

**Supplementary Information for**

**Structural basis for the inactivation of cytosolic DNA sensing by the vaccinia virus**

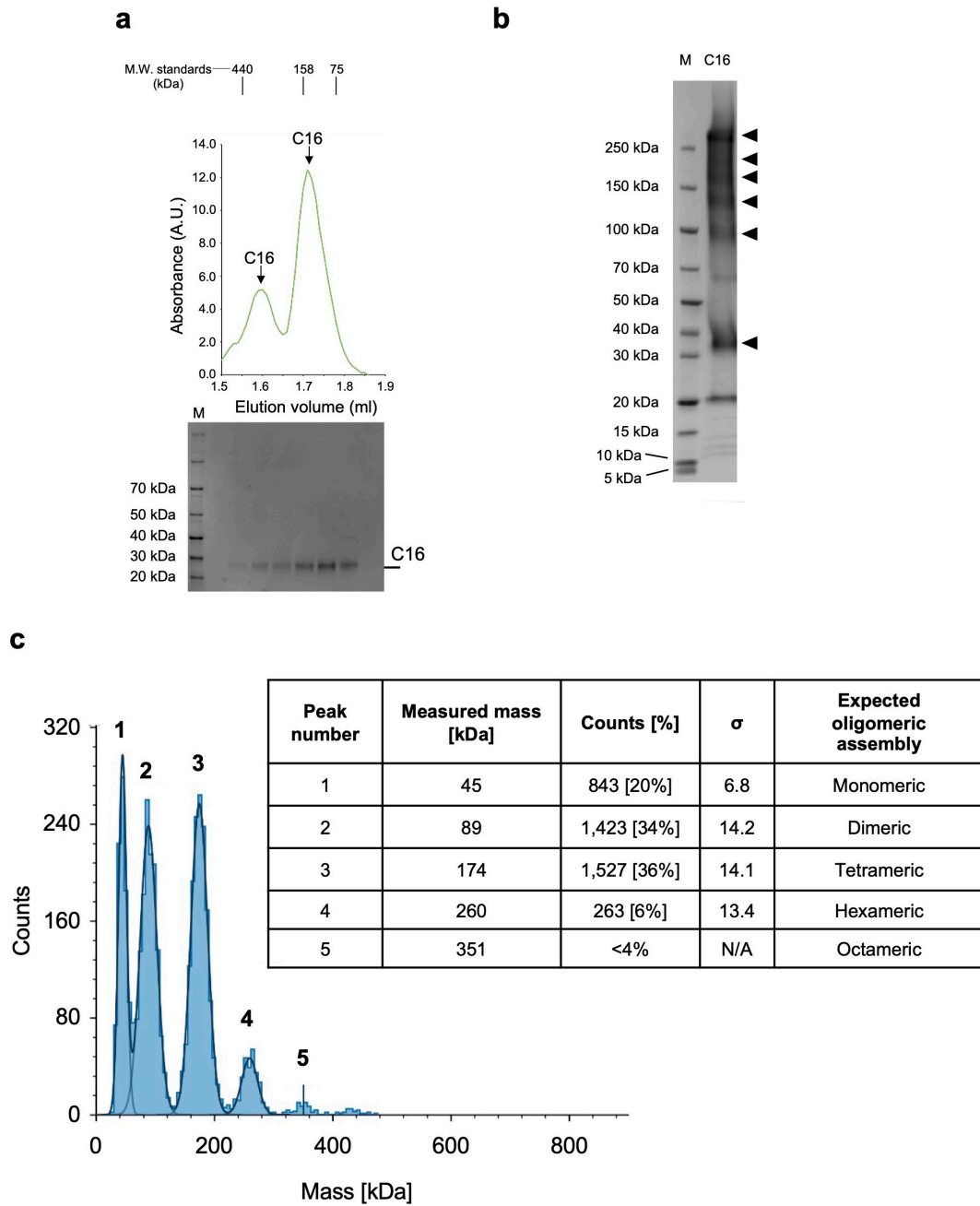
Angel Rivera-Calzada<sup>1&</sup>, Raquel Arribas-Bosacoma<sup>2,3&</sup>, Alba Ruiz-Ramos<sup>1</sup>, Paloma Escudero-Bravo<sup>1</sup>, Jasminka Boskovic<sup>1</sup>, Rafael Fernandez-Leiro<sup>1</sup>, Antony W. Oliver<sup>2</sup>, Laurence H. Pearl<sup>2,3\*</sup> and Oscar Llorca<sup>1\*</sup>

1. Spanish National Cancer Research Centre (CNIO), Melchor Fernández Almagro 3, 28029 Madrid, Spain
2. Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, Brighton BN1 9RQ, UK
3. Division of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW1E 6BT, UK

& These authors contributed equally

\* Correspondence to: LHP - [laurence.pearl@sussex.ac.uk](mailto:laurence.pearl@sussex.ac.uk); OL - [olorca@cnio.es](mailto:olorca@cnio.es)

## Supplementary Figures



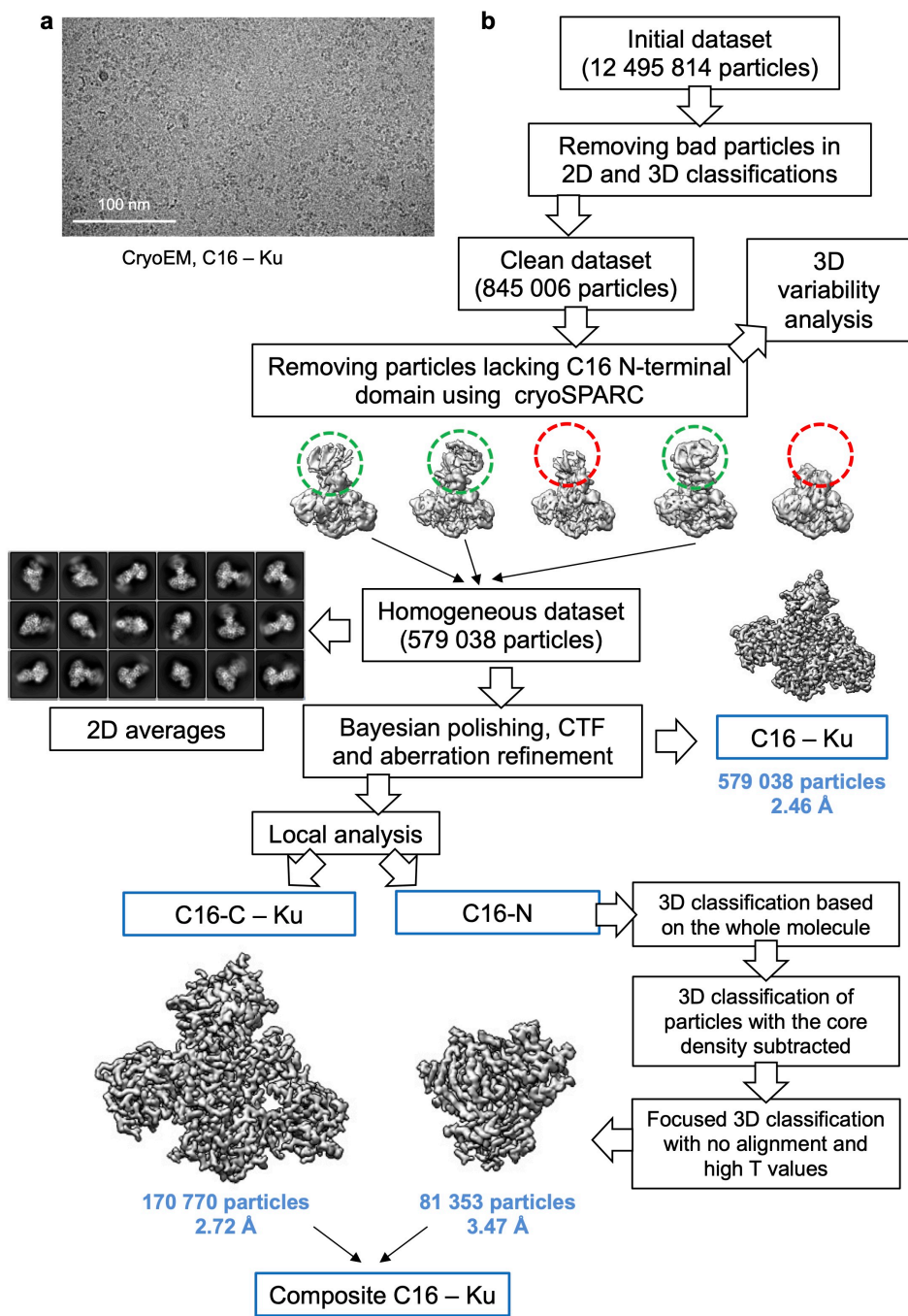
**Supplementary Fig. 1. Purification and oligomeric state of C16**

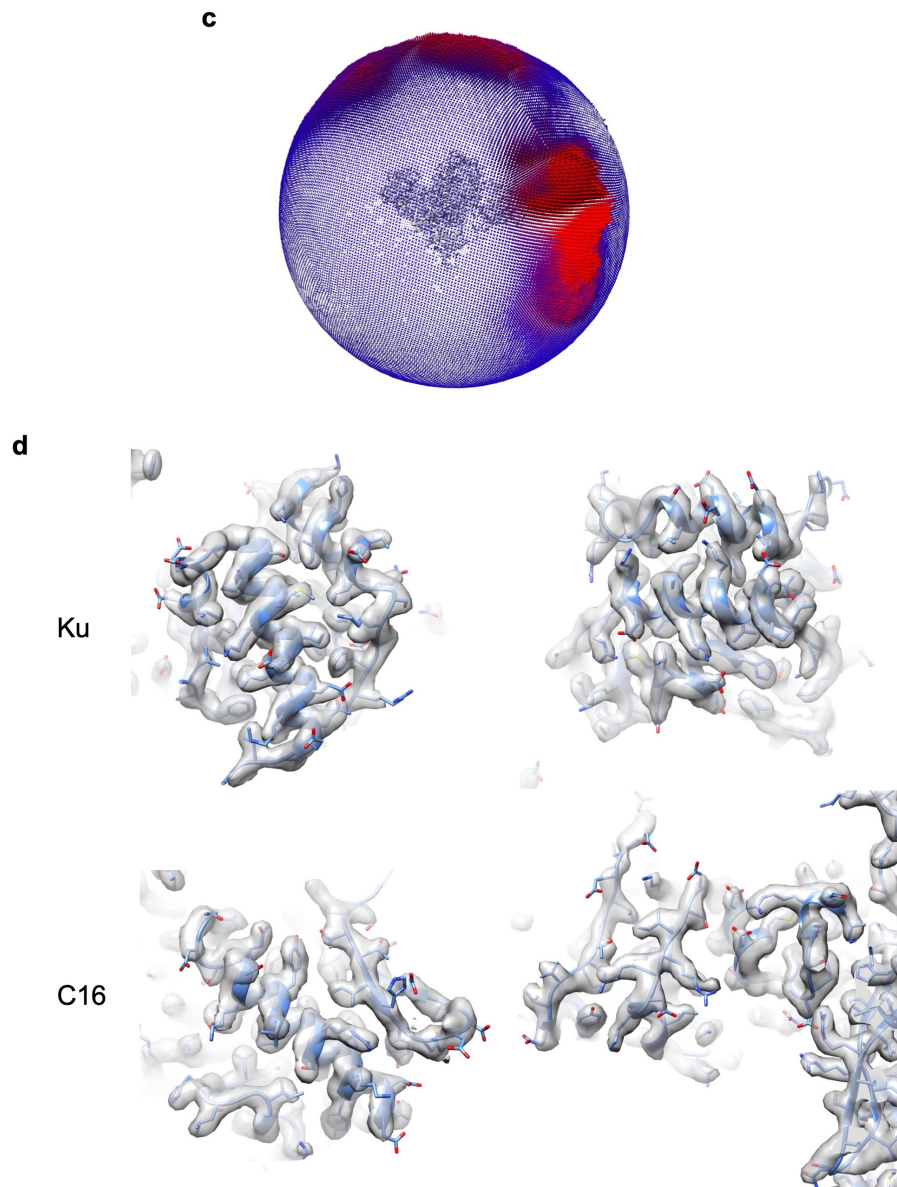
**a** SEC of purified C16. Selected fractions were analysed by SDS-PAGE (bottom panel). The elution volume for the peaks of standard proteins Ferritin (440 kDa), Aldolase (158 kDa) and Conalbumin (75 kDa) are indicated at the top of the chromatogram (M.W. stands for Molecular

Weight). SEC of C16 was repeated at least five times with similar results. Molecular weight marker (M) are shown.

**b** Purified C16 protein exhibits a tendency to oligomerize when analysed by SDS-PAGE under less-stringent denaturing conditions by avoiding boiling the mix of protein and loading buffer. Lanes displayed are molecular weight marker (M) and purified C16. Triangles indicate the different oligomers formed by C16. This was repeated at least three times with similar results.

**c** Mass distribution histogram obtained using mass photometry for the purified C16 protein. Details for each of the peaks are shown in the accompanying table. Purified C16 forms different oligomeric assemblies as observed during its purification and by cryoEM.  $\sigma$  corresponds to the standard deviation of the fit for each peak.





**Supplementary Fig. 2. CryoEM and image processing workflow for the C16 – Ku complex**

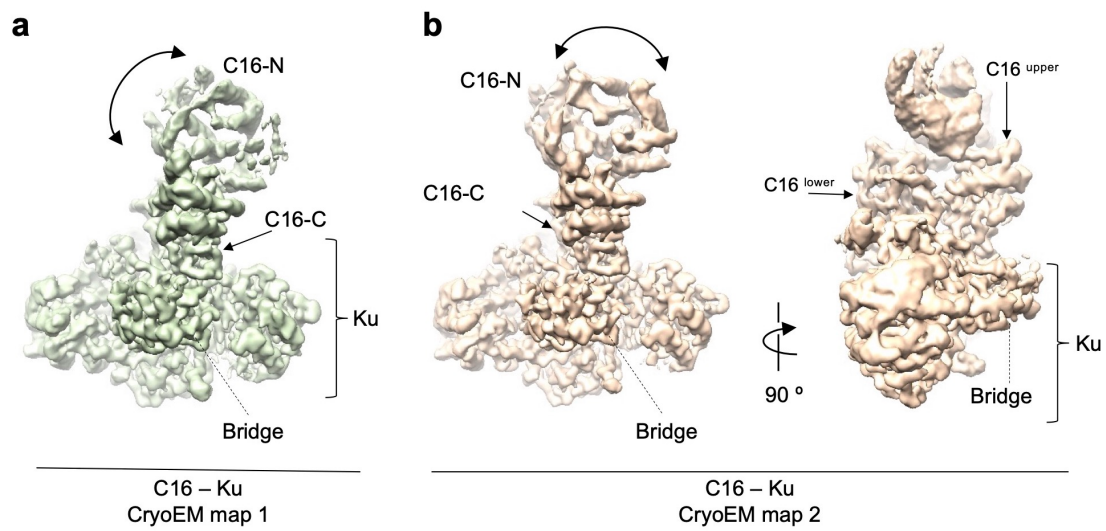
**a** A representative cryoEM field of a total of 13,216 movies collected (see Supplementary Table 1 for a complete list of cryoEM data collection parameters), showing individual particles of the C16 – Ku complex.

**b** Image processing workflow of the cryoEM data for the C16 – Ku complex. Image processing and classification strategies for the cryo-EM data of the C16 – Ku complex. Tools available in Relion 4.0<sup>1,2</sup> and cryoSPARC<sup>3</sup> were used to obtain homogeneous groups of particles that could be refined to high-resolution. cryoSPARC was used to reveal classes of particles in which the N-terminal domain of C16 was absent or poorly defined (encircled in red). The region corresponding to C16 C-terminal domains bound to Ku could be refined to high-resolution (C16 – Ku and C16-

C – Ku maps). The C16-C – Ku map, albeit having a worse overall resolution value than the C16 – Ku map, showed improved detail at two regions: the bridge that encircles the DNA binding region in Ku, and the C-terminal domain of the C16 molecule that protrudes in the DNA binding region. N-terminal domains in C16 are flexibly attached to the rest of the molecule. Therefore, we designed a strategy that combined 3D classifications using masks of the whole molecule, but also 3D classifications using density-subtracted images that only contain density corresponding to C16-N. Finally focused 3D classification combined with high T values revealed a subset that was refined to 3.5 Å resolution (C16-N map). The C16-C – Ku and C16-N maps were merged into a single composite C16 – Ku map. Representative 2D averages are shown.

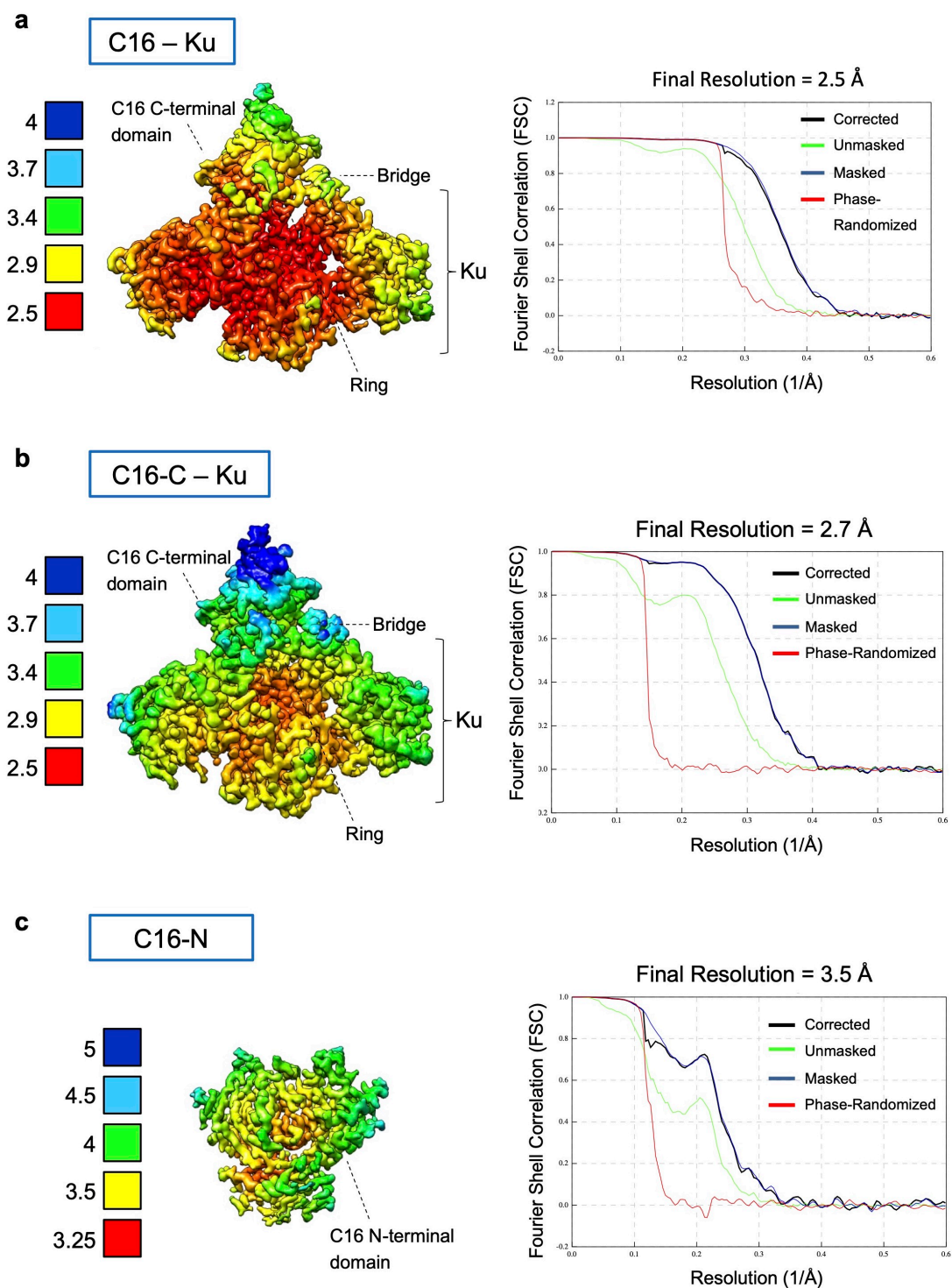
**c** Angular distribution of particles used in the final reconstruction for the C16-C – Ku complex.

**d** Representative regions of the cryoEM map of the C16-C – Ku complex shown as a transparency, with the structural model, in blue colour, fitted. Top panels show regions of the Ku molecule in the complex. Bottom panels show regions of the C16 molecule in the complex.



**Supplementary Fig. 3. Flexibility of the C16 – Ku complex**

**a, b** Two selected cryoEM volumes of the C16 – Ku complex, in green and light orange colour, highlighting the flexibility of the C16 N-terminal region. For one of the volumes, two views are shown and the two C16 monomers labelled (right panel).

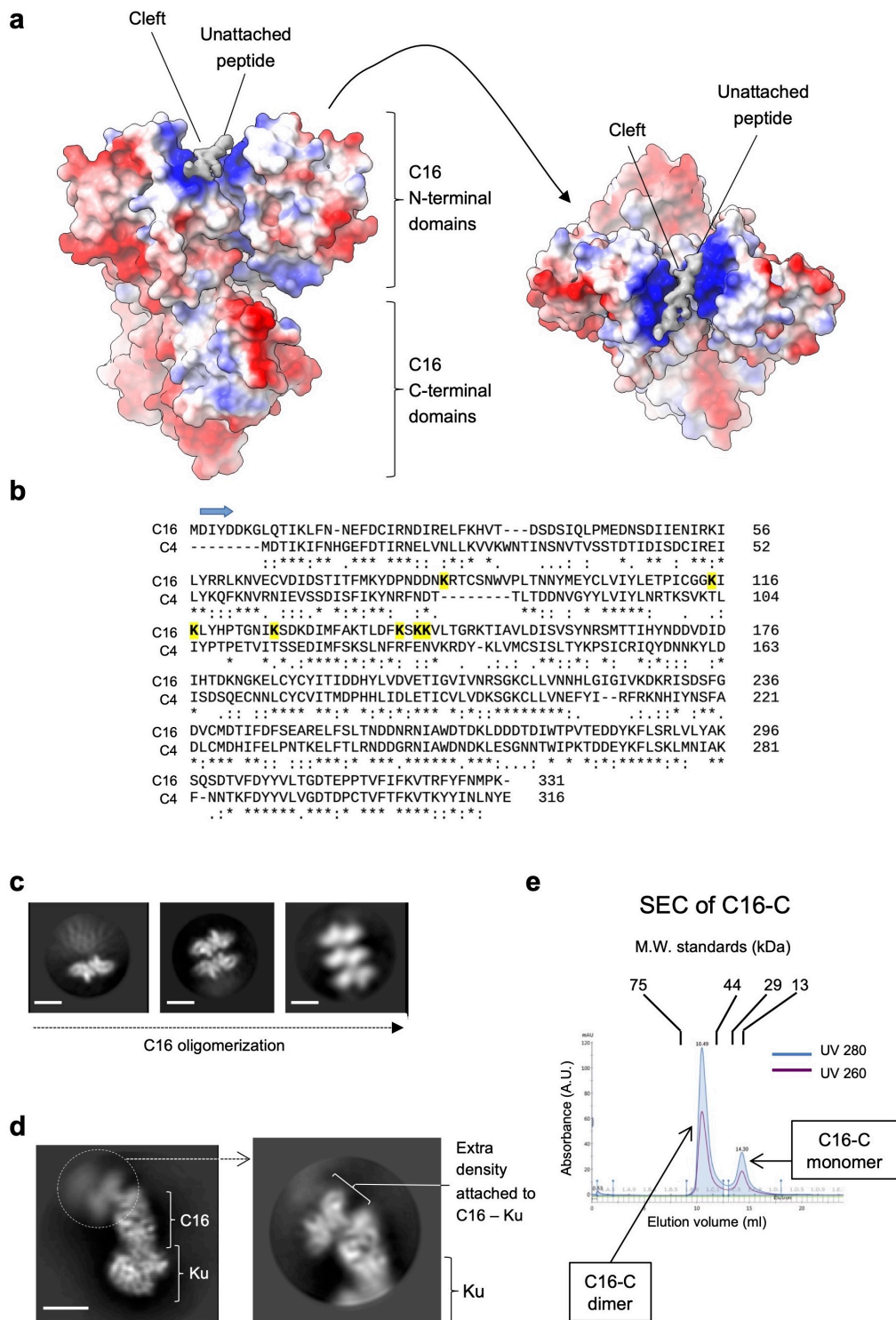


**Supplementary Fig. 4. Resolution estimates of the cryoEM maps**

**a** Right panel - FSC curves for the estimation of the average resolution of the cryoEM map C16 – Ku. Left panel - Local resolution map estimate for the cryoEM map of C16 – Ku.

**b** As in a but for the cryoEM map C16-C – Ku.

**c** As in a but for the cryoEM map C16-N.



**Supplementary Fig. 5. C16 structure and oligomerization**

**a** Coulombic electrostatic potential of the C16 homodimer. The N-terminal domains form a basic cleft where we identified an unattached density not accounted by the C16 molecules forming the

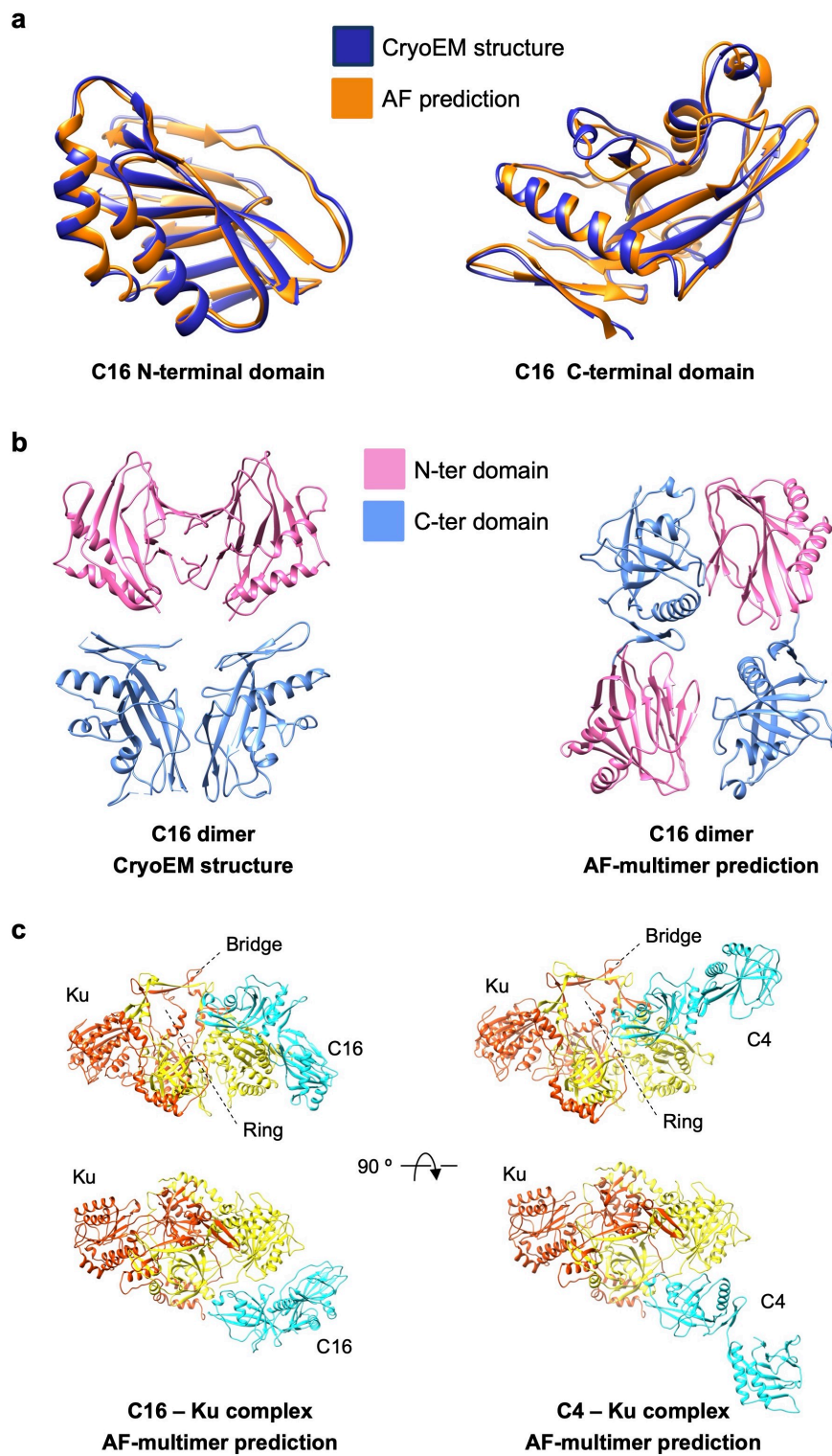
dimer. This density could come from an adjacent C16 dimer as observed in 2D classes corresponding to the C16 – Ku complex that contain additional C16 molecules bound to it.

**b** Sequence alignment of C16 and C4. Both sequences exhibit a high-degree of identity. In C16 residues at the N-terminus form a  $\beta$ -strand that is absent in C4, and it is indicated with a blue arrow. Lysines located at the basic cleft (see panel a) are highlighted in bold and yellow. Alignment was performed using Clustal Omega (1.2.4)(<https://www.ebi.ac.uk/Tools/msa/clustalo/>)<sup>4</sup>.

**c** 2D averages of several oligomeric species for C16. Scale bar corresponds to 5 nm.

**d** 2D averages of C16 – Ku complexes where C16 appears to oligomerize as part of a complex between Ku and C16. Scale bar corresponds to 5 nm.

**e** C16-C can dimerise in the absence of the N-terminal domain. Superdex 75 10/300 GL SEC profile for cterm-C16, which comprises residues from 157-331 and has an expected molecular weight of 23 kDa for a monomer and 46 kDa for dimer. The elution volumes corresponding to standard proteins Conalbumin (75 kDa), Ovalbumin (44 kDa), Carbonic Anhydrase (29 kDa) and Ribonuclease A (13.7 kDa) are indicated on top of the chromatogram.

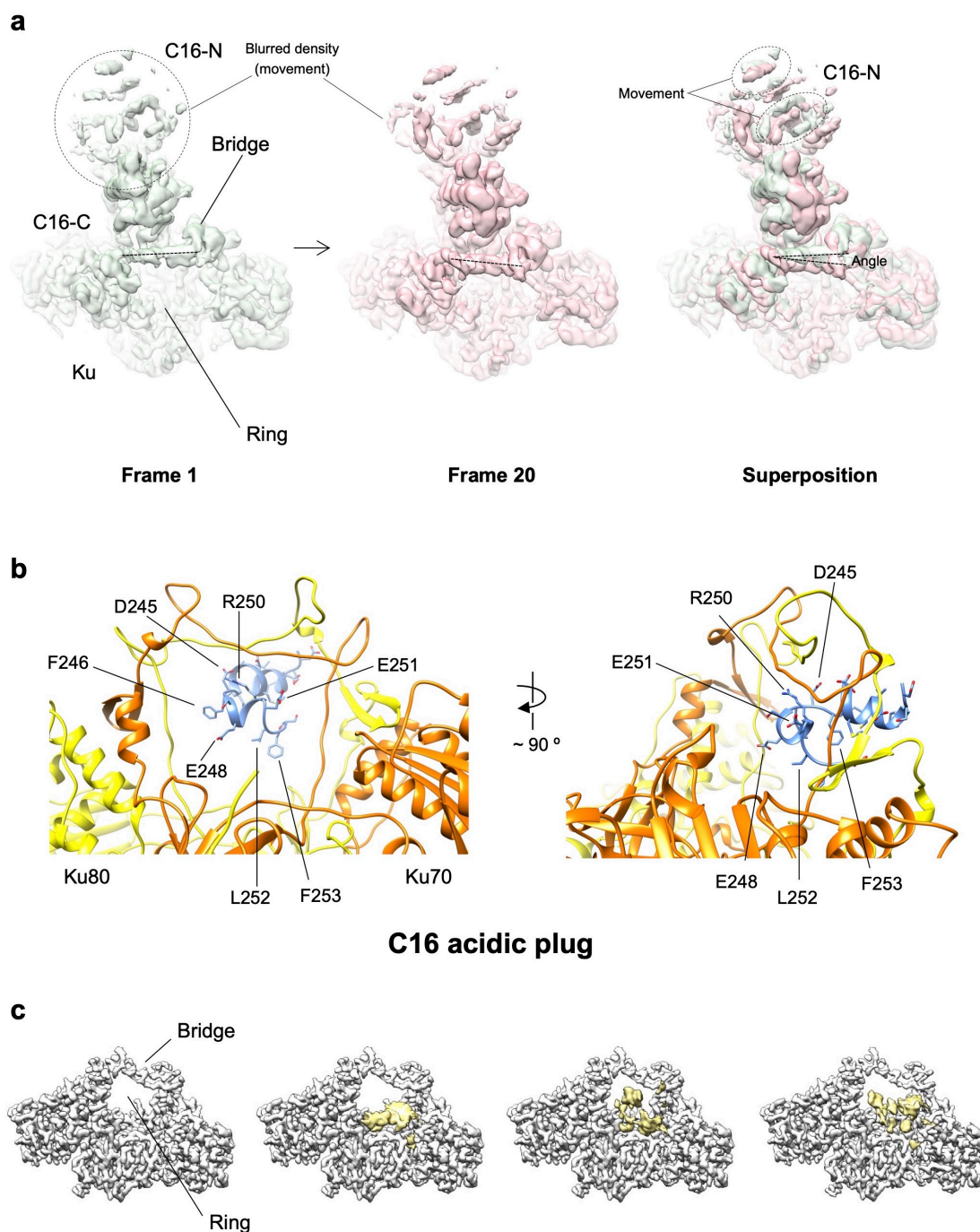


**Supplementary Fig. 6. Comparison between AF prediction and experimental structures**

**a** Comparison between the cryoEM structure of C16 N-terminal and C-terminal domains (blue colour) and a representative top-ranked AlphaFold (AF) prediction<sup>5</sup> (orange colour). Each domain is analysed separately because they are flexibly connected.

**b** Comparison between the cryoEM structure and AF-multimer (<https://www.deepmind.com/publications/protein-complex-prediction-with-alfafold-multimer>)<sup>6</sup> predictions of C16 dimers. For the cryoEM structure, N- and C-terminal domains (coloured in pink and light blue respectively) are showed separately because we find several relative orientations and also to highlight the dimerization interface for each domain. All top-ranked AF predictions suggest the dimerization between the N-terminal and C-terminal domain, and the figure shows a representative top-ranked AF.

**c** AF-multimer<sup>6</sup> predictions of C16 – Ku and C4 – Ku complexes. Predictions using 2 copies of C4 and C16 to allow prediction of dimers did not generate any prediction where C4 or C16 contacted Ku. When predictions were launched using only 1 chain of C4 and C16, some of the top-ranked solutions positioned the proteins in the vicinity of Ku, but none reproduced the interactions determined experimentally. Figure shows a representative top-ranked prediction for each protein.



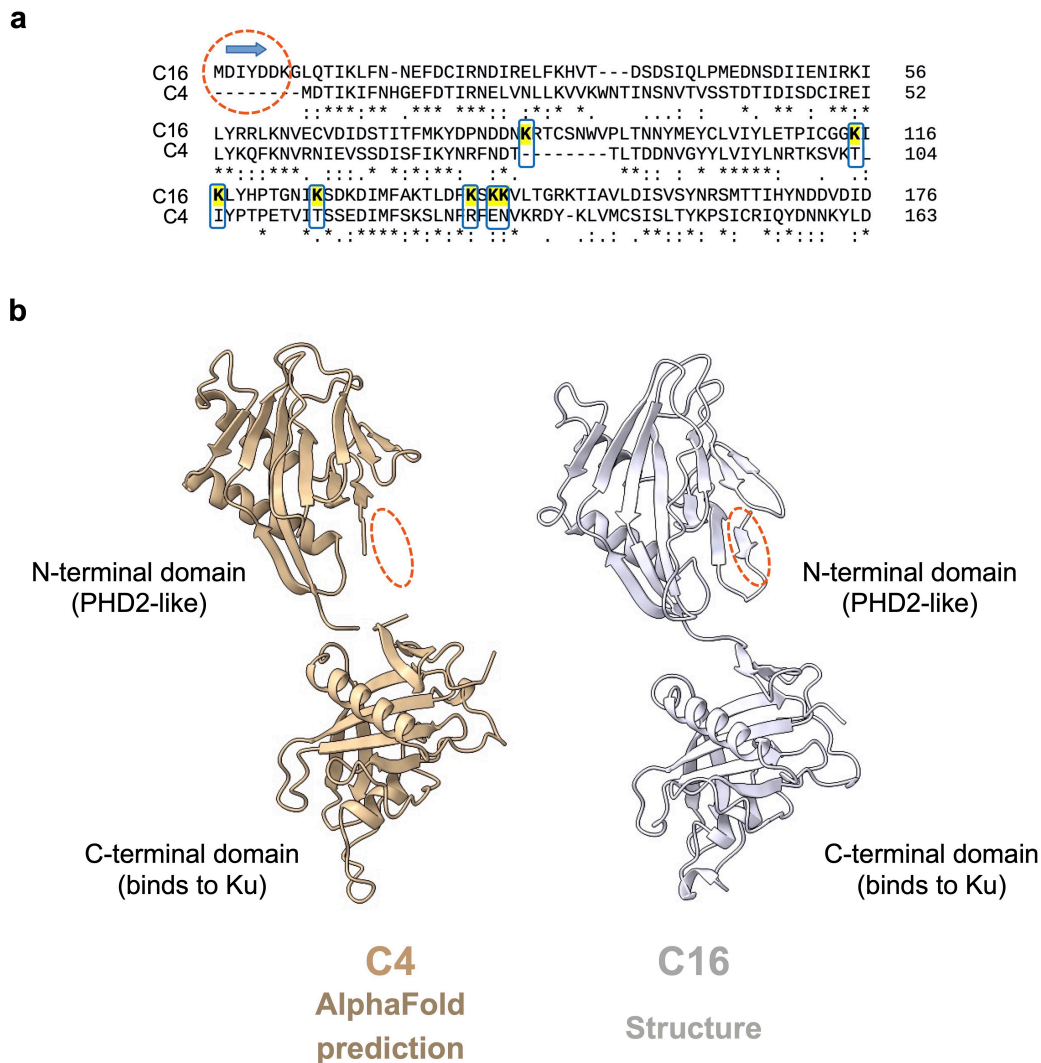
**Supplementary Fig. 7. Structure and heterogeneity in the Ku ring where C16 binds**

**a** The interaction of C16-C with Ku results in some deformation of the bridge forming the DNA binding ring, as revealed by 3D variability analysis performed in cryoSPARC<sup>3</sup> (Supplementary Movies 1, 2 and 3). This figure highlights the conformational changes in the Ku bridge by showing the first and last frame of one of these movies. For the Ku bridge we draw a line along

the density for the bridge to highlight the change in position. Density for C16-N is blurred due to its flexibility and it does not occupy the same position in each of the frames represented.

**b** Two views showing details of the C16 acidic ‘plug’ inserting in the Ku ring. Some residues are indicated.

**c** Particles from the C16 – Ku map were also analysed using focused 3D classification with local masks, not allowing movements during classification, and using a high T value of 200 in Relion 4.0<sup>1</sup>. This analysis revealed different degrees of filling within the ring region in Ku that is at the opposite end to where C16 binds, with a density that exhibited a random orientation. The density was tentatively assigned to the N-terminus of Ku70 because it showed continuity with the rest of the chain in Ku70. This N-terminal region is enriched in negatively charged amino acids which could interact with the positively charged regions at the interior of the DNA binding region in Ku. Figures show several of the maps after 3D classification, with the variable density coloured in yellow.



### Supplementary Fig. 8. Comparison between the structures of C4 and C16

**a** Sequence alignment of the N-terminal region of C16 and C4. A  $\beta$ -strand present in C16 and absent in C4 is indicated with a blue arrow and within a red dashed circle. Several lysine residues at the N-terminus of C16 form the basic cleft (see panel Supplementary Fig. 5) but are absent in C4 and they are highlighted in bold and yellow within a blue square. Alignment was performed using Clustal Omega (1.2.4)(<https://www.ebi.ac.uk/Tools/msa/clustalo/>)<sup>4</sup>.

**b** Comparison between the structure of C16 and the predicted structure of C4 using AlphaFold<sup>5</sup>. An N-terminal  $\beta$ -strand is present in C16 but absent in C4 and its presence and absent is highlighted with a dashed circle.

a

C4_AHKMETA	1	-----MDTIKIFNHGEFVVRNELANLLKSVKWNINSNITVSSDITDISGCISEILYQPKDVENIEVRS	67
C4_CAMELPOX	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C4_COTIA	1	-----MDSVKIFNRSEKFDIIRNELRNISKVIKWNINSNVTSSDITDISRCISEVLYNQPKVNRNIEVSN	67
C4_COWPOX	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C4_MONKEYPOX	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C4_ORTHOPOX	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C4_RACOONPOX	1	-----MGTVKIFNRGEFDDIRNDLVLDLKFVWNTINSNISISSTETIDISQISEILYKQPKVNRNIEVSS	67
C4_SKUNKPOX	1	-----MGNVKIFNRGEFDDIRSDLLDLKLFVWNTINSNISISSTETIDVSRISSEILYKQPKVNRNIEVSS	67
C4_TATERAPOX	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C4_VARIOLA	1	-----MDTIKIFNHGEFDTIRNKLVLNLLKVKVWNTINSNVTSSDITDISNCRSEILYKQPKVNRNIEVSS	67
C4_VOLEPOX	1	-----MGTVKIFNRGEFDDIRNDLVLDLKFVWNTINSNISISSTETVDAISISEILYKQPKVNRNIEVSS	67
C4_VACCINIA	1	-----MDTIKIFNHGEFDTIRNELVNLKVKVWNTINSNVTSSDITDISDCIREILYKQPKVNRNIEVSS	67
C16_VACCINIA	1	MDIYDDKGLQTIKLFNN-EFDCIRNDIRELFKHVTD--SDS-IQLPMDNSDIENIRKILYRRLKNVCEVDID	71
C16_AHKMETA	1	MDIYDDSGIQTIKLFNN-DFDCVRNIRELRFKHVTD--STSIIQHPIDDDSDIENIRKILYRRLKNVCEVDID	72
C16_COWPOX	1	MDIYDDKGLQTIKLFNN-EFDCIRNDIRELRFKHVTD--SDS-IQLPMDKSDIENIRKILYRRLKNVCEVDID	71
C16_ECTROMELIA	1	MDIYDDKGLQTIKLFNN-EFDCIRNDIRELRFKHVTD--FDS-IQLPMDNSDIENIRKILYRRLKNVCEVDID	71
C16_ORTHOPOX	1	MDIYDDKGLQTIKLFNN-EFDCIRNDIRELRFKHVTD--SDS-IQLPMDNSDIENIRKILYRRLKNVCEVDID	71
C16_VARIOLA	1	MDIYDDKGLQTIKLFNN-EFDCIRNDIRELRFKHVTD--SDS-IQLPMDNSDIENIRKILYRRLKNVCEVDID	71
C4_AHKMETA	68	DISFKYNYRN-----CTLTRDDEAYYLIRLYQSKSVKTIIFPTPETVSSDDIMFSSKLNFRFENV	131
C4_CAMELPOX	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_COTIA	68	SISFKYNSKYNL--NEYIPDRRIINNDLEYCLVCLDNTSVKTKIYPSYETTITSSDDIMFSSKLNFRFENV	139
C4_COWPOX	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_MONKEYPOX	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_ORTHOPOX	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_RACOONPOX	68	GITFKYNYRFDN--N--YELNLSNDNMEYVLCIDK-STIKTICTPSTDTVSSDDIMFSSKLNFRFENV	135
C4_SKUNKPOX	68	GITFKYNYRFDN--D--YEVNLSNDNMEYVLCIDKSTVTKIICPTSETTISSEDDIMFSSKLNFRFENV	136
C4_TATERAPOX	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_VARIOLA	68	NISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C4_VOLEPOX	68	GITFKYNYRFDN--D--YELNLSNDNMEYVLCIDKSTVTKIICPTSETTISSEDDIMFSSKLNFRFENV	136
C4_VACCINIA	68	DISFKYNYRN-----DTLTDDNVGYLVIYLNRTKSVKTLIYPTPETVSSDDIMFSSKLNFRFENV	132
C16_VACCINIA	72	TITFMKYDPNDN--KRTCSNVVPLTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	144
C16_AHKMETA	73	TITFMKYDPNDN--KRTCSNVVPLTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	147
C16_COWPOX	72	TITFMKYDPNDN--KRTCSNVVPLTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	144
C16_ECTROMELIA	72	TITFMKYDPNDN--KRTWSNAPLNTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	144
C16_ORTHOPOX	72	TITFMKYDPNDN--KRTWSNAPLNTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	144
C16_VARIOLA	72	TITFMKYDPNDN--KRTCSNVVPLTNNMEYCLVIYLETPICGGKIKLYHPTGNIKSDKIMFAKTLDFKSKKV	144
C4_AHKMETA	132	KRDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	204
C4_CAMELPOX	133	KRDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_COTIA	140	NEGYK-VVVCVSLTYKFSICRIQYDGRKVDIGDQKGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	213
C4_COWPOX	133	KRDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_MONKEYPOX	133	KREYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_ORTHOPOX	133	KREYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_RACOONPOX	136	KRGYK-IVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	210
C4_SKUNKPOX	137	KCGYK-IVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	208
C4_TATERAPOX	133	KLDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_VARIOLA	133	KRDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C4_VOLEPOX	137	KRGYK-IVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	210
C4_VACCINIA	133	KRDYK-LVVCVSLTYKFSICRIQY-NDKCIDISTROEGNLLCYCVITMDPHLLIDLETICVLVDRSGKCLLVNE	206
C16_VACCINIA	145	LTGRKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	219
C16_AHKMETA	148	ITGHKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	222
C16_COWPOX	145	LSGRKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	219
C16_ECTROMELIA	145	LTGRKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	219
C16_ORTHOPOX	145	LTGCKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	219
C16_VARIOLA	145	LTGRKTI AVLDISVSYNRSMTTHYNDVDVIDIHTDKNGKELCYCYITIDHLYLVDVETIGVIVNRSKGLLVNN	218

▲▲▲▲▲  
DPHLLID  
ED Y V

C16-ACIDIC PLUG

C4_AHKMETA	205	FY--TRFRKNIH YDSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	277
C4_CAMELPOX	207	FY--TRFRKNIH YNSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_COTIA	214	FY--IRFKNSIYNTFDLDCMDYIEFIEPESEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMTV	286
C4_COWPOX	207	YN--TRFRKNIH YDSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_MONKEYPOX	207	FY--TRFRKNIH YNSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_ORTHOPOX	207	CY--TRFRKNIH YNSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_RACOONPOX	209	FY--YRFRKNIH YDSFADLCMDHIFELSDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	280
C4_SKUNKPOX	211	FY--YRFRKNIH YDSFADLCMDHIFELSDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	283
C4_TATERAPOX	207	FY--TRFRKNIH YNSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_VARIOLA	207	FY--TRFRKNIH YDSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C4_VOLEPOX	211	FY--YRFRKNIH YDSFADLCMDHIFELSDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	283
C4_VACCINIA	207	FY--TRFRKNIH YNSFADLCMDHIFELPDTEELFTL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	279
C16_VACCINIA	220	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	294
C16_AHKMETA	223	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	295
C16_COWPOX	220	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	294
C16_ECTROMELIA	220	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	294
C16_ORTHOPOX	220	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	294
C16_VARIOLA	219	HLGIGIVKDKRI SDSFGVCMDFIEFSEARELFSL	RNDGGRDIAWDDDNKLESGNNTWI PKTDDDEYKFLSKLMI	293

▲ ▲ ▲ ▲ ▲  
Y I D F D CMDHIFELPD ELFTL D  
L N T DFN LSM E

C4_AHKMETA	278	A-KFRDTKFDYVVLVGDTPCIVFTFKVTKYVNLNLIH-	313
C4_CAMELPOX	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	315
C4_COTIA	287	TKYKNGKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH--	322
C4_COWPOX	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	316
C4_MONKEYPOX	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	316
C4_ORTHOPOX	279	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	315
C4_RACOONPOX	281	K--SKDHKFEYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	315
C4_SKUNKPOX	284	K--SKDNKFEYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	318
C4_TATERAPOX	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	316
C4_VARIOLA	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	316
C4_VOLEPOX	284	K--SKDNKFEYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	318
C4_VACCINIA	280	A-KFNNTKFDYVYMLVGDTPCIVFTFKVTKYVNLNLIH-	316
C16_VACCINIA	295	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	331
C16_AHKMETA	296	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	332
C16_COWPOX	295	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	331
C16_ECTROMELIA	295	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	331
C16_ORTHOPOX	295	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	331
C16_VARIOLA	294	AKSQSDTVFDYVYMLVGDTEPPVVFVFKVTRFYFNMLPK-	330

M-COFFEE    BAD AVG GOOD  
A.A. involved in interaction with Ku    ▲  
Other A.A. conserved in acidic plug    ▲  
NEGATIVE CHARGE conserved in C4 and C16    ■

▲▲▲▲▲  
DTDPC  
E P

**b**

	NCBI REF. SEQ	UNIPROT ACCESSION	UNIPROT CODE	GENE NAME	VIRUS
<i>C4_Akhmeta</i>	YP_010085477.1	-	-	IL-1 receptor antagonist	Akhmeta
<i>C4_Camelpox</i>	NP_570412.1	Q8V2Z5	Q8V2Z5_CAMPM	CamMLVgp022	Camelpox
<i>C4_Cotia</i>	YP_005296213.1	H6TAH8	H6TAH8_9POXV	IL-1 receptor antagonist	Cotia (SPAn232)
<i>C4_Cowpox</i>	NP_619822.1	Q8QN29	Q8QN29_CWPXB	CPXV033	Cowpox
<i>C4_Monkeypox</i>	YP_010377015.1	-	-	C4L/C10L-like family protein	Monkeypox
<i>C4_Orthopox</i>	YP_010085694.1	-	-	Putative C4L protein	Orthopox
<i>C4_Raccoonpox</i>	YP_009143333.1	A0A0G3FZW3	A0A0G3FZW3_RACVI	IL-1 receptor antagonist	Raccoonpox
<i>C4_Skunkpox</i>	YP_009282717.1	A0A1C9KBH6	A0A1C9KBH6_9POXV	IL-1 receptor antagonist	Skunkpox
<i>C4_Taterapox</i>	YP_717332.1	Q0NPI6	Q0NPI6_9POXV	TATV_DAH68_025	Taterapox
<i>C4_Variola</i>	NP_042055.1	P34012	C4_VAR67	VARVgp011	Variola (Smallpox)
<i>C4_Volepox</i>	YP_009281771.1	A0A1C9KC28	A0A1C9KC28_9POXV	IL-1 receptor antagonist	Volepox
<i>C4_Vaccinia</i>	YP_232906.1	P17370	C4_VACCW	VACWR024	Vaccinia
<i>C16_Vaccinia</i>	YP_232892.1	H2DUG7	H2DUG7_9POXV	VACWR010	Vaccinia
<i>C16_Akhmeta</i>	YP_010085466.1	-	-	IL-1 receptor antagonist	Akhmeta
<i>C16_Cowpox</i>	NP_619811.1	Q8QN39	Q8QN39_CWPXB	CPXV022	Cowpox
<i>C16_Ectromelia</i>	NP_671529.1	Q8JLI8	Q8JLI8_9POXV	ECTVgp011	Ectromelia
<i>C16_Orthopox</i>	YP_010085683.1	-	-	C10L protein	Orthopox
<i>C16_Variola</i>	NP_042047.1	P33861	C10_VAR67	VARVgp003	Variola (Smallpox)

**Supplementary Fig. 9. Multiple amino acid sequence alignment of C4 and C16 proteins from the poxviridae family**

a Output from the M-COFFEE server<sup>7</sup>, hand-annotated to incorporate information about C16 dimer - Ku70/80 interactions obtained using the online tool PDBsum (<http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html>). The alignment is coloured according to the internal algorithm of M-COFFEE, where regions coloured yellow/orange/red indicate that the residues are likely to be correctly aligned (see associated key for additional details). Amino acids involved in interactions with Ku are marked with filled triangles and those conserved in the acidic plug but not directly involved in the interaction with Ku are marked with a pink triangle. The predominant amino acid found at each site of interaction is also shown, with those shown in bold text indicating absolute/strong conservation of identity across both C4 and C16 proteins. The residue range corresponding to the described ‘acidic plug’ of C16 is enclosed by a black rectangle, with regions of conserved negative charge indicated by blocks coloured black (strong conservation). Residues from the acid plug that are directly inside de Ku ring or in contact with Ku80 are highlighted within a green box. Other regions of interaction with Ku outside the acidic plug are indicated with green and red boxes. Black boxes serve to highlight the amino acid sequences of the C4 and C16 proteins from Vaccinia described in this study.

**b** Table providing accession codes and additional information about each amino acid sequence used to generate the multiple sequence alignment.

### Supplementary Tables

**Supplementary Table 1. CryoEM data collection parameters**

CryoEM data collection parameters		
Data set name	CNIO, Spain	Leicester, UK
Microscope	JEM-2200FS (JEOL)	Titan Krios G3 (Thermo Fisher Scientific)
Camera	K3 in counting mode (Gatan)	K3 in counting mode (Gatan)
Magnification	40K	105K
Stage Tilt	0°	0°
Voltage (kV)	200	300
Number of Movies	827	13,216
Exposure time (s)	4	2
Number of frames	40	50
Dose per frame (e <sup>-</sup> /Å <sup>2</sup> /frame)	1	1
Total Exposure (e <sup>-</sup> /Å <sup>2</sup> )	40	50
Defocus range (µm)	-1.3 to -2.3	-0.6 to -2.4
Original pixel size (Å/pixel)	0.9773	0.835

**Supplementary Table 2. CryoEM image processing parameters**

CryoEM image processing parameters				
Data set name	CNIO, Spain	Leicester, UK		
Macromolecular Complex	C16 – Ku	C16 – Ku	C16-C – Ku	C16-N
Processing Pixel Size (Å/pixel)	0.9773	0.835	0.835	0.835
Box Size (pixels)	240 x 240	304 x 304	304 x 304	304 x 304
Imposed Symmetry	C1	C1	C1	C1
Number of Particles	242,817	579,038	170,770	81,353
FSC Threshold	0.143	0.143	0.143	0.143
Resolution (Å)	4.04	2.46	2.72	3.47
Resolution Range (Å)	N/A	2.5-4	2.5-4	3.25-5
Map Sharpening B-factor (Å <sup>2</sup> )	N/A	-65	-65	-88

**Supplementary Table 3. Model building and refinement parameters**

<b>Model building and refinement</b>			
<b>Deposited CryoEM model</b> EMDB code	C16-C – Ku EMD-15415	C16-N EMD-15414	Composite map (C16 – Ku) EMD-15416
<b>Initial model</b>	AlphaFold2 for C16 PDB 1JEY for Ku	AlphaFold2	AlphaFold2 for C16 PDB 1JEY for Ku
<b>Model composition</b> Chains Non-hydrogen atoms Protein residues Ligands	4 11,157 1,384 0	2 2,632 322 0	4 13,762 1,703 0
<b>R.m.s. deviations</b> Bond lengths (Å) Bond angles (°)	0.003 0.460	0.002 0.600	0.003 0.486
<b>Validation</b> MolProbity score Clash score Poor rotamers (%)	1.48 4.35 0.00	1.91 8.85 0.00	1.62 6.46 0.26
<b>Ramachandran plot</b> Favored (%) Allowed (%) Disallowed (%)	96.14 3.79 0.07	93.4 6.60 0.00	96.10 3.90 0.00
<b>Deposited PDB code</b>	8AG4	8AG3	8AG5

**Supplementary Table 4. Sequence of DNA fragments used in this work**

<b>Fragment (5'→3')</b>	<b>Sequence</b>
Twin-StrepTag gBlocks <sup>TM</sup> gene fragment (177 bp)	<u>GTTTAACTTTAAGAAGGAGATATAACCATGGGCGGAGAGAA</u> <u>TCTTTATTTTCAGGGC</u> tggagccaccgcagttcgaaaaggaggaggatctgg aggaggatcaggtggttcaagcgcgtggagccaccgcagttcgaaaag <b>TGATAATT</b> <u>AACCTAGGCTGCTGCCACCGCTGAG</u>
C16_FL_RF_Fwd (53 nt)	<u>TAACCTTTAAGAAGGAGATATAACCATGGATATTTACGACGAT</u> <u>AAAGGTCTACAG</u>
C16_FL_RF_Rev (56 nt)	cca <b>GCCCTGAAAATAAAGATTCTCTCCGCCTTTCGGCATATT</b> <u>AAAGTAAAATCTAG</u>

Note: Underlined nucleotides correspond to sequences present in the original pET-52b(+) plasmid. In bold the ATG start codon. In orange sequence of the TEV cleavage site. In lower-case sequence of the Twin-Strep tag. In bold red stop codons. In blue sequences of the C16 protein.

### Supplementary references

1. Kimanius, D., Dong, L., Sharov, G., Nakane, T. & Scheres, S.H.W. New tools for automated cryo-EM single-particle analysis in RELION-4.0. *Biochem J* **478**, 4169-4185 (2021).
2. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**(2018).
3. Punjani, A., Rubinstein, J.L., Fleet, D.J. & Brubaker, M.A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* **14**, 290-296 (2017).
4. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

5. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
6. Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.10.04.463034 (2022).
7. Notredame, C., Higgins, D.G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-17 (2000).