1	
2	SUPPLEMENTARY INFORMATION APPENDIX
3	
4	
5	Destabilisation of the human RED-SMU1 splicing complex
6	as a basis for host-directed anti-influenza therapy
7	
8	
9	Usama Ashraf <sup>1,2,3</sup> , Laura Tengo <sup>4</sup> , Laurent Le Corre <sup>5</sup> , Guillaume Fournier <sup>1,2,3</sup> , Patricia Busca <sup>5</sup> ,
10	Andrew A. McCarthy <sup>6</sup> , Marie-Anne Rameix-Welti <sup>7,8</sup> , Christine Gravier-Pelletier <sup>5</sup> , Rob W.H.
11	Ruigrok <sup>4</sup> , Yves Jacob <sup>1,2,3</sup> , Pierre-Olivier Vidalain <sup>9</sup> , Nicolas Pietrancosta <sup>9*</sup> , Thibaut Crépin <sup>4*</sup> ,
12	Nadia Naffakh <sup>1,2,3*</sup>
13	
14	
15	
16	
17	

#### **1 Materials and Methods**

2

#### 3 Cells

Human embryonic kidney (HEK-293T) and adenocarcinomic human alveolar basal epithelial
(A549) cells were obtained from Dr. M. Perricaudet (CNRS-IGR, Villejuif, France) and Pr. M.
Schwemmle (Freiburg University), respectively, and were grown in Dulbecco's modified
Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal calf serum (FCS).
Madin Darby Canine Kidney cells stably overexpressing the human 2,6-sialtransferase
(MDCK-SIAT1, (1)) were obtained from Pr. M. Matrosovich (Marburg University) and were
grown in minimum essential Eagle's medium (MEM) supplemented with 5% FCS.

11

#### 12 Viruses

The recombinant influenza A/WSN/33 (WSN) and WSN-PB2-Nanoluc viruses were 13 14 described previously (2). Human seasonal IAVs A/Bretagne/7608/2009 (H1N1pdm09) and A/Paris/1003/2012 (H3N2) were provided by the National Influenza Center at the Institut 15 16 Pasteur (Paris, France). The recombinant respiratory syncytial virus (RSV-Firefly) was described previously (3). The recombinant WSN-PB2-Nanoluc or RSV-Firefly virus were 17 18 used at a m.o.i. of 0.001 or 0.01 PFU/cell, respectively. Following viral adsorption for 2 hours, 19 the inoculum was replaced with medium supplemented with drug or DMSO alone, and luciferase activities were determined 24 hours later using Nano-Glo luciferase or Firefly 20 luciferase assay kits (Promega) and a Berthold Centro XS luminometer. The WSN, 21 22 A/Bretagne/7608/2009, or A/Centre/1003/2012 virus were used at a MOI of 0.0001 (WSN) or 23 0.001 PFU/cell. Following viral adsorption for 2 hours the inoculum was replaced with DMEM supplemented with drug or DMSO alone, and with 0.5 µg/ml TPCK-treated trypsin (Sigma-24 25 Aldrich). At 24, 48, and 72 hpi, cell supernatants were harvested and infectious titers 26 determined using a plaque assay on confluent monolayers of MDCK-SIAT1 cells.

- 27
- 28

2

#### **1** Plasmids for split-luciferase-based interaction assays

2 The pCI-Gluc2-RED and pCI-Gluc1-SMU1 plasmids were described previously (4). 3 Sequences encoding RED/SMU1 subdomains were amplified by PCR, and the resulting 4 amplicons were cloned in place of RED/SMU1 into the pCI-Gluc2-RED/pCI-Gluc1-SMU1 5 plasmids using the Notl and Mlul restriction sites. The mCherry-RED-[206-260] plasmid was 6 obtained by cloning the Cherry sequence in place of Gluc2 into the pCI-Gluc2-RED-[206-7 260] plasmid using the XhoI and NotI restriction sites. The pCI-Gluc1-FOS and pCI-Gluc2-8 JUN plasmids were obtained by Gateway cloning using the LR Clonase enzyme mix 9 (Thermo Fisher Scientific), donor plasmids from the human ORFeome v8.1 library 10 (http://horfdb.dfci.harvard.edu) and destination vectors described in (5). Point mutations and 11 deletions were introduced using the QuikChange II Site-Directed Mutagenesis Kit (Agilent). 12 All constructs were verified by Sanger sequencing. The sequences of the oligonucleotides 13 used for amplification, mutagenesis and sequencing can be provided upon request.

14

#### 15 Plasmids for protein expression in E. coli

16 The E. coli codon-optimized full-length human SMU1 and RED genes were synthesized by GeneArt (Life Technologies). The coding sequence of RED<sub>mid</sub> (residues 206 to 260) was 17 PCR amplified and cloned in pETM11 (EMBL) to express an N-terminal His-tagged (TEV 18 19 cleavable) recombinant construct. The coding sequence of SMU1<sub>Nter</sub> (residues 1 to 196) was 20 PCR amplified and cloned in both pETM11 and pACYC-LIC+ (Addgene) to express 21 respectively the N-terminal His-tagged (TEV cleavable) and non His-tagged recombinant 22 construct. The pACYC-LIC+- SMU1<sub>Nter</sub> plasmid was used to transform E. coli BL21 cells (Life 23 Technologies) and prepare the resulting chemical competent cells (E. coli BL21-RIL-SMU1<sub>Nter</sub>) grown in LB containing chloramphenicol (34 µg.mL<sup>-1</sup>). 24

25

#### 26 **Expression**, purification and characterization of full-length SMU1

The codon-optimized full-length human SMU1 gene was cloned in pFastBac-HtB (Thermo
Fisher Scientific) to express the N-terminal His-tagged (TEV cleavable) recombinant

construct in insect cells (High Five cells; Thermo Fisher Scientific) using the Bac-to-Bac 1 2 baculovirus system (Thermo Fisher Scientific) in standard conditions. Cells were 3 resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 1 mM β-4 mercaptoethanol) and cryolysed. After centrifugation, the supernatant was complemented with imidazole to reach 25 mM and load on a nickel affinity column (NiNTA, GIAGEN). The 5 6 resin was washed with a high salt buffer (50 mM Tris-HCl pH 8.0, 1 M NaCl, 1 mM β-7 mercaptoethanol, 25 mM imidazole). The purified protein was then eluted with 50 mM Tris-8 HCl pH 8.0, 200 mM NaCl, 1 mM β-mercaptoethanol, 300 mM imidazole. The protein was 9 dialyzed with TEV protease overnight against buffer without imidazole. The protein sample 10 was loaded on a second NiNTA, concentrated, and loaded on a Superdex S200 (GE 11 Healthcare) with the running buffer containing 20 mM Tris-HCl pH 8.0, 150 mM NaCl; 5 mM β-mercaptoethanol. The SEC-MALLS-RI analysis was performed using a Superdex<sup>™</sup> 200 12 increase column (10/300 GL, GE Healthcare). Sample injection and buffer flow was 13 controlled by a Hitachi L2130 pump, following the SEC column was a L-2400 UV detector 14 (Hitachi), Optilab T-rEX refractometer (Wyatt technologies) and a DAWN HELEOS-II multi 15 16 angle light scattering detector (Wyatt technologies). Prior to injection, column and system 17 were equilibrated in 5 to 10 column volumes of the running buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl; 5 mM β-mercaptoethanol). 50 µL of the protein sample concentrated at 3.5 18 19 mg/mL injection was injected at a constant flow rate of 0.5 mL/min. Accurate MALLS mass 20 prediction was performed with the Astra software (Wyatt Technologies).

21

#### 22 Reagents

Western blots were performed as described earlier (6). The membranes were incubated with primary antibodies directed against RED (Santa Cruz, sc-135485) or SMU1 (Santa Cruz, sc-100896), GAPDH (Thermo Fisher Scientific, MA5-15738), the Gaussia luciferase (New England Biolabs, E8923S), with peroxidase-conjugated Streptavidin (IBA) and peroxidaseconjugated secondary antibodies (GE Healthcare). Western blot signals were revealed using the enhanced chemiluminescent (ECL) reagent (Pierce). For intra-cellular protein cross1 linking, cell monolayers were washed twice with PBS and treated with disuccinimidyl 2 suberate (Thermo Fisher Scientific, #A39267) at the indicated concentrations for 30 min at 3 room temperature. The reaction was stopped by adding a quenching solution (Tris-HCl 1M, 4 pH7.5) for 15 mn at room temperature. Cell lysates were prepared using a RIPA buffer and 5 subjected to western blot analysis.

6 Cell viability was assessed using the CellTiter-Glo® One Solution Assay kit (Promega)
7 according to the manufacturer's instruction.

8

#### 9 Assessment of chemical database diversity

10 The diversity of our in-house chemical database was computed using the Calculate Diversity 11 Metrics protocol (Discovery Studio Modeling Environment, release 4.5; Dassault Systemes 12 BIOVIA: San Diego, 2015) and compared with the Prestwick and FDA approved databases. All three databases were similar in terms of diversity number of assemblies (LCPBT-DB: 13 0.341; FDA: 0.300; Prestwick: 0.468), diversity number of fingerprint features (LCPBT-DB: 14 15 8.1; FDA: 5.8; Prestwick: 11.2), and normalized fingerprint distances (LCPBT-DB: 0.910; 16 FDA: 0.906; Prestwick: 0.913). The number of assemblies is the number of unique Ring, 17 Chain, Bridge, or Murcko groups normalized by the number of molecules. Murcko groups are contiguous ring systems plus chains that link two or more rings, as defined by Bemis and 18 Murcko (7). The number of fingerprints is measured by the number of unique fingerprint 19 20 features of the selected fingerprint property normalized by the number of molecules. 21 Fingerprint distances report the minimum, maximum, and average fingerprint distance for all 22 pairs of ligands in the set. Chemical compounds were provided by the chemical library of 23 LCPBT (Laboratoire de Chimie, Biochimie Pharmacologiques et Toxicologiques - UMR8601), 24 and compound purity was >80%.

- 25
- 26
- 27
- 28
- 29

**1** Supplementary Figure Legends

Supplementary Figure S1. Disorder predictions on RED and SMU1 sequences. Disorder
profiles of the amino-acid sequence of (A) RED and (B) SMU1 as determined with D-score.
D-score is an algorithm to predict structured and disordered regions in proteins (8). The
predictions are based on 19 predictor web servers. The value 1.0 means that the protein is
fully ordered and and 0.0 means it is fully disordered.

8

2

#### 9 Supplementary Figure S2. X-ray structure of SMU1<sub>Nter</sub>.

10 The X-ray structures of  $SMU1_{Nter}$  (**A**) and  $RED_{mid}$ - $SMU1_{Nter}$  (**B**) have been superimposed 11 according to their dimerization interfaces. The two monomers of  $SMU1_{Nter}$  are coloured 12 yellow and bright orange.  $RED_{mid}$  is coloured red in panel B.

13

# Supplementary Figure S3. Common RED-SMU1 interface (interface II) between *H. sapiens* and *C. elegans* structures.

16 A. Structure of the RED-SMU1 interface within the C. elegans RED[194-221]-SMU1<sub>Nter</sub> 17 complex (9) compared with the corresponding interface (interface II) within the human RED<sub>mid</sub>-SMU1<sub>Nter</sub> complex (this study). The human structure was superimposed with C. 18 19 elegans structure (PDB code 5EN7) with a root mean square deviation (r.m.s.d.) of 0.97 Å 20 (112 C $\alpha$  of SMU1<sub>Nter</sub> aligned) using PDBeFold (10). The two X-ray structures are shown in 21 the same orientation and colour coded as in Figure 2. B-C. Partial sequence alignment of 22 RED (B) and SMU1 (C). Identical residues are white on a red background and similar 23 residues are red in a blue box. B. The boundaries and secondary structures of the RED 24 region present in the human structure (residues 206 to 260, Uniprot ID Q9UK43) and C. 25 elegans structure (residues 163 to 223, Uniprot ID Q9N4U5) are represented schematically 26 above and below the alignment, respectively. C. The boundaries and secondary structures of the SMU1<sub>Nter</sub> region present in the human structure (residues 1 to 196, Uniprot ID A0MNN4) 27 and C. elegans structure (residues 1 to 181, Uniprot ID G5EEG7) are represented 28 29 schematically above and below the alignment, respectively.

#### 1 Supplementary Figure S4. Dimerization status of the full-length SMU1 protein.

A. SEC-MALLS-RI analysis. The Size Exclusion Chromatography (SEC), MALLS (Multi 2 3 Angle Lase Light Scattering) and Refractive Index (RI) analyses was performed as described 4 in the SI Appendix Materials and Methods section. Experimental determination of the molar 5 mass of the full-length SMU1 protein is shown (the theoretical mass is 57.8 kDa for a SMU1 6 monomer). The red line indicates the molecular weight as estimated below the peak. The 7 panel was generated with Graphpad (Prism). B. Cross-linking of the SMU1 dimers in live 8 cells. Western blot analysis was performed on lysates from A549 (left panels) or HEK-293T 9 cells (rights panels) with or without pretreatment with the membrane permeable cross-linking 10 agents DSS at the indicated concentrations. The membrane was first probed with anti-SMU1 antibodies (upper panels), then with anti-GAPDH antibodies (lower panels). (-): DMSO 11 12 control.

13

#### 14 Supplementary Figure S5. Effect of RED<sub>mid</sub> overexpression

A. RED<sub>mid</sub>-SMU1 interaction in the absence or presence of the V216D mutation on RED<sub>mid</sub>. 15 16 HEK-293T cells were transfected with Gluc1-SMU1 together with the Gluc2-RED, -RED-V216D, -RED<sub>mid</sub> or RED<sub>mid-</sub>V216D plasmids. The normalized luciferase activities were 17 expressed as percentages relative to the activity measured in the presence of Gluc2-RED. 18 19 The data shown are the mean +/- SD of three independent experiments in triplicates. **B-F.** 20 Effect of RED<sub>mid</sub> overexpression. B. Effect on RED<sub>mid</sub>-SMU1 interaction. The split-luciferase RED-SMU1 interaction assay was performed as described in the Methods section, in the 21 presence of a co-transfected mCherry-RED<sub>mid</sub> or control mCherry-RED<sub>mid</sub>-V216D plasmids 22 (100 ng). The normalized luciferase activities were expressed as percentages relative to the 23 24 activity measured in the mock pCI-transfected control. The data shown are the mean +/- SD 25 of three independent experiments in triplicates. \*\* : p<0.01 (parametric paired t-test). C. HEK-26 293T cells were transfected with 100 ng of pCI, mCherry-RED<sub>mid</sub> or mCherry-RED<sub>mid</sub>-V216D 27 plasmids. At 48 hours post-transfection (hpt), the mCherry fluorescence signals were 28 observed using a fluorescence microscope (Olympus CKX41). The data shown are

representative of two independent experiments. D. Effect on the steady-state levels of the 1 2 endogenous RED and SMU1 proteins. HEK-293T cells were transfected with an expression 3 plasmid coding for RED<sub>mid</sub> or RED<sub>mid</sub>-V216D mutant fused to mCherry, or with the control 4 plasmids pCI / pCI-mCherry (mC). Total cell extracts were prepared at 48 hpt and were analysed by western blot, using antibodies specific for RED, SMU1 and GAPDH. One 5 representative experiment out of three is shown. E. Effect on IAV replication. HEK-293T cells 6 7 were transfected as in A with the indicated plasmids and infected at 48 hpt with the WSN-8 Nanoluc virus at a MOI of 0.001 PFU/cell. The luciferase activities measured at 24 hpi are 9 expressed as percentages relative to the pCI-transfected control. The data shown are the 10 mean +/- SD of three independent experiments in triplicates. F. Effect on NS1 mRNA splicing. HEK-293T cells were transfected as in A with the indicated plasmids and infected at 11 12 48 hpt with the WSN virus at a MOI of 5 PFU/cell. At 6 hpi, polyA+ RNA were extracted and 13 the levels of NS2 and NS1 mRNAs were analyzed by guantitative real-time PCR and normalized with respect to GAPDH mRNA levels. The NS2/NS1 mRNA ratios shown are the 14 mean +/- SD of three independent experiments in duplicates. \*\*p<0.01, \*\*\*p<0.001, unpaired 15 16 t-test (A-B), paired t-test (E-F).

17

#### 18 Supplementary Figure S6. Effect of incubation with LSP61 on cell monolayers.

19 Monolayers of A549 (**A**) or MDCK-SIAT (**B**) cells were incubated with the indicated 20 concentrations of LSP61 or DMSO alone. They were imaged at 10x magnification at the 21 onset of the experiment (T0), and following 48 hours of incubation (T48).

22

Supplementary Figure S7. Effects of compounds LSP641 and LSP61 on splicing of the viral M1 and PB2 mRNAs. A. M1 mRNA splicing. The same RNA samples that were tested for NS1 mRNA splicing in Figure 6F were used. The levels of M2 and M1 mRNAs were analyzed by RT-qPCR and normalized with respect to GAPDH mRNA levels, as described in (4). The M2/M1 mRNA ratios shown are the mean +/- SD of three independent experiments in duplicates. ns : non significant, unpaired t-test. **B.** PB2 mRNA splicing. The same RNA samples that were tested for NS1 and M1 splicing were used. The levels of PB2 and PB2-S1
mRNAs were analyzed by semi-quantitative RT-PCR as described in (11). The results of one
representative experiment out of three is shown.

4

Supplementary Figure S8. Refined spliceosome complex obtained from cryo-EM structure (PDB ID:509Z) superimposed to RED<sub>mid</sub>-SMU1 co-crystal (this study). Side view (A) and front view after a 90° rotation around the horizontal axis (B) of the spliceosome complex showing the pivotal position of SMU1 dimer (blue) and RED (red) in the spliceosome complex.

10

11

#### 1 References

Matrosovich M, Matrosovich T, Carr J, Roberts NA, & Klenk HD (2003)
 Overexpression of the alpha-2,6-sialyltransferase in MDCK cells increases influenza virus
 sensitivity to neuraminidase inhibitors. *J Virol* 77(15):8418-8425.

5 2. Diot C, *et al.* (2016) Influenza A Virus Polymerase Recruits the RNA Helicase DDX19 6 to Promote the Nuclear Export of Viral mRNAs. *Sci Rep* 6:33763.

Rameix-Welti MA, *et al.* (2014) Visualizing the replication of respiratory syncytial virus
 in cells and in living mice. *Nat Commun* 5:5104.

9 4. Fournier G, et al. (2014) Recruitment of RED-SMU1 complex by Influenza A Virus
10 RNA polymerase to control Viral mRNA splicing. *PLoS Pathog* 10(6):e1004164.

15. Cassonnet P, *et al.* (2011) Benchmarking a luciferase complementation assay for 12 detecting protein complexes. *Nat Methods* 8(12):990-992.

13 6. Moisy D, *et al.* (2012) HMGB1 protein binds to influenza virus nucleoprotein and 14 promotes viral replication. *Journal of virology* 86(17):9122-9133.

15 7. Bemis GW & Murcko MA (1996) The properties of known drugs. 1. Molecular 16 frameworks. *J Med Chem* 39(15):2887-2893.

17 8. Gerard FC, *et al.* (2009) Modular organization of rabies virus phosphoprotein. *J Mol* 18 *Biol* 388(5):978-996.

Ulrich AKC, Schulz JF, Kamprad A, Schutze T, & Wahl MC (2016) Structural Basis for
 the Functional Coupling of the Alternative Splicing Factors Smu1 and RED. *Structure* 24(5):762-773.

10. Krissinel E & Henrick K (2004) Secondary-structure matching (SSM), a new tool for
fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* 60(Pt
12 Pt 1):2256-2268.

11. Yamayoshi S, Watanabe M, Goto H, & Kawaoka Y (2016) Identification of a Novel
Viral Protein Expressed from the PB2 Segment of Influenza A Virus. *J Virol* 90(1):444-456.

27

#### **Supplementary table 1: Crystallographic data collection and structure refinement statistics.** Values in parentheses are for highest resolution shell.

Crystal parameters Native SMU1<sub>Nter</sub> Se-Met SMU1<sub>Nter</sub>. SMU1<sub>Nter</sub>-RED<sub>mid</sub> SMU1<sub>Nter</sub>-LSP641 Data collection Beamline BM14 ID29 ID29 ID30B 0.9791 0.97918 Wavelength (Å) 0.95373 0.97625 P21 P6522 Space group **P2**<sub>1</sub> P21 Cell a, b, c (Å) 40.2, 41.2, 123.6 40.4, 41.2, 123.9 128.0, 68.1, 145.3 72.2, 72.2, 183.2 α, β, γ (°) 90, 91.5, 90 90, 91.1, 90 90, 109.4, 90 90, 90, 120 Molecules per AU 2 2 16 1 Resolution (Å) 40 - 1.9 (1.95 - 1.9) 60 - 2.1 (2.2 - 2.1) 49 - 3.17 (3.25 - 3.17) 44 - 1.8 (1.88 - 1.8) Observed reflections 175 799 (6 523) 114 457 (7 606) 93 149 (7 046) 139 252 (16 734) Unique reflections 31 423 (1956) 41 829 (3 047) 38 425 (2 828) 26 383 (3 140) Completeness (%) 97.2 (81.5) 92.5 (64.8) 94.3 (97.2) 97.4 (97.7)  $R_{meas}$  (%) 6.4 (53.6) 8.4 (82.2) 14.4 (114.6) 5.1 (76.2) 22.3 (2.36) 10.2 (1.2) 16.86 (2.2) <l/o(l)> 6.9 (1.14) Model quality indicators FOM (centric/acentric) 0.13/0.41 R<sub>crvst</sub> (%) 17.6 (25.8) 25.6 (25.2) 16.0 (21.9)  $\mathsf{R}_{_{\mathrm{free}}}(\%)$ 21.4 (26.5) 27.7 (28.0) 22.0 (28.6) Mean B-factor (Å<sup>2</sup>) 25.7 132.3 43.9 rms deviations, bonds 0.007 / 1.49 0.008 / 1.71 0.007 / 1.45 (Å) / angles (°)









SMU1<sub>Nter</sub>

 $\mathsf{RED}_{\mathsf{mid}}\text{-}\mathsf{SMU1}_{\mathsf{Nter}}$ 







### A. A549 cells



### **B. MDCK-SIAT cells**







В



