively 1781 and 2761 transcripts from Raw 264.7 cells with and without LPS stimulation, respectively (Figure 3b and Supporting Information Figure 7). Notably, several lncRNAs were found to be enriched as targets of NLRP3, including *Gm43660*, *Gm13833*, *Morrbid*, *Neat1*, *Jpx*, *Gm38414* and *Malat1* (Figure 3c). Since the conserved lncRNA *Malat1* has been reported to regulate the innate immune responses,[19] we analyzed the expression level of *Malat1* by using RT-qPCR before and after LPS stimulation. As expected, an upregulation trend of *Malat1* was observed in Raw 264.7

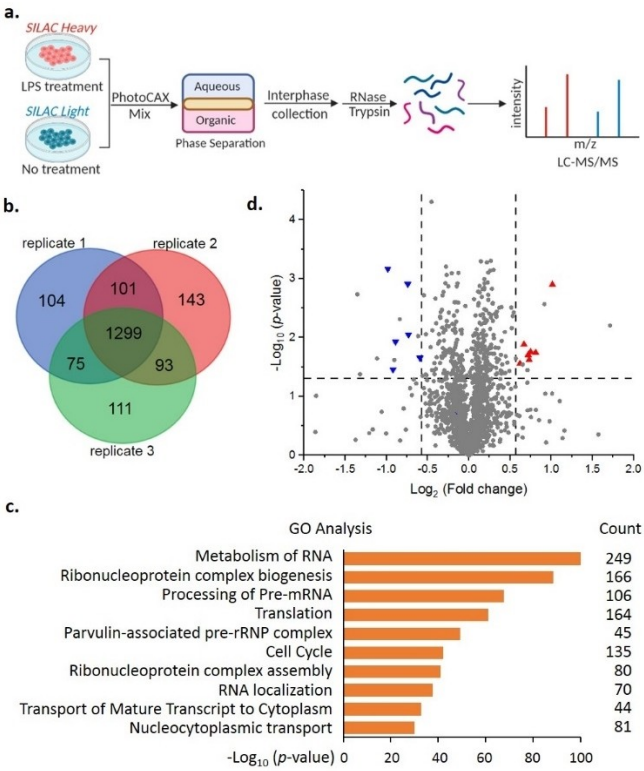

Figure 2. PhotoCAX-mediated global profiling of RBPome in LPS-treated macrophage cells. a) Schematic illustration of PhotoCAX-MS experimental workflow. RAW 264.7 cells cultured in medium containing heavy isotope (SILAC Heavy) were treated with LPS, while cells cultured in medium containing light isotope (SILAC Light) received no treatments. Following PhotoCAX and cell lysis, RBPs were extracted through the phase separation protocol, followed by LC-MS/MS analysis. b) Venn diagram of the RBPomes identified in three independent PhotoCAX experiments in RAW 264.7 cells. c) GO analysis of the RBPs identified by our PhotoCAX strategy in RAW 264.7 cells. d) Volcano plot showing the fold changes and p-values of the RNA-interacting proteome after LPS treatment. Proteins with p-value less than 0.05 and a minimum of 1.5-fold increase (red) or decrease (blue) were considered as significantly changed.

cells after LPS treatment (Figure 3d). Since the RP-linker used in our PhotoCAX strategy can capture direct RNA–protein interactions, we next pulled down the crosslinked *Malat1*-protein complexes by using biotinylated antisense oligos. Control experiments were conducted with non-PhotoCAX conditions, pre-pull-down RNase treatment of lysate, as well as non-interacting control probes. Silver staining showed that *Malat1* antisense oligo probes pulled down a significant amount of proteins from cell lysates, whereas all control samples showed negligible signals (Supporting Information Figure 8). Indeed, immunoblotting analysis confirmed that *Malat1* directly interacted with NLRP3 (Figure 3e). Interestingly, we also found that the *Malat1* formed direct complex with DDX3X, which have been recently shown to interact with NLRP3 under both LPS-primed and inflammasome-activating conditions.[20] These observations further indicated the role of *Malat1* in the inflammasome priming and activation processes.

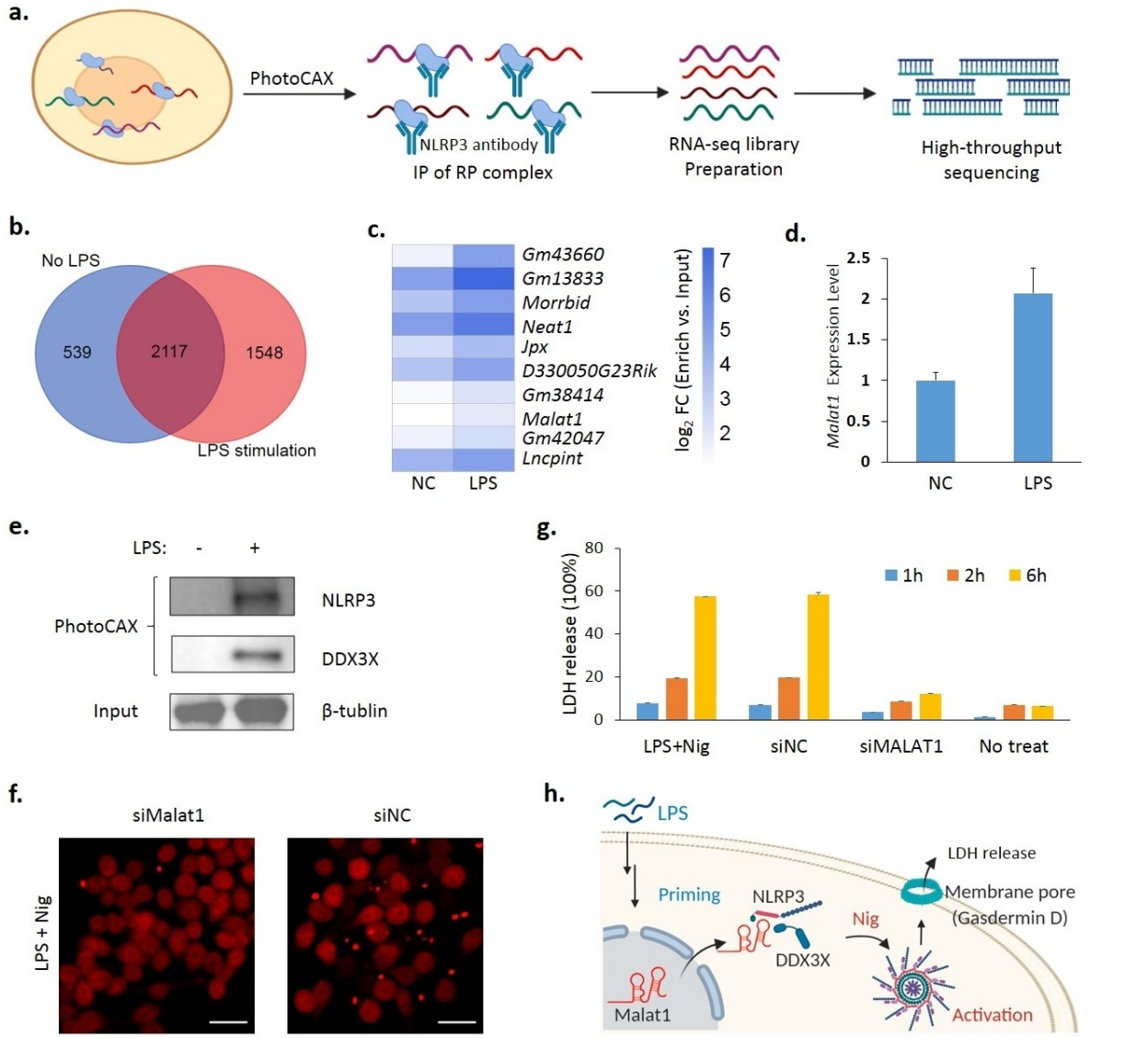

Figure 3. Transcriptome-wide identification and analysis of NLRP3-interacting RNAs. a) Schematic representation of the PhotoCAX-seq procedure identifying target transcripts bound by NLRP3 in LPS-activated RAW 264.7 cells. Following PhotoCAX and cell lysis, NLRP3-RNA complexes were pulled down using NLRP3-specific antibodies and subsequently treated with Proteinase K to release NLRP3-associated RNAs, which were analyzed by high-throughput sequencing. b) Venn diagram of target transcripts identified in two independent PhotoCAX experiments in the RAW 264.7 cells with or without LPS stimulation, respectively. c) Enrichment of NLRP3-associated long noncoding RNAs (lncRNAs) in the negative control (NC, omitting LPS) and LPS-activated RAW 264.7 cells. d) RT-qPCR analysis of the lncRNA *Malat1* transcription level upon LPS treatment. Data are the mean of three technical replicates ±1 s.d. e) Western blot analysis of *Malat1*-interacting proteins in RAW 264.7 after LPS treatment. Following PhotoCAX and cell lysis by sonication, RNA–protein complexes were pulled out by biotinylated antisense oligos against lncRNA *Malat1*. Associated proteins were eluted with free biotin and analyzed by western blot against DDX3X and NLRP3. f) Visualization of inflammasome assembly in *Malat1*-knockdown cells. RAW 264.7 cells stably expressing mCherry-fused ASC protein were co-stimulated with LPS and nigericin (Nig). Scale bar: 5 μm. g) *Malat1* promotes caspase-1-dependent pyroptosis. LDH release assay showed suppressed pyroptosis after LPS/Nig treatment in *Malat1*-knockdown cells. h) Proposed mechanism of *Malat1*-promoted NLRP3 inflammasome assembly. The *Malat1* was up-regulated and released from nucleus to the cytoplasm upon LPS treatment, where it further interacted with NLRP3 and DDX3X to facilitate the priming process. Endogenous danger signals such as Nig can induce activation of the NLRP3 inflammasome, followed by the recruitment of the adaptor protein ASC, which undergoes prion-like polymerization that causes caspase-1 activation and Gasdermin D (GSDMD) cleavage. The assembly of the N-terminal domain of GSDMD (N-GSDM) on cell membrane as oligomeric pores ultimately led to pyroptotic cell death.

Whereas murine *Malat1* and its human ortholog are lncRNAs known to reside in the nucleus of unstressed cells,[21] the assembly of various canonical inflammasomes occurs in the cytoplasm. We next monitored the subcellular localization of *Malat1* by RNA fluorescence in situ hybridization (FISH) upon LPS stimulation. Indeed, *Malat1* was found in the cytoplasm (Supporting Information Figure 9), suggesting that *Malat1* was transported from the nucleus to the cytoplasm during NLRP3 inflammasome priming. To further evaluate the role of *Malat1* in NLRP3 inflamma-

some, we used small interfering RNAs (siRNAs) to knock down the expression level of *Malat1* in RAW-ASC cells, an engineered cell line that stably expresses the inflammasome adapter protein ASC fused to mCherry (Supporting Information Figure 10). After LPS priming, the canonical activator nigericin (Nig) elicited robust inflammasome activation in RAW-ASC cells, as reflected by the formation of ASC specks. Notably, we observed that *Malat1* silencing dramatically suppressed the ASC speck formation in LPS/Nig co-stimulated RAW-ASC cells (Figure 3f). Similarly, LDH release assay showed suppressed pyroptosis after LPS/Nig treatment in *Malat1*-knockdown cells (Figure 3g), indicating that *Malat1* likely promoted caspase-1-dependent pyroptotic cell death. Taken together, our data demonstrated that *Malat1* was up-regulated and released from the nucleus to the cytoplasm, where it interacted with NLRP3 and DDX3X to facilitate the priming process, which ultimately led to Gasdermin D (GSDMD) cleavage and cell pyroptosis (Figure 3h)

Finally, we applied PhotoCAX-MS to identify host proteins associated with the 5′ UTRs of SARS-CoV-2 RNA. As the causative virus of the ongoing COVID-19 pandemic, the SARS-CoV-2 uses its RNA's UTR region to interact with proteins in host cells to facilitate the viral replication and transcription in the life cycle.[22] For example, the 5′ UTR of the viral genome is involved in translation initiation of viral proteins.[23] We first expressed the 5′ UTR of SARS-CoV-2 RNA in HEK293 cells (Supporting Information Figure 11). RNA–protein complexes were crosslinked in situ with PhotoCAX, followed by 5′ UTR pull-down using biotinylated antisense oligos. After streptavidin (SA) beads enrichment, associated proteins were eluted with free biotin and the fractions separated through electrophoresis were subjected to LC/MS-MS analysis (Figure 4a). Silver staining of enriched samples showed effective isolation of target proteins (Figure 4b). A total of 193 proteins were identified using MaxQuant, where 100 of them exsisted in two independent replicates (Figure 4c). As a positive control, we co-expressed the SARS-CoV-2 nucleocapsid (N) protein, which has been shown to bind with the coronavirus 5′ UTR to facilitate translation of viral RNAs, in HEK293 cells.[24] Tandem mass spectrometry identified over 33 peptides of the N protein (Supporting Information Figure 12). GO analysis showed a strong enrichment of the Nop56p-associated pre-rRNA complex, ribonucleoprotein complex biogenesis, telomere maintenance, RNA localization, mRNA splicing and regulation of translation (Figure 4d).

Built upon this expanded interactome for the viral 5′ UTR RNA, we constructed a protein-protein association network to assist the analysis of their function in virus-associated biological processes Network-based analysis revealed broad functional categories of host proteins that associate with the 5′ UTR, including ribosomal proteins, translation factors, heterogeneous nuclear ribonucleoproteins (hnRNPs), RNA helicases, minichromosome maintenance proteins (MCMs), as well as eukaryotic chaperonin t-complex proteins (Supporting Information Figure 13). In agreement with the crucial roles of mRNA translation, we identified 19 ribosomal proteins (Supporting Information

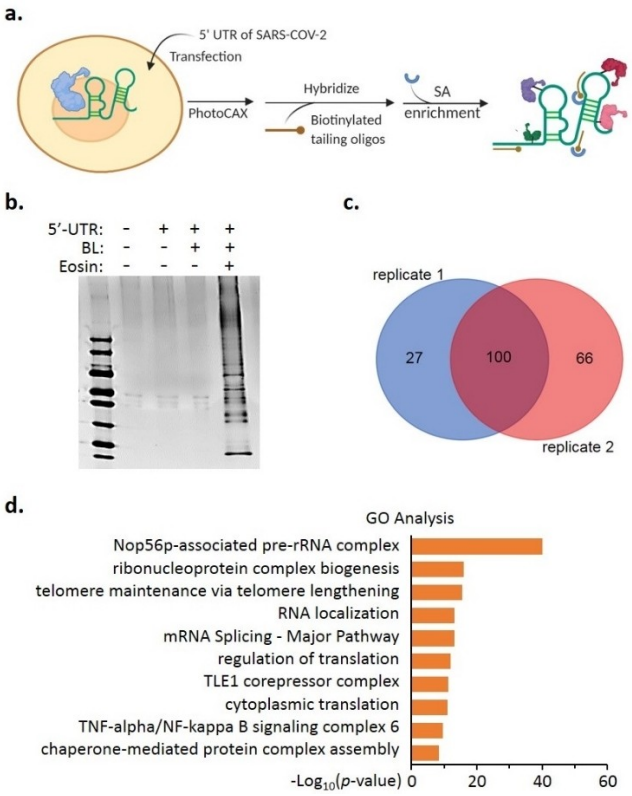

Figure 4. PhotoCAX-MS identification of host proteins associated with 5′ UTR of SARS-CoV-2 RNA. a) Schematic illustration of proteomic experimental workflow. The 5′ UTR of SARS-CoV-2 RNA was expressed in HEK293 cells. Following PhotoCAX and cell lysis by sonication, target RNAs were captured by biotinylated antisense oligos. RNA-associated proteins were eluted and analyzed by LC-MS/MS. b) Silver stain revealing enrichment of proteins associated with 5′ UTR of SARS-CoV-2 RNA. c) Venn diagram showing a summary of the host proteins identified in two independent PhotoCAX experiments in HEK293 cells. d) GO analysis of host interaction proteins of 5′ UTR of SARS-CoV-2 RNA identified by PhotoCAX-MS.

Figure 13, blue cluster) which have previously been used to block the viral replication or act as activators for the host immune factors.[25] Among the translation factors, we found that the eukaryotic translation initiation factors (EIF3A/3B/5A) and the eukaryotic translation elongation factors (EEF1G/1D/1B2) were both strongly enriched in our experiments. Meanwhile, the polypyrimidine tract-binding protein 1 (PTBP1) was also observed, which can interact with the 5′ UTR and canonical translation factors to facilitate the recruitment of ribosomes.[26] The hnRNPs have been shown to bind to the 5′ UTR as well. For example, hnRNP A1 binds to the complement of the 5′ leader sequence (negative-strand leader) and to the complement of the TRS-B sequences.[23] In this study, we identified hnRNP A1, A2B1, K, M and U (Supporting Information Figure 13, red cluster), which have attracted great attentions as potential antiviral targets.[27] Besides, RNA helicases such as DDX21 and DHX9 play vital roles in transcription as well as the regulation of host innate immunity during virus infection.[28] The role of MCMs in viral replication has been confirmed

by depleting MCMs and assaying transient and long-term maintenance of the viral episomes (Supporting Information Figure 13, red cluster).[29] Eukaryotic chaperonin T-complex protein (CCT1-7) was identified as a cellular factor required for late events in the replication of mammalian reovirus (Supporting Information Figure 13, green cluster).[30] In addition, we also identified splicing factor 3B subunit 3 (SF3B3) as a new host RBP of SARS-CoV-2, which may be involved in viral replication and transcription.[31] Taken together, our network analysis connected the host RBPome to the emerging SARS-CoV-2 biology and provided a map of putative regulatory hubs in SARS-CoV-2 infections.

Conclusion

In summary, we developed the photocatalytic chemistry enabled PhotoCAX strategy for global and in-depth profiling and mechanistic investigation of direct interactions between RNA and RBPs in living cells. By coupling with phase separation-based enrichment protocol and the protein mass spectrometry, we identified a total of 2044 RBPs in HEK293 cells, in which 827 were newly discovered RBPs. Furthermore, by taking the advantage of this chemical approach, RBPome dynamic of hard-to-transfect RAW 264.7 cells could be characterized. NLRP3 as a newly defined RNA-binder has been validated in RAW 264.7 cells upon LPS stimulation, which interacts with lncRNA *Malat1* to facilitate the inflammasome priming and activation. Finally, we also profiled the host proteins associated with 5′ UTR of SARS-CoV-2 RNA. Our method unveils a diverse array of 5′ UTR binding proteins, including both well-characterized host proteins involved in viral replication/transcription and several other protein families with RNA-binding capacity, thus demonstrated the power of PhotoCAX for discovering pathogen RNA's interactions with host proteome during viral infection. The advantages of PhotoCAX include: i) the unique dual-functional RP-linker ensures the labelling efficiency of different biomacromolecules as well as the labelling fidelity of direct interactions; ii) the designed catalytic activation provided sufficient reactive intermediates for efficient labelling; iii) the applicability to hard-to-transfect cells; and iv) spatial and temporal resolution are conveniently achievable through a light trigger. In addition, a spatial-resolved approach for RNA–protein interactome profiling could yield valuable insight into local functions of RNA–protein complexes. Here, we adopted triphenylphosphonium (TPP) for the uptake of eosin into mitochondria (Supporting Information Figure 14). On the other hand, given the potential issues caused by the inherent high reactivity of $^1O_2$, certain cell damages are unavoidable along with the RNA-RBP crosslinking processes, making the development of more biocompatible techniques or faster labeling reactions highly desirable. Continued efforts on the optimization of the crosslinking kinetics as well as other bioorthogonal methods are currently under investigation in our lab.

Acknowledgements

We acknowledge funding from the Ministry of Science and Technology (2018YFA0507600, 2016YFA0501500 and 2017YFA0503600), the National Natural Science Foundation of China (22007001, 32088101, 21727806, 91753131, 91753000, 91957101, 22077004), Beijing Natural Science Foundation (Z200010). P.Z. is sponsored by Li Ge-Zhao Ning Life Science Junior Research Fellowship and Bayer Investigator Award. We thank Profs. Na Dong (China Agricultural University) for providing the Raw-ASC cell. We thank National Center for Protein Sciences at Peking University in Beijing, China, for assistance with Fragment Analyzer 12.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available in the Supporting Information of this article.

Keywords: Crosslinking · Photocatalysis · RNA-Binding Protein · RNA–Protein Interactome · SARS-CoV-2

[1] a) L. Nair, H. Chung, U. Basu, *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 123–136; b) M. Müller-McNicoll, K. M. Neugebauer, *Nat. Rev. Genet.* **2013**, *14*, 275–287.

[2] a) E. Jankowsky, M. E. Harris, *Nat. Rev. Mol. Cell Biol.* **2015**, *16*, 533–544; b) A. E. Brinegar, T. A. Cooper, *Brain Res.* **2016**, *1647*, 1–8.

[3] S. Gerstberger, M. Hafner, T. Tuschl, *Nat. Rev. Genet.* **2014**, *15*, 829–845.

[4] A. Castello, R. Horos, C. Strein, B. Fischer, K. Eichelbaum, L. M. Steinmetz, J. Krijgsveld, M. W. Hentze, *Nat. Protoc.* **2013**, *8*, 491–500.

[5] a) X. Bao, X. Guo, M. Yin, M. Tariq, Y. Lai, S. Kanwal, J. Zhou, N. Li, Y. Lv, C. Pulido-Quetglas, X. Wang, L. Ji, M. J. Khan, X. Zhu, Z. Luo, C. Shao, D. H. Lim, X. Liu, N. Li, W. Wang, M. He, Y. L. Liu, C. Ward, T. Wang, G. Zhang, D. Wang, J. Yang, Y. Chen, C. Zhang, R. Jauch, Y. G. Yang, Y. Wang, B. Qin, M. L. Anko, A. P. Hutchins, H. Sun, H. Wang, X. D. Fu, B. Zhang, M. A. Esteban, *Nat. Methods* **2018**, *15*, 213–220; b) R. Huang, M. Han, L. Meng, X. Chen, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 3879–3887.

[6] a) J. Trendel, T. Schwarzl, R. Horos, A. Prakash, A. Bateman, M. W. Hentze, J. Krijgsveld, *Cells* **2019**, *176*, 391–403.e319; b) R. M. L. Queiroz, T. Smith, E. Villanueva, M. Marti-Solano, M. Monti, M. Pizzinga, D. M. Mirea, M. Ramakrishna, R. F. Harvey, V. Dezi, G. H. Thomas, A. E. Willis, K. S. Lilley, *Nat. Biotechnol.* **2019**, *37*, 169–178; c) E. C. Urdaneta, C. H. Vieira-Vieira, T. Hick, H. H. Wessels, D. Figini, R. Moschall, J. Medenbach, U. Ohler, S. Granneman, M. Selbach, B. M. Beckmann, *Nat. Commun.* **2019**, *10*, 990.

[7] a) K. M. Meisenheimer, T. H. Koch, *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 101–140; b) J. Ule, K. B. Jensen, M. Ruggiu, A. Mele, A. Ule, R. B. Darnell, *Science* **2003**, *302*, 1212–1215; c) D. D. Licatalosi, A. Mele, J. J. Fak, J. Ule, M. Kayikci, S. W.

Chi, T. A. Clark, A. C. Schweitzer, J. E. Blume, X. Wang, J. C. Darnell, R. B. Darnell, *Nature* **2008**, *456*, 464–469; d) X. Li, J. Song, C. Yi, *Genomics Proteomics Bioinf.* **2014**, *12*, 72–78.
[8] a) M. Ramanathan, K. Majzoub, D. S. Rao, P. H. Neela, B. J. Zarnegar, S. Mondal, J. G. Roth, H. Gai, J. R. Kovalski, Z. Siprashvili, T. D. Palmer, J. E. Carette, P. A. Khavari, *Nat. Methods* **2018**, *15*, 207–212; b) W. Yi, J. Li, X. Zhu, X. Wang, L. Fan, W. Sun, L. Liao, J. Zhang, X. Li, J. Ye, F. Chen, J. Taipale, K. M. Chan, L. Zhang, J. Yan, *Nat. Methods* **2020**, *17*, 685–688.
[9] Z. Huang, Z. Liu, X. Xie, R. Zeng, Z. Chen, L. Kong, X. Fan, P. R. Chen, *J. Am. Chem. Soc.* **2021**, *143*, 18714–18720.
[10] a) P. Di Mascio, G. R. Martinez, S. Miyamoto, G. E. Ronsein, M. H. G. Medeiros, J. Cadet, *Chem. Rev.* **2019**, *119*, 2043–2086; b) Y. Li, M. B. Aggarwal, K. Ke, K. Nguyen, R. C. Spitale, *Biochemistry* **2018**, *57*, 1577–1581.
[11] a) M. Op de Beeck, A. Madder, *J. Am. Chem. Soc.* **2011**, *133*, 796–807; b) M. Op de Beeck, A. Madder, *J. Am. Chem. Soc.* **2012**, *134*, 10737–10740.
[12] M. J. Schmidt, D. Summerer, *Angew. Chem. Int. Ed.* **2013**, *52*, 4690–4693; *Angew. Chem.* **2013**, *125*, 4788–4791.
[13] D. P. Hari, B. König, *Chem. Commun.* **2014**, *50*, 6688–6699.
[14] H. Song, Y. Wang, R. Wang, X. Zhang, Y. Liu, G. Jia, P. R. Chen, *Cell Chem. Biol.* **2020**, *27*, 283–291.e286.
[15] V. A. K. Rathinam, Y. Zhao, F. Shao, *Nat. Immunol.* **2019**, *20*, 527–533.
[16] H. Kitamura, M. Ito, T. Yuasa, C. Kikuguchi, A. Hijikata, M. Takayama, Y. Kimura, R. Yokoyama, T. Kaji, O. Ohara, *Physiol. Genomics* **2008**, *33*, 121–132.
[17] T. Uehata, O. Takeuchi, *Cells* **2020**, *9*, 1701.
[18] N. Kelley, D. Jeltema, Y. Duan, Y. He, *Int. J. Mol. Sci.* **2019**, *20*, 3328.
[19] a) W. Liu, Z. Wang, L. Liu, Z. Yang, S. Liu, Z. Ma, Y. Liu, Y. Ma, L. Zhang, X. Zhang, M. Jiang, X. Cao, *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 23695–23706; b) H. Cui, S. Banerjee, S. Guo, N. Xie, J. Ge, D. Jiang, M. Zörnig, V. J. Thannickal, G. Liu, *JCI Insight* **2019**, *4*, e124522; c) M. R. Hadjicharalambous, M. A. Lindsay, *Non-coding RNA* **2019**, *5*, 34.
[20] P. Samir, S. Kesavardhana, D. M. Patmore, S. Gingras, R. K. S. Malireddi, R. Karki, C. S. Guy, B. Briard, D. E. Place, A. Bhattacharya, B. R. Sharma, A. Nourse, S. V. King, A. Pitre, A. R. Burton, S. Pelletier, R. J. Gilbertson, T. D. Kanneganti, *Nature* **2019**, *573*, 590–594.
[21] G. Arun, D. Aggarwal, D. L. Spector, *Non-coding RNA* **2020**, *6*, 22.
[22] a) A. Baldassarre, A. Paolini, S. P. Bruno, C. Felli, A. E. Tozzi, A. Masotti, *Epigenomics* **2020**, *12*, 1349–1361; b) Z. Li, P. D. Nagy, *RNA Biol.* **2011**, *8*, 305–315.
[23] D. Yang, J. L. Leibowitz, *Virus Res.* **2015**, *206*, 120–133.
[24] M. T. Khan, M. Irfan, H. Ahsan, A. Ahmed, A. C. Kaushik, A. S. Khan, S. Chinnasamy, A. Ali, D. Q. Wei, *Intervirology* **2021**, *64*, 55–68.
[25] S. Li, *Cells* **2019**, *8*, 0.
[26] R. E. Lloyd, *Virology* **2015**, *479–480*, 457–474.
[27] Q. Wan, D. Song, H. Li, M. L. He, *Signal Transduction Targeted Ther.* **2020**, *5*, 125.
[28] a) A. Mishra, S. Chanchal, M. Z. Ashraf, *TH Open* **2020**, *4*, e403–412; b) Z. Chen, Z. Li, X. Hu, F. Xie, S. Kuang, B. Zhan, W. Gao, X. Chen, S. Gao, Y. Li, Y. Wang, F. Qian, C. Ding, J. Gan, C. Ji, X. W. Xu, Z. Zhou, J. Huang, H. H. He, J. Li, *Adv. Sci.* **2020**, *7*, 2000532.
[29] A. Kawaguchi, K. Nagata, *EMBO J.* **2007**, *26*, 4566–4575.
[30] J. J. Knowlton, I. Fernández de Castro, A. W. Ashbrook, D. R. Gestaut, P. F. Zamora, J. A. Bauer, J. C. Forrest, J. Frydman, C. Risco, T. S. Dermody, *Nat. Microbiol.* **2018**, *3*, 481–493.
[31] a) N. Schmidt, C. A. Lareau, H. Keshishian, S. Ganskih, C. Schneider, T. Hennig, R. Melanson, S. Werner, Y. Wei, M. Zimmer, J. Ade, L. Kirschner, S. Zielinski, L. Dölken, E. S. Lander, N. Caliskan, U. Fischer, J. Vogel, S. A. Carr, J. Bodem, M. Munschauer, *Nat. Microbiol.* **2021**, *6*, 339–353; b) R. A. Flynn, J. A. Belk, Y. Qi, Y. Yasumoto, J. Wei, M. M. Alfajaro, Q. Shi, M. R. Mumbach, A. Limaye, P. C. DeWeirdt, C. O. Schmitz, K. R. Parker, E. Woo, H. Y. Chang, T. L. Horvath, J. E. Carette, C. R. Bertozzi, C. B. Wilen, A. T. Satpathy, *Cells* **2021**, *184*, 2394–2411.e2316.

Manuscript received: February 6, 2022
Accepted manuscript online: April 22, 2022
Version of record online: May 5, 2022