SUPPLEMENTAL INFORMATION

A network of SMG-8, SMG-9 and SMG-1 C-terminal insertion domain regulates UPF1 substrate recruitment and phosphorylation

Aurélien Deniaud[1,2], Manikandan Karuppasamy[1,2], Thomas Bock[3], Simonas Masiulis[1,2], Karine Huard[1,2], Frédéric Garzoni[1,2], Kathrin Kerschgens[4,5], Matthias W. Hentze[5,6], Andreas E. Kulozik[4,5], Martin Beck[3], Gabriele Neu-Yilik[4,5] and Christiane Schaffitzel[1,2,7] *

[1] European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, 38042 Grenoble, France

[2] Unit of Virus Host-Cell Interactions, Univ. Grenoble Alpes-EMBL-CNRS, UMI 3265, 71 Avenue des Martyrs, 38042 Grenoble, France

[3] European Molecular Biology Laboratory, Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[4] Department of Pediatric Oncology, Hematology and Immunology, University of Heidelberg, Im Neuenheimer Feld 430, 69120 Heidelberg, Germany

[5] Molecular Medicine Partnership Unit, University of Heidelberg and European Molecular Biology Laboratory, Im Neuenheimer Feld 350, 69120 Heidelberg, Germany

[6] European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[7] School of Biochemistry, University of Bristol, Bristol, BS8 1TD, United Kingdom

* To whom correspondence should be addressed. Tel: 0033 (0)4 7620 7505; Fax: 0033 (0)4 7620 7199;

Email: *schaffitzel@embl.fr*

SUPPLEMENTARY MATERIAL AND METHODS

Chemical protein cross-linking coupled to mass spectrometry analysis

Cross-linking and proteolysis of the complexes

Cross-linking reactions were performed as described previously (47, 48). In brief, SMG-1-8-9 complexes (0.5 mg/ml) alone or with a three-fold excess of UPF1 were cross-linked by 0.5 mM isotope-labeled DSS (Disuccinimidyl-suberate, Creative Molecules Inc.) or DSG cross-linker (Disuccinimidyl-glutarate, Creative Molecules Inc.) in a buffer containing 20 mM Hepes pH 8.0, 100 mM KCl, 1 mM DTT, 5 % sucrose for 30 minutes at 37 ºC and 600 rpm. The crosslinking reaction was quenched using 100 mM ammonium bicarbonate at 35 °C and 600 rpm for 10 min. For tryptic digestion, cross-linked samples were further denatured in 4 M Urea / 100 mM ammonium bicarbonate containing buffer. Carbamidomethylation of cysteines was performed by 10 mM DTT at 37 °C for 30 min and treatment with 15 mM IAA at room temperature in the dark for 30 min. Samples were digested using the endoproteinase Lys-C (Wako) at a ratio of 1:100 (w/v) at 37 °C for 4 h. Samples were diluted to 1.5 M Urea and further digested by Trypsin (Promega) at a ratio of 1:50 (w/v) at 37 °C for 12 h. Digestion was stopped by acidification to 0.5% (v/v) TFA and the peptide mixture was desalted by C-18 Macro SpinColumns (Harvard Apparatus) according to the manufacturer's procedure. Purified peptides were dried in a vacuum concentrator and stored at -20 °C until further use.

Size exclusion chromatography enrichment of cross-linked peptides

Cross-linked peptides were enriched using gel-filtration (49). Purified peptides were resuspended in SEC buffer (30% (v/v) ACN in 0.1% (v/v) TFA) and fractionated using a Superdex Peptide PC 3.2/30 column (GE) on an Äktamicro LC system (GE) at a flow rate of 50 µL/min. According to 215 nm absorbance, fractions eluting between 0.9 and 1.4 mL were dried in a vacuum concentrator and reconstituted in 20-40 µL water containing 5% (v/v) ACN and 0.1% (v/v) FA.

LC-MS analysis of cross-linked peptides

Purified cross-linked peptide fractions were separated with a BEH300 C18 (75 µm x 250 mm, 1.7 µm) nanoAcquity UPLC column (Waters) using a stepwise 60 min gradient from 3% to 85% (v/v) acetonitrile in 0.1% (v/v) formic acid at a flow rate of 0.3 µl/min. The nano-LC system was connected to a LTQ-Orbitrap Velos Pro instrument (Thermo Scientific) that was operated in data-dependent mode. The collision induced dissociation method used one survey MS scan followed by up to 20 fragmentation scans (TOP20) of the most abundant ions

analyzed. Single and double charged ions were excluded from MS/MS analysis. Important MS parameters were: full MS: AGC = 10^6, maximum ion time = 500 ms, m/z range = 375-1600, resolution = 30,000 FWHM; MS2: AGC = 30,000, maximum ion time = 10 ms, minimum signal threshold = 100, dynamic exclusion time = 30 s, isolation width = 2 Da, normalized collision energy = 40, activation Q = 0.25. All fractions were analysed in technical duplicates.

Identification of cross-links from fragmentation spectra

MS data was processed using the xQuest/xProphet packages (48). MS raw data files were converted to mzXML using Mass Matrix file converter and searched against a manually created fasta database that contained the sequences of the cross-linked proteins using xQuest. Posterior probabilities were calculated by xProphet. The following filters were applied to the search results: false-discovery rate = 0.05, min delta score = 0.95, MS1 tolerance window -3 to 3 ppm, 1d score equal or larger than 20. All cross-links were manually checked to contain a convincing series of both, Xlink ions and common ions identified for both cross-linked peptide instances.

Analysis of conformational heterogeneity of the EM reconstructions

The conformational heterogeneity of the SMG-1-8-9-UPF1 reconstructions was analyzed by hypergeometrically stratified resampling (HGSR) using SPARX (50). The average and variance were obtained by sx3dvariability.py for the full dataset and for the classified volumes. The color code is based on the voxel values of the variance map.

SEC-MALLS experiments

Sample preparation

SMG-1 C-terminal insertion, SMG-8/SMG-9 complex and SMG-9 were expressed in insect cells. Cells were lysed in 50 mM Hepes pH 8.0, 300 mM NaCl, 5 mM β-mercaptoethanol and 10 mM imidazole supplemented with protease inhibitors (Roche). Hexahistidine-tagged proteins were purified using Ni-NTA affinity chromatography followed by ion exchange chromatography using a Q-XI column (GE-Healthcare) and then SP-sepharose column in the case of SMG-8-9. Finally, proteins were purified by size exclusion chromatography (Superose-6, GE-Healthcare) in 20 mM Hepes pH 8.0, 250 mM KCl and 2 mM β-mercaptoethanol for SMG-9 and the complex SMG-8-9 and in 20 mM Hepes pH 8, 150 mM NaCl and 2 mM β-mercaptoethanol for the SMG-1 C-terminal

insertion. The monomer peak was concentrated, flash-frozen in liquid nitrogen and stored at -80 °C.

SEC-MALLS

SEC was performed with a Superose-6 column (GE Healthcare) equilibrated with the final buffer of any protein studied. The column was calibrated with globular standard proteins. Separations were performed at 20°C with a flow rate of 0.5 mL.min-1. On-line multi-angle laser light scattering (MALLS) detection was performed with a DAWN-HELEOS II detector (Wyatt Technology Corp.) using a laser emitting at 690 nm, and protein concentration was measured on-line by the use of differential refractive index measurements using an Optilab T-rEX detector (Wyatt Technology Corp.) and a refractive index increment, dn/dc, of 0.185 mL.g-1. Weight-averaged molar masses were calculated using the ASTRA software (Wyatt Technology Corp.) (1). For size determination, the column was calibrated with proteins of known Stokes radius (2).

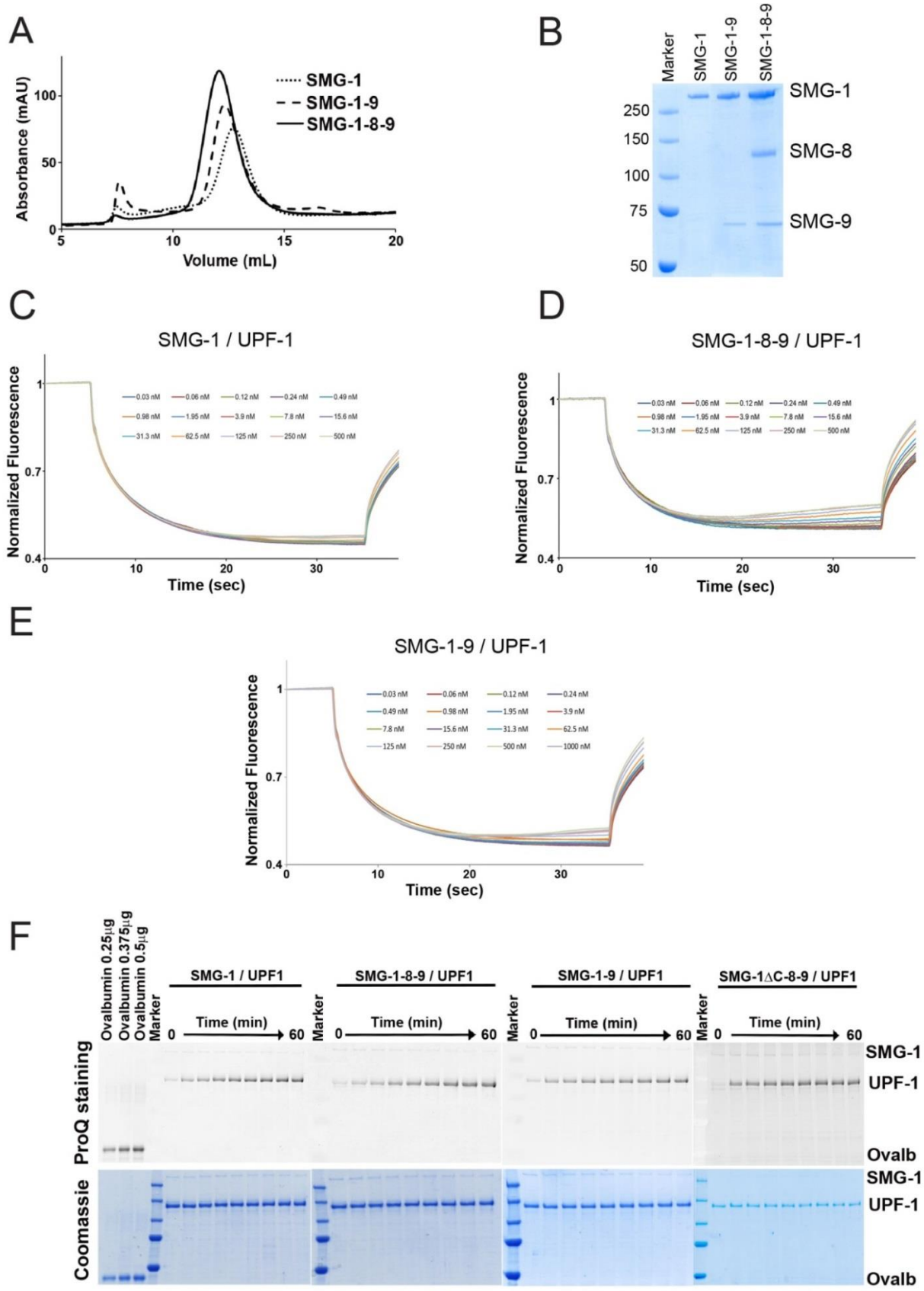

A
Absorbance (mAU)
Volume (mL)
SMG-1
SMG-1-9
SMG-1-8-9
B
Marker
SMG-1
SMG-1-9
SMG-1-8-9
250
150
100
75
50
SMG-1
SMG-8
SMG-9
C
SMG-1 / UPF-1
Normalized Fluorescence
Time (sec)
0.03 nM
0.06 nM
0.12 nM
0.24 nM
0.49 nM
0.98 nM
1.95 nM
3.9 nM
7.8 nM
15.6 nM
31.3 nM
62.5 nM
125 nM
250 nM
500 nM
D
SMG-1-8-9 / UPF-1
Normalized Fluorescence
Time (sec)
E
SMG-1-9 / UPF-1
Normalized Fluorescence
Time (sec)
1000 nM
F
Ovalbumin 0.25μg
Ovalbumin 0.375μg
Ovalbumin 0.5μg
Marker
SMG-1 / UPF1
SMG-1-8-9 / UPF1
SMG-1-9 / UPF1
SMG-1ΔC-8-9 / UPF1
Time (min)
ProQ staining
Coomassie
SMG-1
UPF-1
Ovalb

Figure S1: **Sample preparation, microscale thermophoresis experiments and *in vitro* phosphorylation raw data.**

(A) Size exclusion chromatograms for SMG-1, SMG-1-9 and SMG-1-8-9 complexes. **(B)** Coomassie-stained 8% SDS-PAGE of the final SMG-1, SMG-1-9 and SMG-1-8-9 samples. One representative series of raw data curves presenting the normalized fluorescence as a function of the time (sec) for all individual **(C)** SMG-1 kinase/ UPF1 conditions, **(D)** SMG-1-8-9 complex / UPF1 conditions and **(E)** SMG-1-9 complex / UPF1 conditions. **(F)** Kinetic *in vitro* phosphorylation assay of UPF1 by SMG-1, SMG-1-8-9, SMG-1-9 and SMG-1ΔC-8-9 kinase samples. UPF1 was phosphorylated by SMG-1, SMG-1-8-9, SMG-1-9 or SMG-1ΔC-8-9 for different periods of time. Ovalbumin was used as a standard. The reactions were separated on SDS-PAGE gels, analyzed by ProQ staining and then Coomassie staining. One representative gel ProQ-stained (above) and then Coomassie-stained (below) is shown for each experimental series.

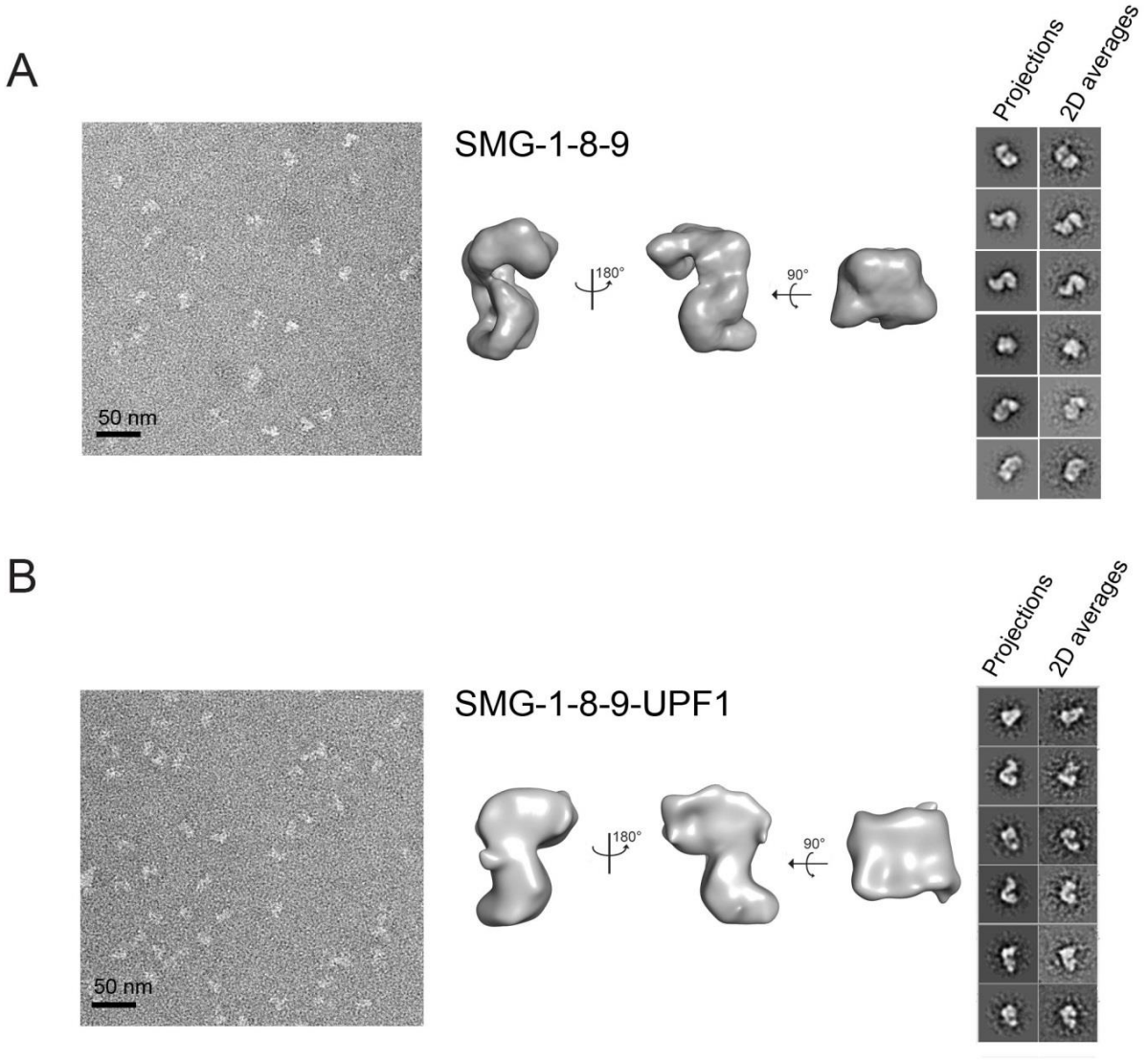
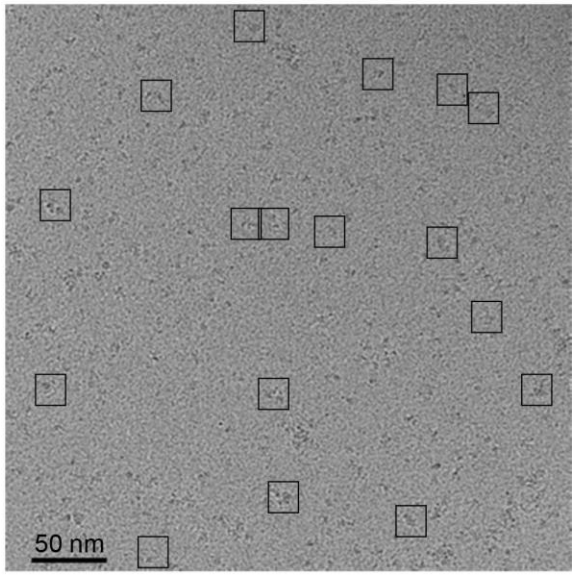

A
SMG-1-8-9
Projections
2D averages
50 nm
180°
90°
B
SMG-1-8-9-UPF1
Projections
2D averages
50 nm
180°
90°
C
D
2D class averages
10 nm

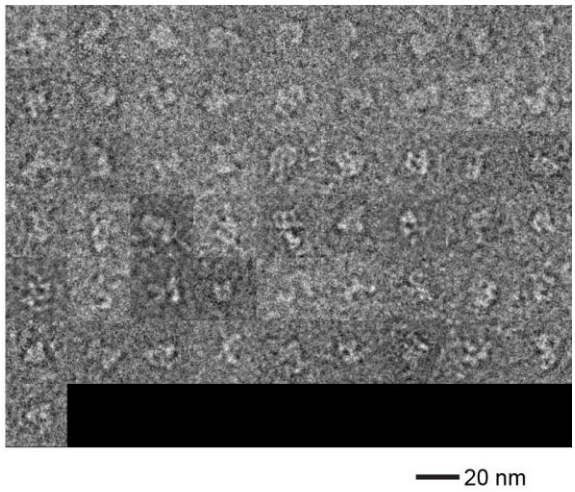
Figure S2: SMG-1-8-9 and SMG-1-8-9-UPF1 RCT reconstructions and UPF1 localization.

(A) Negative stain EM micrograph of the SMG-1-8-9 complex (left). The final SMG-1-8-9 negative stain volume is depicted in three different views (middle), and volume projections are compared with reference-free 2D class averages (right). **(B)** Negative stain EM micrograph of SMG-1-8-9-UPF1 sample (left). The SMG-1-8-9-UPF1 negative stain volume is shown in three different views (middle), and volume projections are depicted next to reference-free 2D class averages (right). The bar corresponds to 50 nm. **(C)** Superimposition of the two RCT volumes, SMG-1-8-9 (violet) and SMG-1-8-9-UPF1 (transparent light blue) in a side (left) and a top view (right). The extra-density present in the SMG-1-8-9-UPF1 head domain is encircled in red (position of UPF1) and blue (conformational changes in the head domain). **(D)** Antibody labeling of SMG-1-8-9-UPF1 complex using a polyclonal antibody against UPF1 residues 250-300 (CH domain). Negatively stained 2D class averages with the antibody bound to the SMG-1-8-9-UPF1 complex are shown. The antibody density is encircled in red.

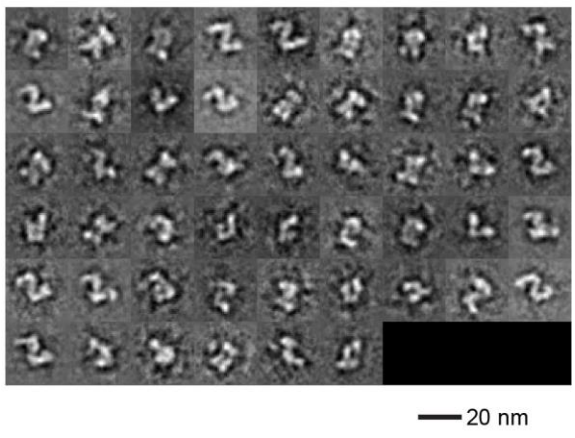
A



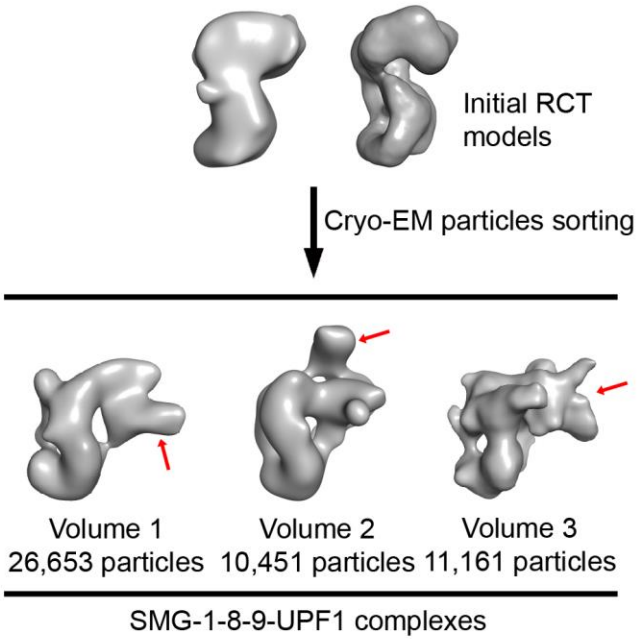
B



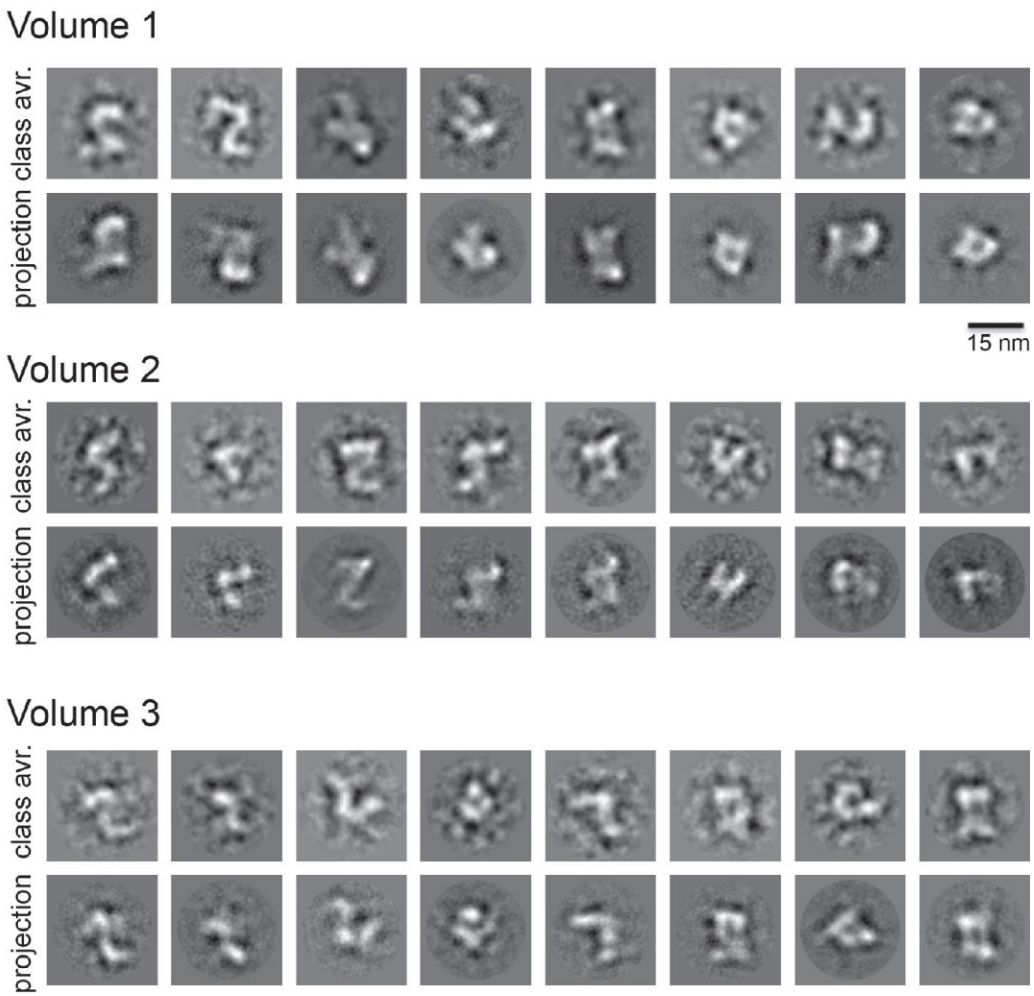
C



D


Initial RCT models
Cryo-EM particles sorting
Volume 1
26,653 particles
Volume 2
10,451 particles
Volume 3
11,161 particles
SMG-1-8-9-UPF1 complexes

E


Volume 1
projection class avr.
15 nm
Volume 2
projection class avr.
Volume 3
projection class avr.

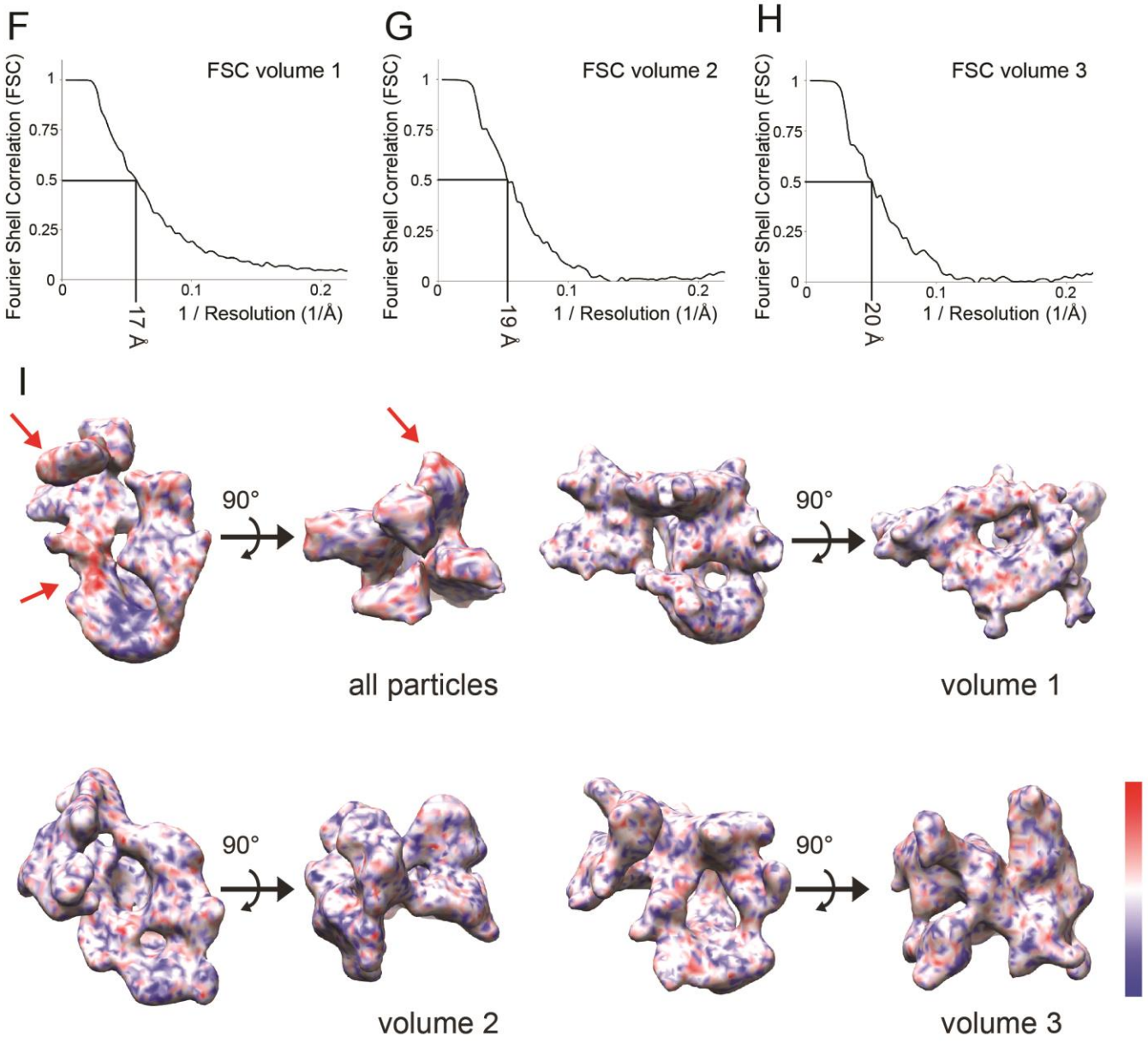


Figure S3: **SMG-1-8-9-UPF1 cryo-EM reconstructions.**

(A) Cryo-EM micrograph of SMG-1-8-9-UPF1 imaged under low dose conditions at 100 kV and a magnification of 59,000× with a Polara (FEI) microscope equipped with a 4k x 4k CCD camera. Representative individual particles are boxed. The bar corresponds to 50 nm. **(B)** Selected individual vitrified SMG-1-8-9-UPF1 particles from the complete dataset. The bar corresponds to 20 nm. **(C)** Selected reference-free 2D class averages of the complete cryo-EM dataset. The bar corresponds to 20 nm. **(D)** Sorting of 131,456 cryo-EM particles using the SMG-1-8-9 and SMG-1-8-9-UPF1 RCT reconstructions as initial models. Volumes 1, 2 and 3 showed defined additional density next to the head domain, highlighted with red arrows. These volumes were subsequently further sorted to clean up the data set and refined using iterative projection matching and maximum likelihood 3D refinement until convergence. Numbers indicate the final particle pool sizes. **(E)** For each volume, reference-free 2D class

averages from the sub-pool of particles that were used to calculate it (top) are compared to corresponding projections of the cryo-EM volume (below). **(F)** Fourier Shell Correlation curve of the SMG-1-8-9-UPF1 cryo-EM volume 1. The resolution of the map is 17 Å according to the FSC 0.5 criterion. **(G)** FSC curve of SMG-1-8-9-UPF1 volume 2 indicates a resolution of 19 Å (FSC 0.5 criterion). **(H)** FSC curve of SMG-1-8-9-UPF1 volume 3 indicates a resolution of 20 Å (FSC 0.5 criterion). **(I)** 3D Variance maps of SMG-1-8-9-UPF1 reconstructions calculated from hypergeometrically stratified resampled volumes in SPARX. The reconstruction from all 131,456 cryo-EM particles (upper left; resolution 26 Å) is compared to the volumes after computational sorting. Highly homogeneous regions (lower voxel variances) are colored in blue, flexible or heterogeneous regions (higher voxel variances) are displayed in red and highlighted with a red arrow. The surface representations were prepared in Chimera.

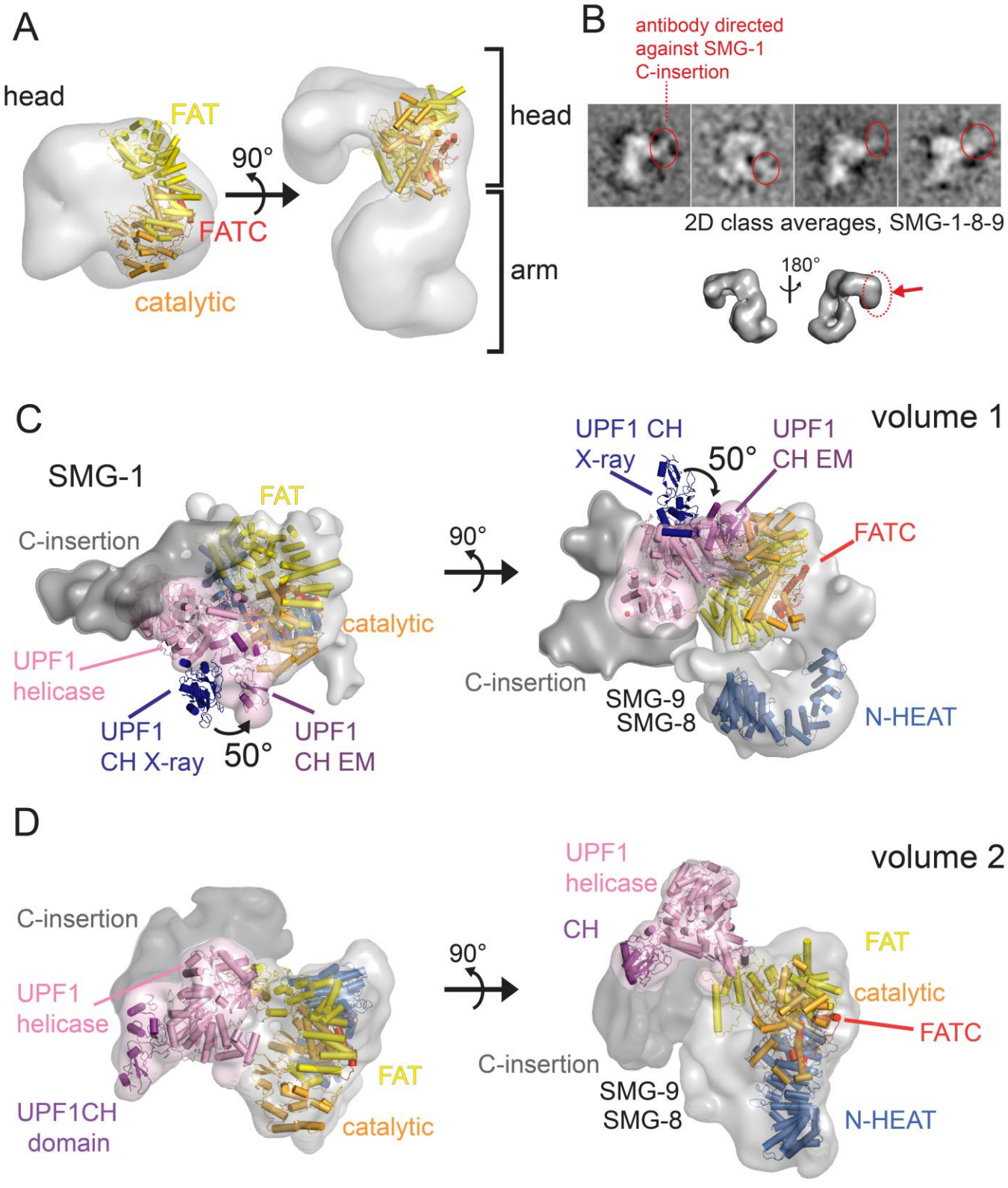

A
head
FAT
90°
FATC
catalytic
head
arm
B
antibody directed
against SMG-1
C-insertion
2D class averages, SMG-1-8-9
180°
C
SMG-1
FAT
C-insertion
catalytic
UPF1
helicase
UPF1
CH X-ray
50°
UPF1
CH EM
90°
volume 1
UPF1 CH
X-ray
50°
UPF1
CH EM
FATC
C-insertion
SMG-9
SMG-8
N-HEAT
D
C-insertion
UPF1
helicase
FAT
UPF1CH
domain
catalytic
90°
volume 2
UPF1
helicase
CH
FAT
catalytic
FATC
C-insertion
SMG-9
SMG-8
N-HEAT

Figure S4: **Fitting of homology models into the EM density.**

(A) Fitting of the homology model of the SMG-1 FAT-catalytic-FATC domain into the head domain of the negative stain 3D reconstruction of the SMG-1-8-9 complex (cc = 0.81, determined using Chimera). The placement is in agreement with previous mapping of the kinase domain using an antibody recognizing an epitope in the FATC domain (23,31) leaving unfilled density in the head domain. **(B)** Localization of the SMG-1 C-insertion domain in the SMG-1-8-9 complex using a monoclonal antibody and negative stain electron microscopy. The antibody density is encircled in red in the 2D class averages with the antibody bound to the SMG-1-8-9 complex head domain. Below: negative stain 3D reconstruction of the SMG-1-8-9 complex; the binding site of the antibody against the C-terminal insertion domain is highlighted in red. **(C)** Comparison of the UPF1 domain orientation in the crystal structure of UPF1 in complex with the UPF2 C-terminus (13) and of the UPF1 domain orientation in the EM quasi-atomic model where the CH domain has been fitted into the EM density next to the SMG-1 catalytic domain. For this comparison, the UPF1 helicase domains of the X-ray structure and the EM quasi-atomic model were superimposed. The CH domain as observed in the crystal structure (UPF1 CH X-ray) is depicted in dark blue; the model (UPF1 CH EM) was generated by rotation of the CH domain by ~50° towards the kinase catalytic domain as indicated by the black arrow. **(D)** Model of the alternative conformation of the SMG-1-8-9-UPF1 complex in volume 2. The structure and model are shown in a top (left) and a side view (right). The EM density corresponding to UPF1 and the SMG-1 C-insertion domain are shown in transparent pink and dark grey, respectively, in panels C and D. The atomic models for UPF1 CH domain and helicase are colored purple and light pink. The homology model of SMG-1 N-terminal HEAT repeats is depicted in blue, the FAT domain in yellow, the catalytic domain in orange, the FATC domain in red.

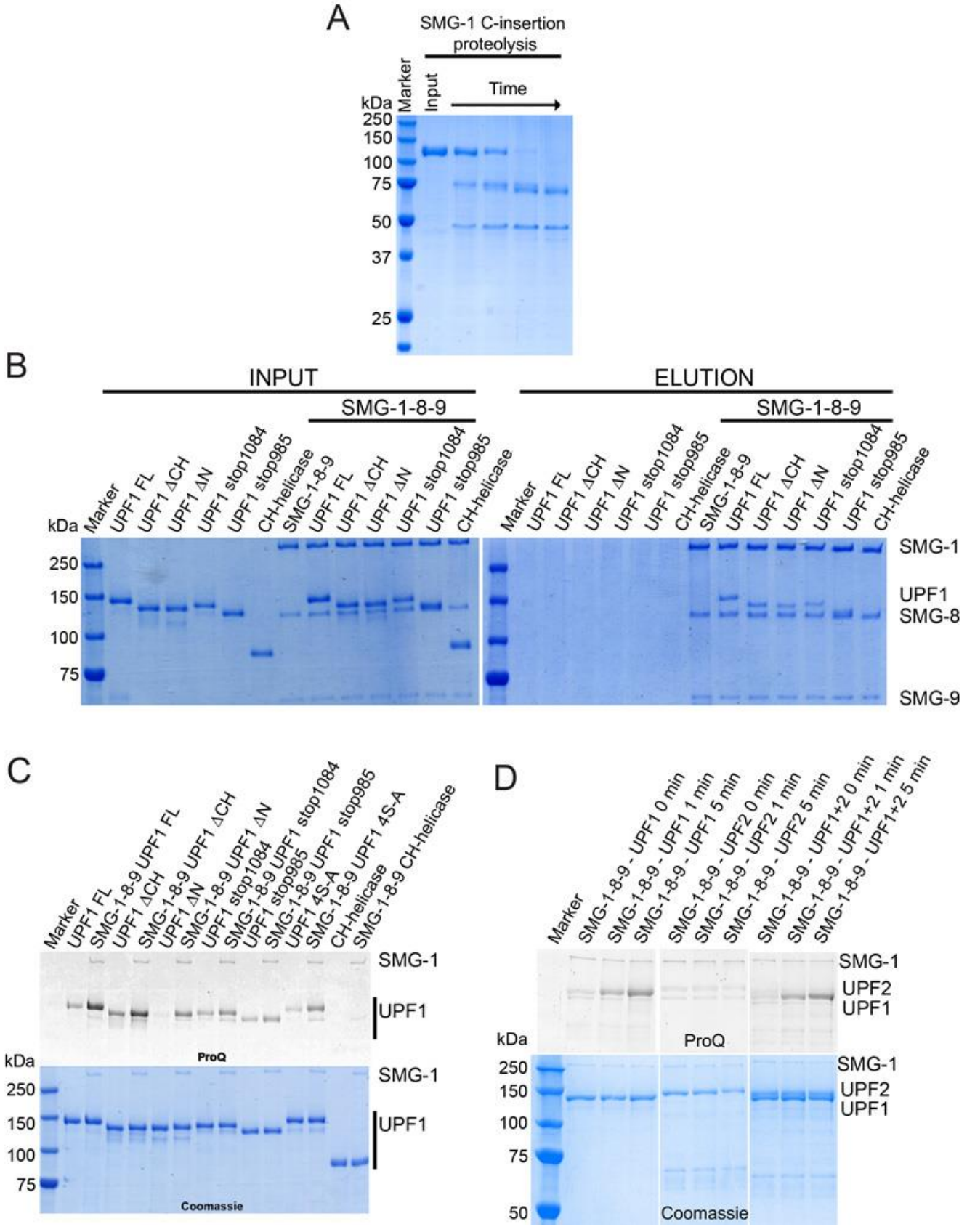

A
SMG-1 C-insertion proteolysis
Marker
Input
Time
kDa
250
150
100
75
50
37
25
B
INPUT
ELUTION
SMG-1-8-9
Marker
UPF1 FL
UPF1 ΔCH
UPF1 ΔN
UPF1 stop1084
UPF1 stop985
CH-helicase
SMG-1-8-9
kDa
250
150
100
75
SMG-1
UPF1
SMG-8
SMG-9
C
Marker
UPF1 FL
SMG-1-8-9 UPF1 FL
UPF1 ΔCH
SMG-1-8-9 UPF1 ΔCH
UPF1 ΔN
SMG-1-8-9 UPF1 ΔN
UPF1 stop1084
SMG-1-8-9 UPF1 stop1084
UPF1 stop985
SMG-1-8-9 UPF1 stop985
UPF1 4S-A
SMG-1-8-9 UPF1 4S-A
CH-helicase
SMG-1-8-9 CH-helicase
SMG-1
UPF1
ProQ
kDa
250
150
100
75
Coomassie
D
Marker
SMG-1-8-9 - UPF1 0 min
SMG-1-8-9 - UPF1 1 min
SMG-1-8-9 - UPF1 5 min
SMG-1-8-9 - UPF2 0 min
SMG-1-8-9 - UPF2 1 min
SMG-1-8-9 - UPF2 5 min
SMG-1-8-9 - UPF1+2 0 min
SMG-1-8-9 - UPF1+2 1 min
SMG-1-8-9 - UPF1+2 5 min
SMG-1
UPF2
UPF1
ProQ
kDa
250
150
100
75
50
Coomassie

Figure S5: **Biochemical analyses of SMG-1, SMG-1-8-9-UPF1 and SMG-1-8-9-UPF1-UPF2 complexes.** **(A)** Trypsinolysis of the SMG-1 C-insertion domain for 120 minutes at room temperature resulting in stable domains of ~70 kDa and 45 kDa. **(B)** Streptavidin pull-down of UPF1 deletion constructs. The Coomassie-stained gel on the left shows the input samples, UPF1 deletion constructs alone (control) and mixed with SMG-1-8-9. The SBP-tag of SMG-1 binds to the streptavidin beads. The gel on the right shows the eluate fractions, confirming that the UPF1 constructs do not interact with the streptavidin beads. **(C)** UPF1 *in vitro* phosphorylation by SMG-1-8-9. Samples of the different UPF1 constructs were incubated with SMG-1-8-9 for 15 minutes, separated on a SDS-PAGE gel and analyzed by ProQ staining (upper panel) and Coomassie staining (lower panel). 'FL' stands for full-length. **(D)** UPF1 *in vitro* phosphorylation by SMG-1-8-9 in the presence of UPF2. Samples of UPF1 alone or in the presence of a stoichiometric amounts of UPF2 and control samples with UPF2 alone were incubated with SMG-1-8-9 for 0, 1 or 5 minutes, separated on SDS-PAGE gel and analyzed by ProQ staining (upper panel) and Coomassie staining (lower panel).

Proteins	Molecular weight of expressed protein (kDa)	Molecular weight determined by MALLS (kDa)	Hydrodynamic radius (nm)
SMG-1	416	-	8.2
SMG-1 C-terminal insertion	121	118.9	5.4
SMG-1-9	477	-	9.4
SMG-1-8-9	590	-	9.4
SMG-9	61	63.6	5.3
SMG-8-9	174	171.2	6.9
UPF1	127	128.5	5.2

Table S1. Theoretical molecular weight of different proteins and complexes purified. Molecular weight and hydrodynamic radii of the different samples determined by size exclusion chromatography and multiple angle laser light scattering (MALLS). The dash stands for not determined.

SUPPLEMENTARY REFERENCES

47. Bui, K.H., von Appen, A., DiGuilio, A.L., Ori, A., Sparks, L., Mackmull, M.T., Bock, T., Hagen, W., Andres-Pons, A., Glavy, J.S. *et al.* (2013) Integrated structural analysis of the human nuclear pore complex scaffold. *Cell*, **155**, 1233-1243.

48. Walzthoeni, T., Claassen, M., Leitner, A., Herzog, F., Bohn, S., Forster, F., Beck, M. and Aebersold, R. (2012) False discovery rate estimation for cross-linked peptides identified by mass spectrometry. *Nat. Methods*, **9**, 901-903.

49. Leitner, A., Reischl, R., Walzthoeni, T., Herzog, F., Bohn, S., Forster, F. and Aebersold, R. (2012) Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell Proteomics*, **11**, M111 014126.

50. Penczek, P.A., Kimmel, M. & Spahn, C.M. Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. *Structure* **19**, 1582-1590 (2011).