Supplementary Materials for

Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization

Liman Zhang, Shuobing Chen, Jianbin Ruan, Jiayi Wu, Alexander B. Tong, Qian Yin, Yang Li, Liron David, Alvin Lu, Wei Li Wang, Carolyn Marks, Qi Ouyang, Xinzheng Zhang, Youdong Mao,* Hao Wu*

*Corresponding author. E-mail: wu@crystal.harvard.edu (H.W.); youdong_mao@dfci.harvard.edu (Y.M.)

This PDF file includes:

Materials and Methods

Figs. S1 to S8

Table S1

Materials and Methods

Plasmid construction

The genes of full-length mouse NAIP2 and NAIP5 were purchased from Open Biosystems. Mouse NLRC4 gene was reverse transcribed from mouse RAW macrophage cells. The genes of FliC and PrgJ were PCR-amplified from the *Salmonella typhimurium* bacteria. For protein expression by baculoviruses, all constructs were cloned into pFastBac1 plasmids (Invitrogen) using the transfer-PCR method (*31*). The purification tags used include N-terminal His-MBP (HM) or His-Sumo (HS), both with a PreScission protease cleavage site, and C-terminal His (His), generating pFastBac-HM-NAIP2, pFastBac-NAIP2ΔBIR-His, pFastBac-HM-NAIP5, pFastBac-HS-NLRC4, pFastBac-HS-NLRC4Δ, pFastBac-HS-PrgJ and pFastBac-FliC-His. For expression and pull-down experiments in HEK 293T cells, the constructs were cloned into pCMV vectors with appropriate tags, including Flag-NAIP5, His-MBP-NLRC4Δ, and His-FliC. All mutagenesis experiments were performed with the QuikChange method.

Protein expression and purification

The baculoviruses were prepared using the Bac-to-Bac® system (Invitrogen), and the proteins were expressed in sf9 cells following the manufacturer's instructions. To express the PrgJ/NAIP2/NLRC4Δ complex, 10 ml of HS-PrgJ, HM-NAIP2 and HS-NLRC4Δ baculoviruses were co-infected into 1 liter sf9 cells and harvested 48 hrs post-infection. Cells were lysed by sonication in buffer A (50 mM Tris at pH 8.0, 300 mM NaCl, 20 mM imidazole, 10 μg/μl DNase I and 1 mM βME) with freshly added protease inhibitor cocktail (sigma). After centrifugation, the supernatant was incubated with 3 ml Ni-NTA agarose beads at 4°C for 1 hr. The beads were washed with buffer A, and the bound protein was eluted with 500 mM imidazole in buffer A in 3 ml fractions. The fractions were pooled, placed on a 20 % sucrose cushion, and centrifuged at 40,000 rpm for 12 hrs. The resulting pellet was re-suspended with buffer A and placed on a step-gradient of sucrose with 30 %, 35 %, 40 % and 45 % sucrose in 10 mM HEPES at pH 7.5 and 300 mM NaCl. The sucrose gradient was centrifuged for 14 hrs at 40,000 rpm and fractions were collected manually at 400 μl/tube. The fractions were buffer-exchanged using Zeba^{TM} spin desalting columns (Thermo, #89882) and examined under negative staining EM (next section). Those fractions with complete inflammasome particles were selected for cryo-EM studies. Expression and purification of the PrgJ/NAIP2/NLRC4 and FliC/NAIP5/NLRC4Δ complexes were performed similarly using co-infection using the appropriate baculoviruses The PrgJ/NAIP2/NLRC4 sample used in the caspase-1 polymerization assay was concentrated and exchanged to Assay Buffer (25 mM Tris at pH 8.0, 150 mM NaCl and 2 mM DTT) using Zeba^{TM} spin desalting column. HS-NLRC4 was expressed using baculoviruses and purified sequentially by Ni-NTA agarose, HiTrap Q and gel filtration chromatography with separated monomers and aggregates. His-GFP-NLRC4^{CARD} was expressed in *E. coli* and purified by Ni-NTA agarose and gel filtration chromatography to isolate the monomers.

For expression of the FliC/NAIP5/NLRC4Δ complex in HEK 293T cells, His-FliC, MBP-NLRC4Δ, Flag-NAIP5 constructs were co-transfected, harvested at 48 hrs with a rubber scraper, and sedimented at ~2,600 x g at 4 °C for 20 min. Cells were washed once with 20 ml of ice-cold PBS to remove serum proteins. The cell pellet was resuspended in 15 ml of PD buffer (50 mM Tris, 50 mM NaCl, 1 % Triton 100 and freshly added protease inhibitor cocktail) and incubated at 4 °C for 10 min. After centrifugation at 4,000 rpm at 4 °C for 30 min, the supernatant was incubated with amylose resin (New England Biolabs) for 2 hrs at 4 °C, washed with 10 ml of PD buffer and eluted with 25 mM maltose in PD buffer. The eluate sample was further purified with 100 μl Flag resin, incubated for 2 hrs at 4 °C, washed with 10 ml of PD buffer, and then eluted with 200 μg/ml Flag peptides in PD buffer.

Negative staining EM

For negative staining, 10 μl of an inflammasome complex was placed onto a glow-discharged copper grid, washed twice with H_2O and stained with 2 % uranyl formate for 20 seconds and air-dried. The grids were imaged using a JEOL 1200EX 80 kV Transmission Electron Microscope (TEM) and recorded with an AMT 2k CCD camera (Harvard Medical School Electron Microscopy Facility).

Cryo-EM data collection

A 3 μl drop of the PrgJ/NAIP2/NLRC4Δ inflammasome complex at 3 mg/ml was applied to a glow-discharged C-flat grid (R 1/1, 400 Mesh, Protochips, CA, USA), blotted for 3 seconds, plunged into liquid ethane, and flash frozen using an FEI Vitrobot Mark IV. The blotting time was 1 second for the cryo-grids used for data collection of enriched side views of the inflammasome. The cryo-grids were imaged in an FEI Tecnai Arctica microscope, equipped with an Autoloader, at a nominal magnification of 21,000 and an acceleration voltage of 200 kV. Coma-free alignment was manually conducted prior to data collection. Cryo-EM data were collected semi-automatically using the Leginon version 3.1 software (*32*) on a Gatan K2 Summit direct detector camera (Gatan Inc., CA, USA) in a super-resolution counting mode, with 9.0 seconds of total exposure time and 250 millisecond per frame. This resulted in movies of 36 frames per exposure and an accumulated dose of 47.6 electrons/$Å^2$. The physical and the super-resolution pixel sizes are 1.72 Å and 0.86 Å, respectively. The defocus level in the data collection was set in the range of 1.0-3.0 μm. A total of 12,328 movies were collected, among which 9,113 movies were selected for further data analysis.

Cryo-EM data analysis and structure determination

Raw movie frame data were first corrected for their gain reference and each movie was used to generate a micrograph that was corrected for sample movement and drift. These drift-corrected micrographs were used for the determination of actual defocus of each micrograph with the CTFFind3 program (*33*). A power spectrum from a representative micrograph showed Thon-rings extending to ~3.0 Å resolution (fig. S3A). 180,585

particles were initially selected from the micrographs with combination of manual picking in EMAN2 (*34*) and automated picking in SPIDER (*35*). The automated picking was only used to pick particles for the top views from micrographs at relatively large defocus. All side views and tilted views were picked manually. Manual particle picking was also practiced for top view particles on micrographs at relatively low defocus and with low particle densities. The picked particles were binned by two times and then subjected to reference-free 2D classification in RELION 1.3 (*36*). Bad particles were rejected as a whole class upon inspection of class average quality. Good class averages were selected to perform common-line reconstruction in SPIDER. The resulting initial model was low-pass filtered to 60 Å, and was used for reconstruction of a consensus model in RELION using all selected particles. The consensus model, low-pass filtered to 60 Å, was used as the input reference to conduct unsupervised 3D classifications in RELION without assumption of any symmetry, using initially selected particles (fig. S2, S3B). Junk reconstructions were rejected. Data sets corresponding to good 3D reconstructions were further classified separately within each 3D class by reference-free 2D classification in RELION. The mis-classified particles (about 10 % initially) were re-sorted to the different 3D classes for additional rounds of 3D classification. Such iterative 2D and 3D classifications were performed until no symmetry-mismatched class averages were observed. In later cycles, reference-free 2D classification into 450 classes within each 3D class was performed using a machine learning algorithm generative topographic mapping (*37*) to verify the homogeneity and correct mis-classification. These operations resulted in three final data sets corresponding to 10-fold, 11-fold and 12-fold, with 4,489, 75,114 and 30,430 particles, respectively.

Although each inflammasome disk should contain one PrgJ/NAIP2 complex, there is no observable difference among the individual blades to determine which might represent the PrgJ/NAIP2 complex, probably due to the similar domain organizations of NAIP2 and NLRC4, and the small size of PrgJ (Fig. 1A). Upon imposing the apparent symmetry, the C11, C12 and C10 reconstructions reached to resolutions of 4.7 Å, 7.5 Å and 12.5 Å, respectively, as measured by Fourier shell correlation (FSC) of a gold-standard refinement in which two halves of the dataset were refined separately and combined only when building the final map (fig. S3C-E) (*38*). The final 11-fold reconstruction was refined with the super-resolution pixel size (0.86 Å), whereas the final 10-fold and 12-fold reconstructions were refined with the two-fold-binned pixel size (1.72 Å), which corresponds to the physical pixel size of the detector sensor. For the refined 11-fold reconstruction, the orientation distribution of particles and tilt-pair validation of chirality are shown (fig. S3F-G). To access the local resolution in the refined 11-fold reconstruction, we used three different approaches (fig. S4). First, we used the blocres program in Bsoft package (*39*) to calculate the local resolution by the local FSC = 0.3 cutoff between the two half maps refined separately in the gold-standard procedure in RELION. Second, we used the blocres program in the Bsoft package to calculate the local resolution by the local FSC = 0.3 cutoff between the final 11-fold map and the refined atomic model (see below). Third, we used ResMap (*40*) to assess the local resolution using the two half maps refined separately instead of using the single merged final map. In all three cases, the local resolutions of the inner ring and the outer ring were estimated to be largely in the range of 4.0-6.0 Å and 4.5-7.5 Å, respectively (fig. S4). The final 11-fold map was corrected for its amplitude by a B-factor of -150 Å^2 in the Bfactor

program, and was low-pass filtered at 4.7 Å with a cosine edge of 8-Fourier-pixel width before any further atomic model fitting.

Atomic model fitting and analysis

Because the cryo-EM density should mainly represent NLRC4 instead of NAIP2, we utilized the crystal structure of NLRC4Δ in the inactive conformation (PDB ID: 4KXF) for model building (*16*). From the 8 molecules in 4KXF, a composite structure was generated starting from one molecule and replacing the missing loops with corresponding observed loops in other molecules, leaving only one break in the entire structure. The composite structure was divided into two regions at reside 356. Resides 93-356 correspond to NBD and HD1 domains, and residues 357-909 correspond to WHD, HD2 and LRR domains. Each region was fitted into the cryo-EM maps as a rigid-body in UCSF Chimera (*41*).

For the C11 reconstruction, refinement was first carried out on a single NLRC4 subunit model in real space using Modeller (*20*) and Rosetta (*42*). Refinement was then performed on three consecutive NLRC4 subunits to take into account subunit interactions using the cryo-EM map segmented in Chimera (*41*), which includes only density for such three subunits. The cryo-EM map and the atomic model were placed into a pseudo-unit cell and the refinement was performed in Phenix (*43*) in Fourier space using both amplitudes and phases, with restraints by non-crystallographic symmetry, reference model and secondary structures. The crystal structure of inactive NLRC4Δ monomer was used as the reference model. The final refinement statistics are shown in table S1. The FSC curve between the map and the final refined atomic model gave an estimated resolution of 4.7 Å at FSC = 0.5 (fig. S3C).

To model the NAIP2 structure based on the C11 cryo-EM map, the Swiss-Model Server (http://swissmodel.expasy.org) was first utilized to predict the homology structure of inactive NAIP2ΔBIR using the composite NLRC4Δ crystal structure as the template. Fitting of NAIP2 into the cryo-EM map was performed similarly as for NLRC4Δ and refined in Modeller (*20*), resulting in a pseudo-atomic model of active NAIP2. All figures of the structures were plotted in Chimera (*41*) and Pymol (*44*). Interaction analysis was performed using PISA (*45*).

Pulldown

For mutational analysis, His-FliC, His-MBP-NLRC4Δ (WT and mutants), Flag-NAIP5 (WT and mutants) proteins were co-expressed in HEK 293T cells and pulled down using amylose resin (New England Biolabs) or anti-Flag M2 affinity gel (Sigma). Proteins were resolved on 10 % SDS-PAGE gels and Western blotted with anti-His, anti-MBP, and/or anti-Flag antibodies.

Caspase-1 polymerization assay

Preparation of fluorescently labeled caspase-1^{CARD} and the fluorescence polarization (FP) experiment were performed as previously described (*24*). Briefly, the sortase method (*46*)

was used to label the monomeric "sandwich"-tagged His-MBP-caspase-1[CARD]-Sumo-LPETGG ("LPETGG" is the sortase motif) with the fluorophore TAMRA. A TEV cleavage site is present after the His-MBP tag. Labeling reagents, including an engineered, Ca^{2+}-independent sortase and the peptide-fluorophore conjugate Gly-Gly-Gly-TAMRA (GGG-TAMRA), were kindly provided by Dr. Hidde Ploegh. The polymerization assay was performed in a black round-bottom 384-well Greiner Bio-One plate in 20 μl volumes. For each well, one part of reaction buffer containing 20 mM HEPES at pH 8.0, 150 mM NaCl, 1 mM TCEP and 1 % Triton-X100 was mixed with three parts of the protein solution containing labeled caspase-1[CARD] and one of the following: the purified PrgJ/NAIP2/NLRC4 complex at different concentrations, His-GFP-NLRC4[CARD], NLRC4 monomers, or NLRC4 aggregates from the void fraction. The final concentration of labeled caspase-1[CARD] was 2.5 μM, and those of GFP-NLRC4[CARD], NLRC4 monomers, and NLRC4 aggregates were 2.0 μM. TEV protease at a final concentration of 50 μg/ml was then added immediately before starting the measurements. Data were collected with the SpectraMax M5e Multi-Mode Microplate Reader (Molecular Devices) using excitation at 561 nm and emission at 585 nm with an auto-cutoff filter at 570 nm.
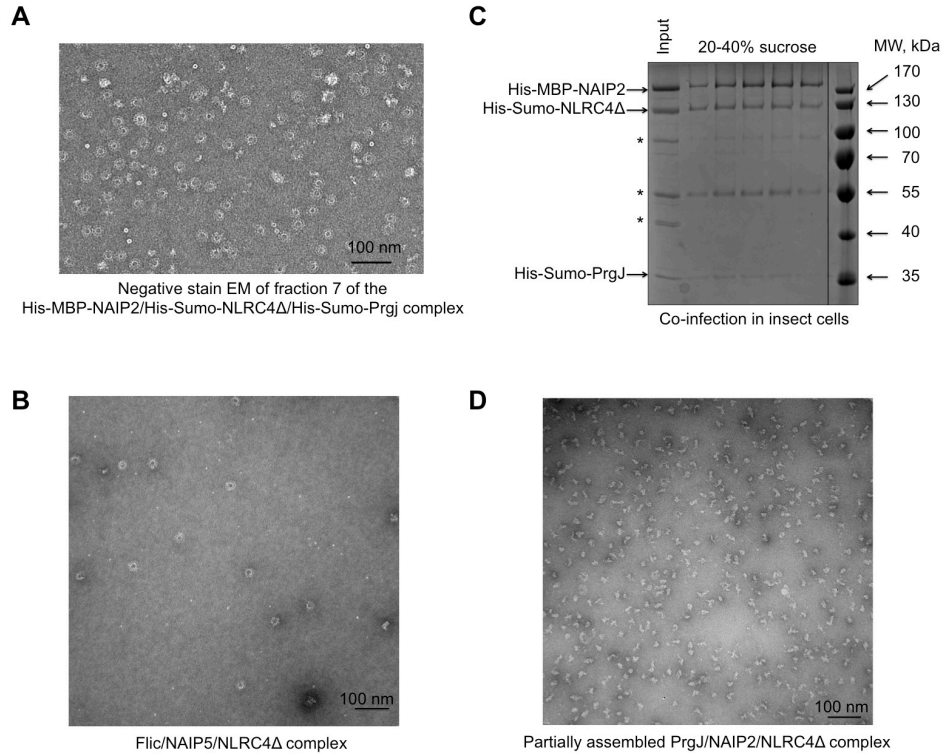
A

Negative stain EM of fraction 7 of the
His-MBP-NAIP2/His-Sumo-NLRC4Δ/His-Sumo-Prgj complex

C

Co-infection in insect cells

B

Flic/NAIP5/NLRC4Δ complex

D

Partially assembled PrgJ/NAIP2/NLRC4Δ complex

Fig. S1. **Examples of purification and negative staining EM images of the FliC/NAIP5/NLRC4Δ complex and the PrgJ/NAIP2/NLRC4Δ complex.** (**A**) A representative negative staining EM image from fraction 7 in Figure 1B of the PrgJ/NAIP2/NLRC4Δ complex. It is from the same EM image as Fig. 1C but contains a larger area. (**B**) A negative staining EM image of the FliC/NAIP5/NLRC4Δ complex, showing similar disk-like structures as those of the PrgJ/NAIP2/NLRC4Δ complex. (**C**) SDS-PAGE gel of fractions from sucrose gradient ultracentrifugation, showing apparent similar molar ratios between NAIP2 and NLRC4Δ. Asterisks indicate contaminating bands. (**D**) A corresponding negative staining EM image showing the partially assembled disk-like particles.
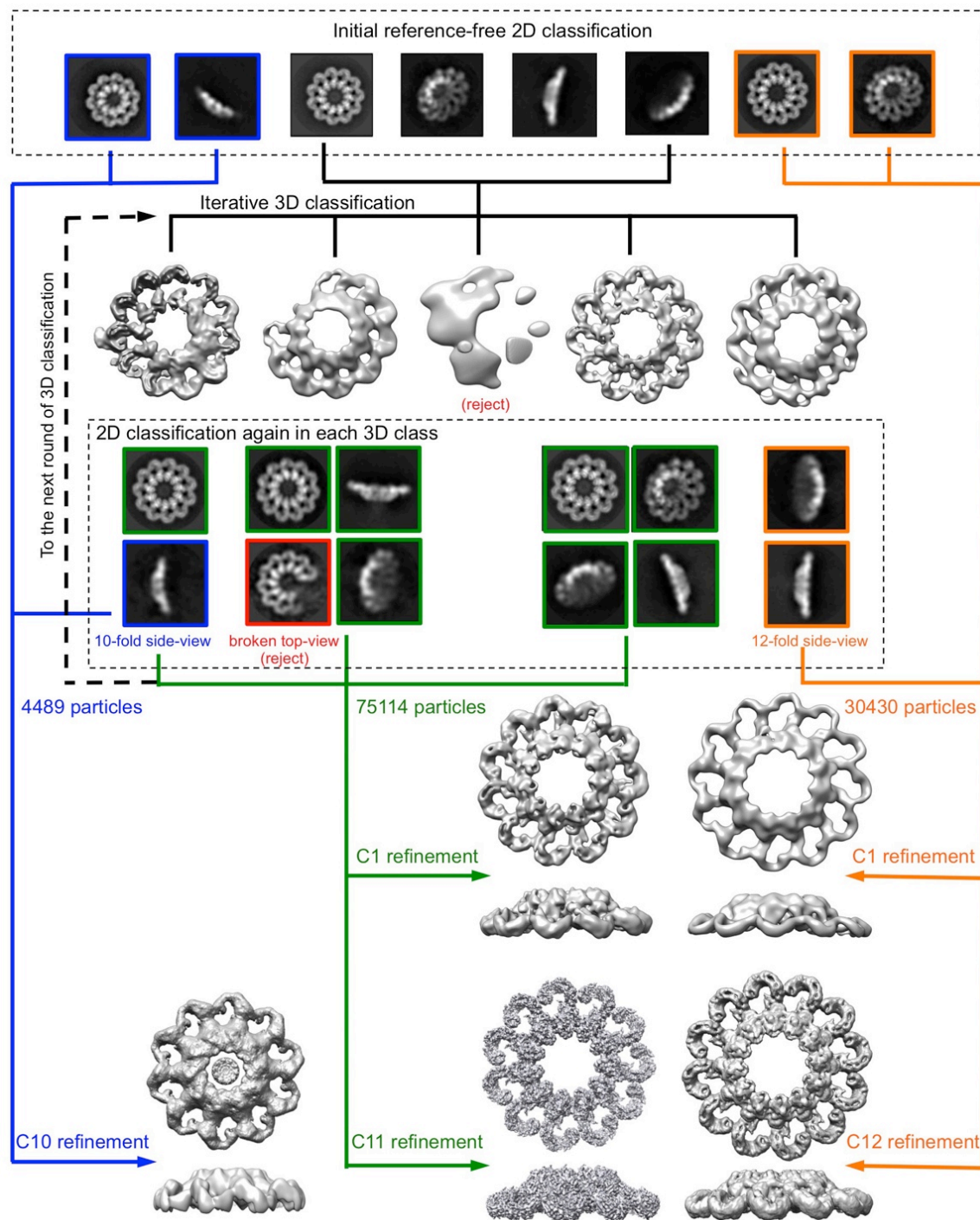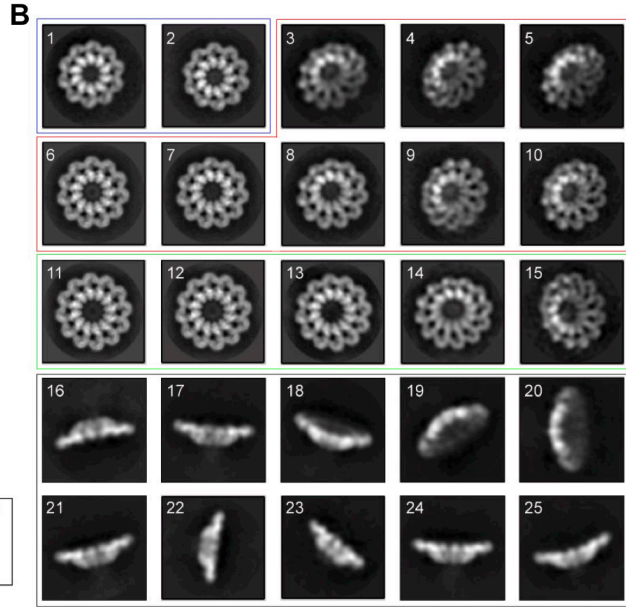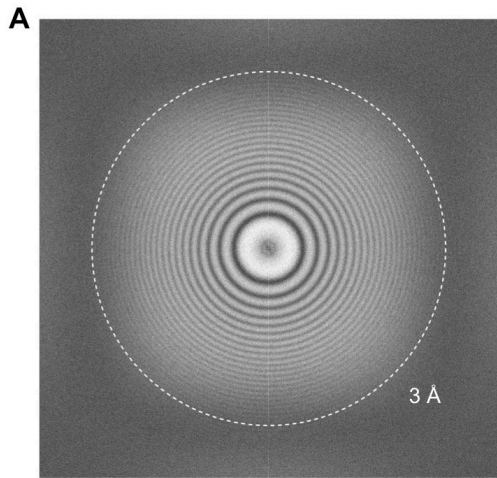
Fig. S2. Schematic workflow of 2D/3D classification to sort out conformational heterogeneity. The workflow diagram illustrates the basic strategy practiced in this study. In total, we completed 8 rounds of 3D classification interleaved with 2D classification in each 3D class. After each round of 3D and 2D classification, bad classes were rejected by visual inspection. The finally classified 11-bladed and 12-bladed data sets were first subjected to 3D refinement with C1 symmetry. The refined maps clearly

indicated the 11-fold and 12-fold symmetry for the two classes. We then applied C11 and C12 symmetry and repeated the refinement using a gold-standard procedure in RELION. For the 10-fold map, the data size was too small for reliable refinement without symmetry assumption. Given that the 2D class averages of 10-bladed disks clearly indicated 10-fold symmetry, we conducted the gold-standard refinement on this small dataset with C10 symmetry directly.

A

3 Å

B

Blue box: Top/bottom views of 10-bladed complexes
Red: box: Top/bottom views of 11-bladed complexes
Green box: Top/bottom views of 12-bladed complexes
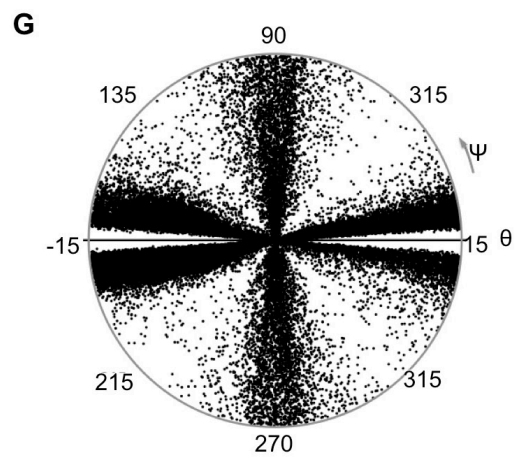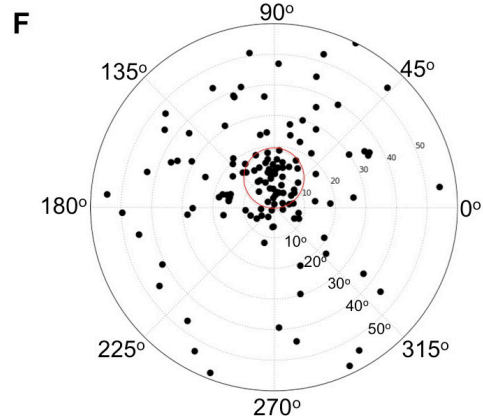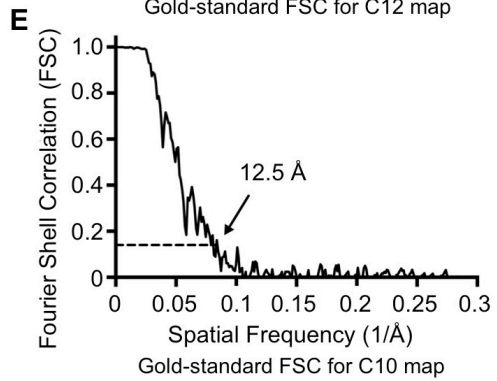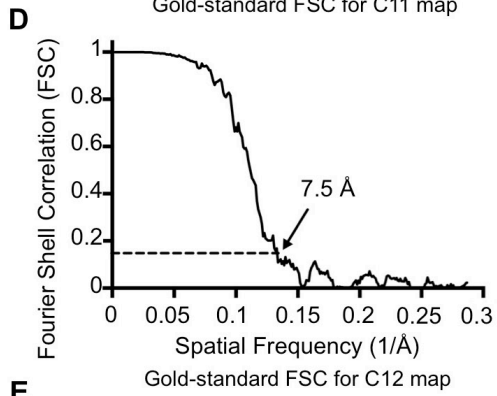Black box: Side views of all complexes

C

Merged vs. atomic model
Half 1 vs. atomic model
Half 2 vs. atomic model
Half 1 vs. half 2

4.7 Å

Fourier Shell Correlation (FSC)
Spatial Frequency (1/Å)
Gold-standard FSC for C11 map

D

7.5 Å

Fourier Shell Correlation (FSC)
Spatial Frequency (1/Å)
Gold-standard FSC for C12 map

E

12.5 Å

Fourier Shell Correlation (FSC)
Spatial Frequency (1/Å)
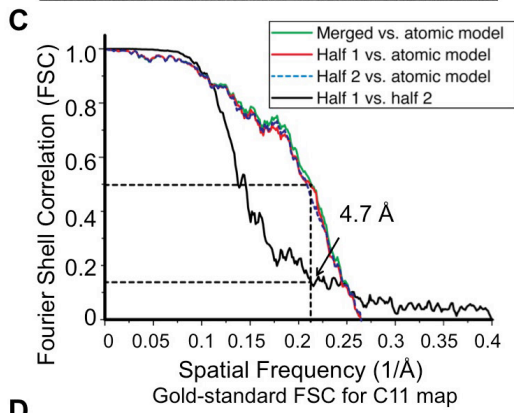Gold-standard FSC for C10 map

F

G

Fig. S3. **Cryo-EM structure determination.** (**A**) The power spectrum of a cryo-EM image, showing that the Thon rings extend beyond 3Å. (**B**) 2-D classification showing the 10-, 11- and 12-bladed inflammasome complexes. The dimension of each image is 43.5 nm. (**C-E**) The gold-standard Fourier Shell Correlation (FSC) plots from two half-reconstructions refined separately in RELION for the C11 (C), C12 (D) and C10 (E) averaged PrgJ/NAIP2/NLRC4Δ complex. For the 11-fold reconstruction, the FSCs of the first half map (red), the second half map (dashed blue) and the merged map (green) versus the fitted atomic model are shown. (**F**) Tilt-pair validation of the cryo-EM map of the 11-bladed PrgJ/NAIP2/NLRC4Δ complex, showing that the data points cluster around the experimental tilt angle (10°) and tilt axis (90°). The average tilt angle and tilt axis calculated by the data points in the red circle is 10.1° ± 3.8° and 92.5° ± 5.2°. There are a total of 126 tilt-pairs used in this validation. (**G**) Angular distribution of particle orientations for reconstruction of the 11-bladed PrgJ/NAIP2/NLRC4Δ complex.
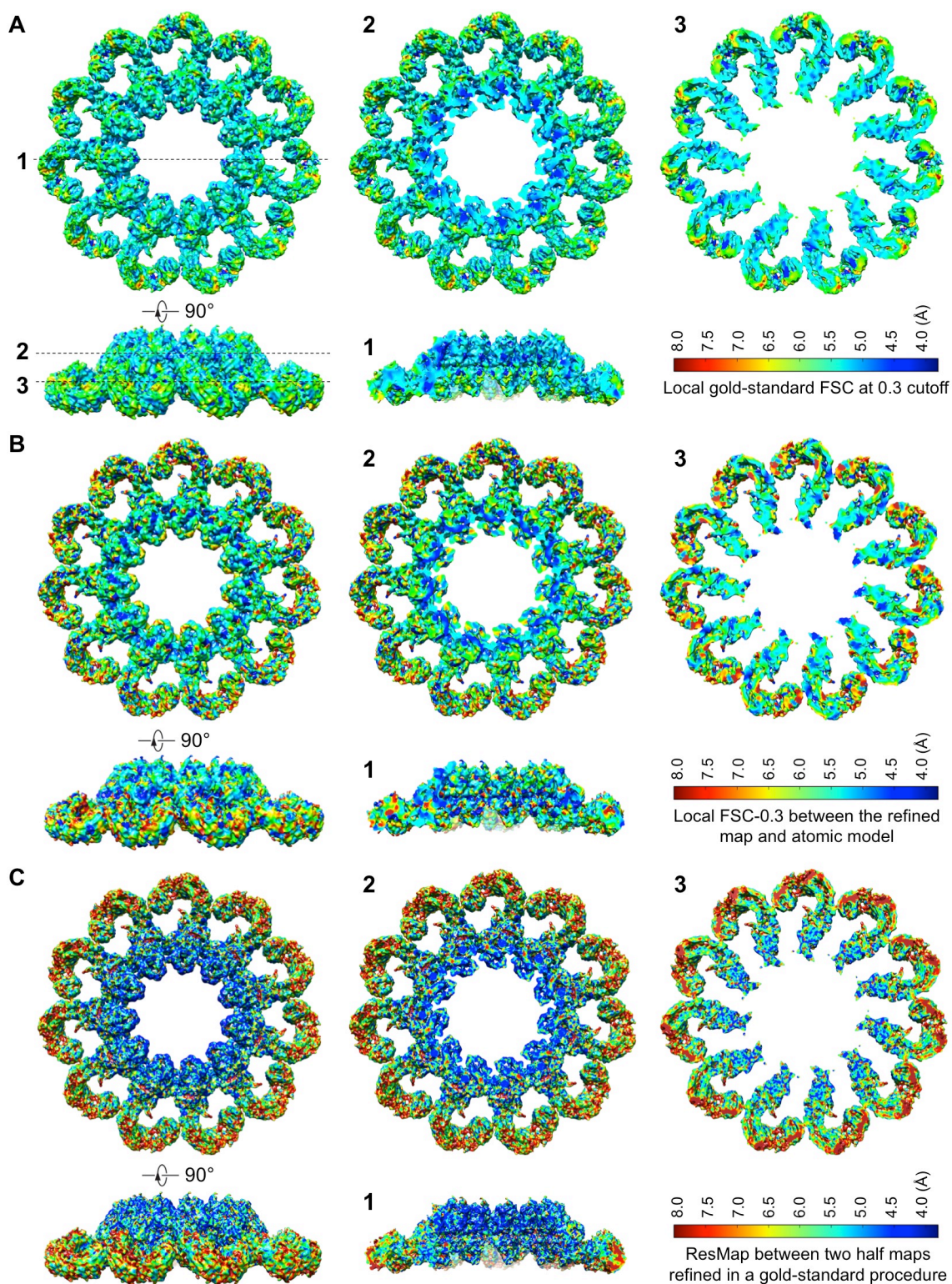
Fig. S4. Local resolution estimation of the 11-fold reconstruction through three different methods. (**A**) The local resolution was estimated by local FSC between two half maps separately refined in a gold-standard procedure in RELION, using the blocres

program in Bsoft. Panels 1-3 show the local resolution at different cross-section positions in the 3D map. The position of the cross section is illustrated in the left panel. (**B**) The local resolution was estimated by local FSC between the final refined map and the atomic model, using the blocres program in Bsoft. (**C**) The local resolution was estimated by ResMap on the two half maps separately refined in a gold-standard procedure in RELION. The left image of this panel is also presented in Fig. 2A. For the convenience of comparison, the color bar for resolution remains at the same range and scale in all three cases. Although the surface resolution appears to vary from case to case, cross-sections 1 and 2 among all three cases show consistently higher resolution (blue) in the inner ring, corresponding to 4-6 Å, than in the outer ring.
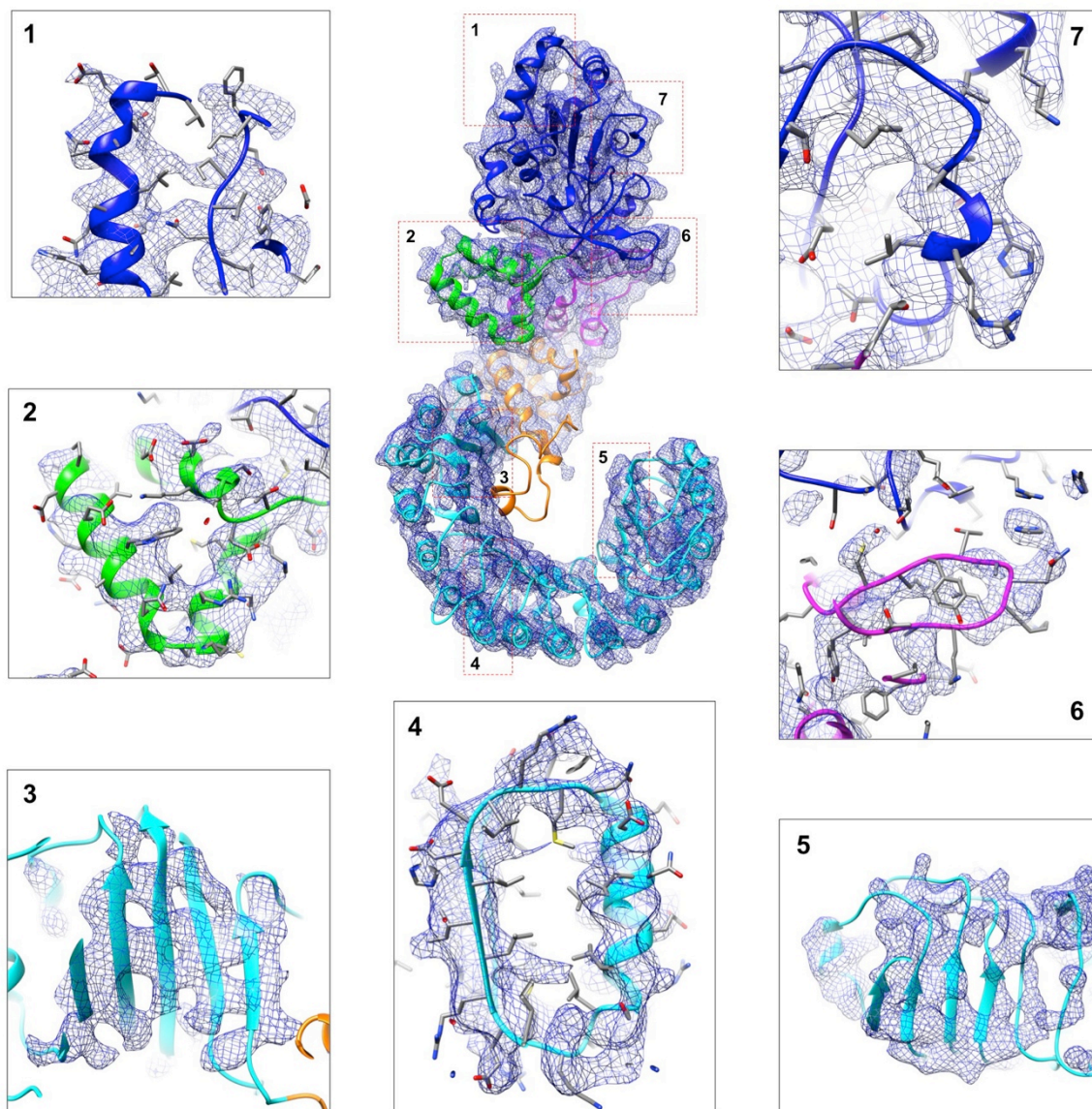
Fig. S5. Close-up evaluation of 11-fold map quality. In the central panel, only one subunit of NLRC4Δ is shown with superposition of the refined atomic model on the cryo-EM density map. Seven local areas, labeled by the numbers, are shown with close-up views. The relative position of each close-up view in the cryo-EM map is indicated using a dashed red box labeled with the identical number. However, each close-up view may assume different orientation for the convenience of visual inspection. These close-up views exhibit the expected secondary features at the estimated resolution, including the bulky side-chain bumps (1, 2, 4, 6 and 7), the partial separation between neighboring β-strands (3, 5 and 6), the helical pitches (1, 2, 4 and 7), and the hole at the tip of a loop (6).
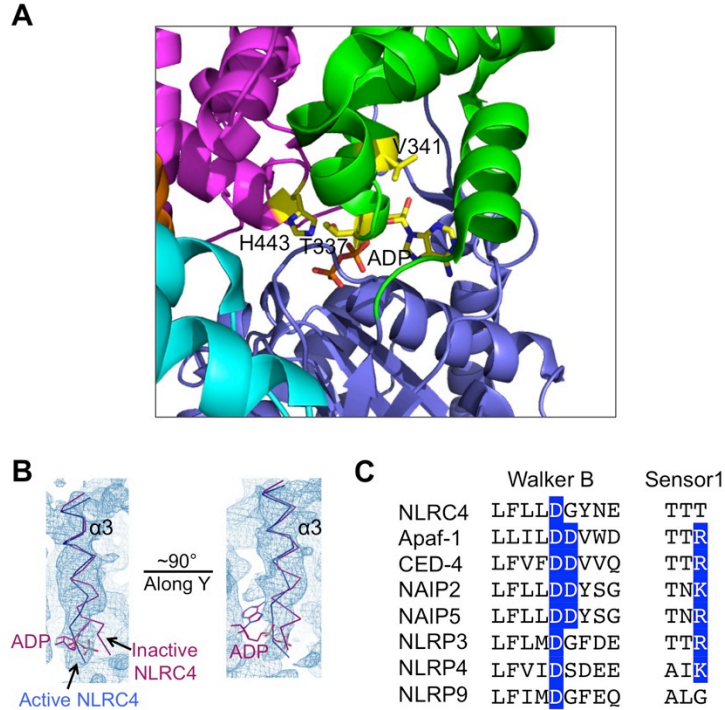
Fig. S6. The nucleotide-binding pocket of NLRC4. (**A**) Residues that are mutated in human auto-inflammatory conditions (*12-14*) map to the interface between WHD and the NBD-HD1 module, or the core of HD1. They are situated around the ADP-binding site in the crystal structure of closed NLRC4Δ and at the region with conformational changes during NLRC4 activation. (**B**) Superimposed active NLRC4 (blue) and inactive NLRC4 (purple) at the ADP-binding site from inactive NLRC4. The cryo-EM map of active NLRC4 is shown. The adenine base and ribose of the bound ADP in inactive NLRC4 are out of density and the di-phosphates of the bound ADP would have been in clash with the slightly moved position of helix α3 in the active NLRC4 structure. (**C**) Sequence alignment among several members of the signal transduction ATPases with numerous domains (STAND) class of P-loop NTPases (*47, 48*), showing the modified Walker B motif and lack of the conserved Arg in NLRC4. NLRC4 and many other NLR members such as NLRP3 have a modified Walker B motif in which the second Asp in the conserved Asp-Asp sequence is a Gly (*19*). NLRC4 also lacks a conserved Arg in the sensor I motif.
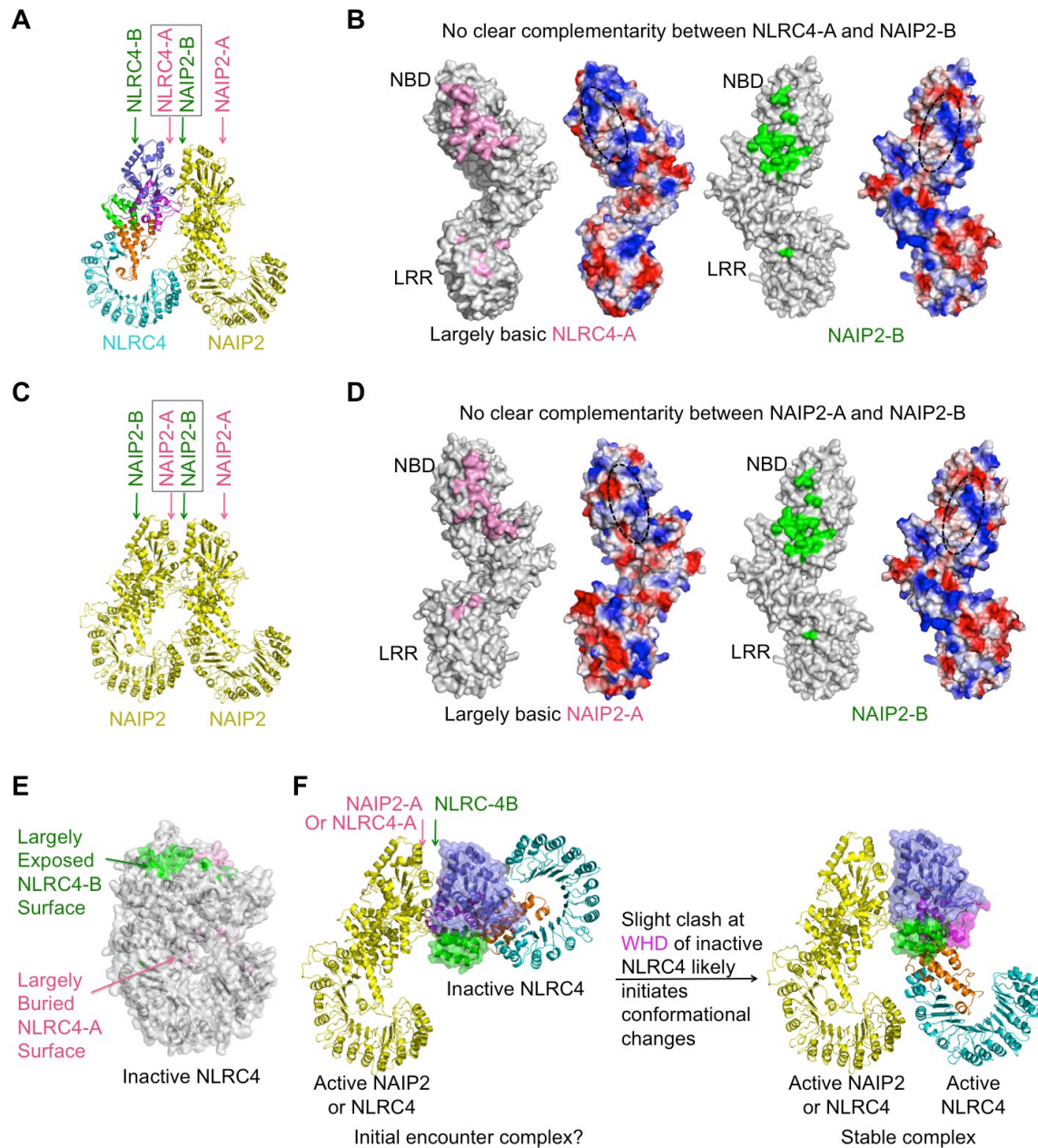
Fig. S7. Analysis of NAIP2 and NLRC4 interactions. (**A**) Locations of NLRC4-A and NAIP2-B surfaces. (**B**) Mapped interactions at the NLRC4-A (pink) and NAIP2-B (green) surfaces and their electrostatic potentials, showing no obvious complementarity. Dotted ovals show the approximate locations of the interface on the electrostatic surfaces. (**C**) Locations of NAIP2-A and NAIP2-B surfaces. (**D**) Mapped interactions at the NAIP2-A (pink) and the NAIP2-B (green) surfaces and their electrostatic potentials, showing no obvious complementarity. Dotted ovals show the approximate locations of the interface on the electrostatic surfaces. (**E**) Crystal structure of inactive NLRC4Δ showing the largely buried A surface (pink) and the largely exposed B surface (green). (**F**) Proposed initial encounter of NAIP2-A (or NLRC4-A) with the B surface of auto-inhibited NLRC4 triggers the conformational change at the WHD.
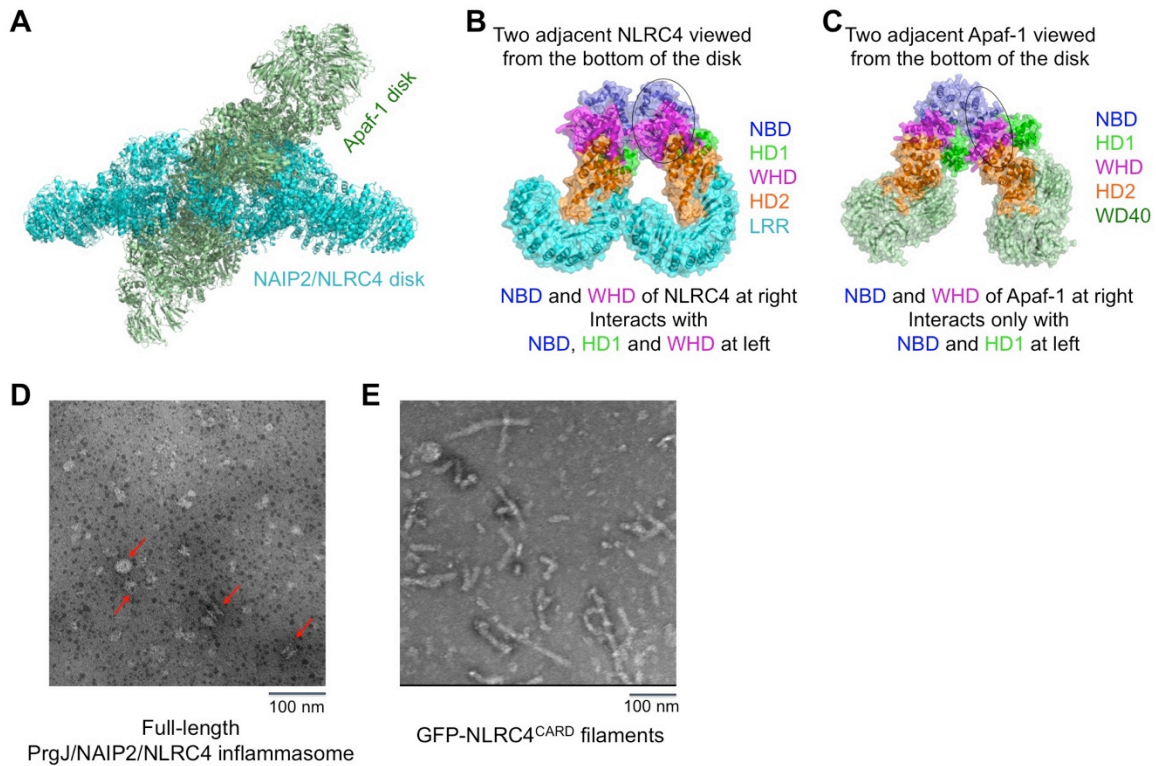
16

Fig. S8. Comparison of NLRC4 and Apaf-1 structures. (**A**) Comparison of the overall structure of the C11 NLRC4 disk (cyan) and the Apaf-1 disk (pale green, 3J2T) by superimposing the NBDs of one NLRC4 subunit and one Apaf-1 subunit. An angle of approximately 50-60° is intercepted by the disk planes, indicating that the NBDs use different surfaces for oligomerization. (**B**) Two adjacent NLRC4 subunits shown in ribbon and transparent surface with indicated domain colors. The subunit at right in approximately the region marked by the oval interacts with the entire NBD-HD1-WHD module of the subunit at left. The subunits are more parallel to each other. (**C**) Two adjacent Apaf-1 subunits shown in ribbon and transparent surface with indicated domain colors. The subunit at right in approximately the region marked by the oval does not interact with the WHD, only the NBD-HD1 module of the subunit at left. The subunits are more oblique to each other. (**D**) Negative staining EM image of the PrgJ/NAIP2/NLRC4[FL] complex. Red arrows denote examples of disk-like and stacked disk-like structures. (**E**) His-GFP-NLRC4[CARD] shows filamentous structures under negative staining EM.

Table S1. Structure determination statistics

Cryo-EM reconstruction and refinement of the C11 particles	
Number of particles in the final reconstruction	75,114
Number of particles in side views	8,304
Pixel size in the super-resolution counting mode of K2 Summit (Å)	0.86
Pixel size corresponding to the physical detector sensor (Å)	1.72
Defocus range (μm)	-1 to -3
Electron dose (electrons / Å^2)	47.6
Number of frames in each movie	36
Electron energy (kV)	200
Pixel dimension of single-particle image	448
Isotropic B-factor for amplitude correction (Å^2)	-150

Pseudo-crystallographic refinement	
Cell dimensions a, b, c (Å)	138.5, 205.2, 102.6
Cell angle α, β, γ (º)	90.0, 90.0, 90.0
Space group	P1
Bragg spacing limits (Å)	199.5 to 4.5
Total observations of reflections	67,409
Number of protein residues	2,727
Number of atoms	21,945
R_{work} (R_{free}) factor (%)	36.4 (38.3)
Geometric parameters (RMSD[†])	
Bond length (Å)	0.0073
Bond angle (º)	2.048
Ramachandran statistics	
Residues in favored regions (%)	90.83
Residues in allowed regions (%)	7.48
Residues in disallowed regions (%)	1.69
Rotamer outliers (%)	3.17

[†]RMSD: root-mean-square deviation