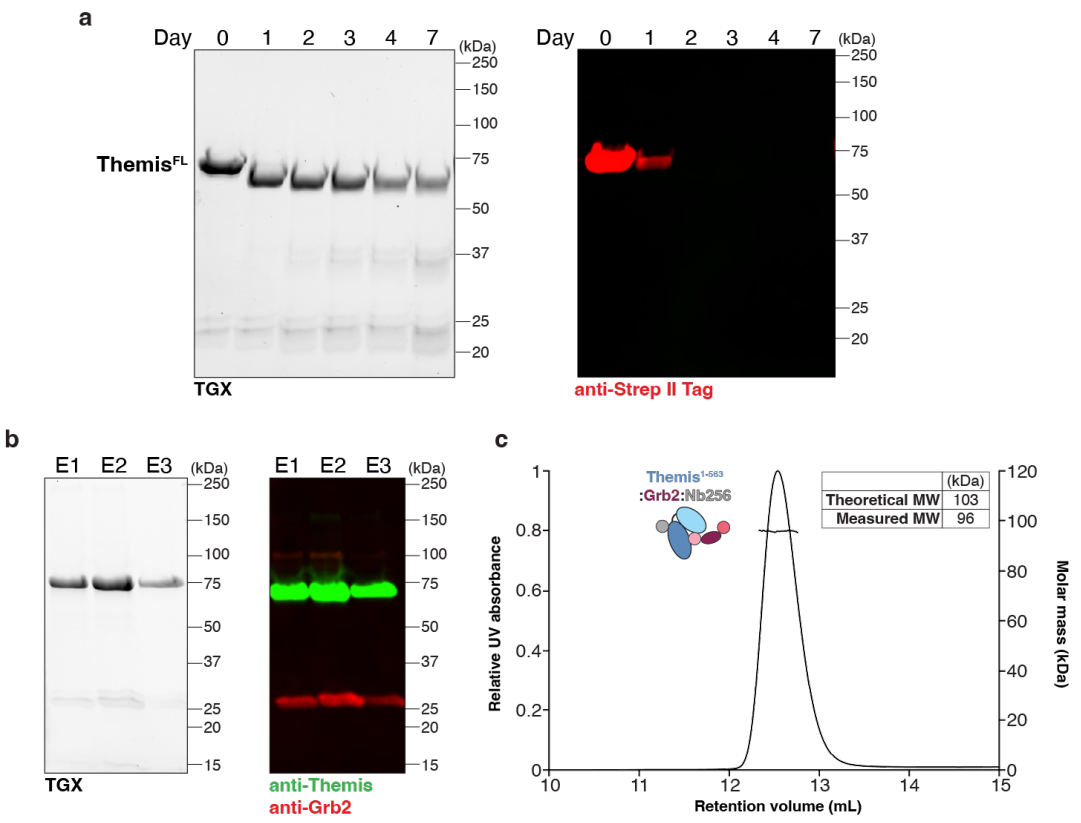
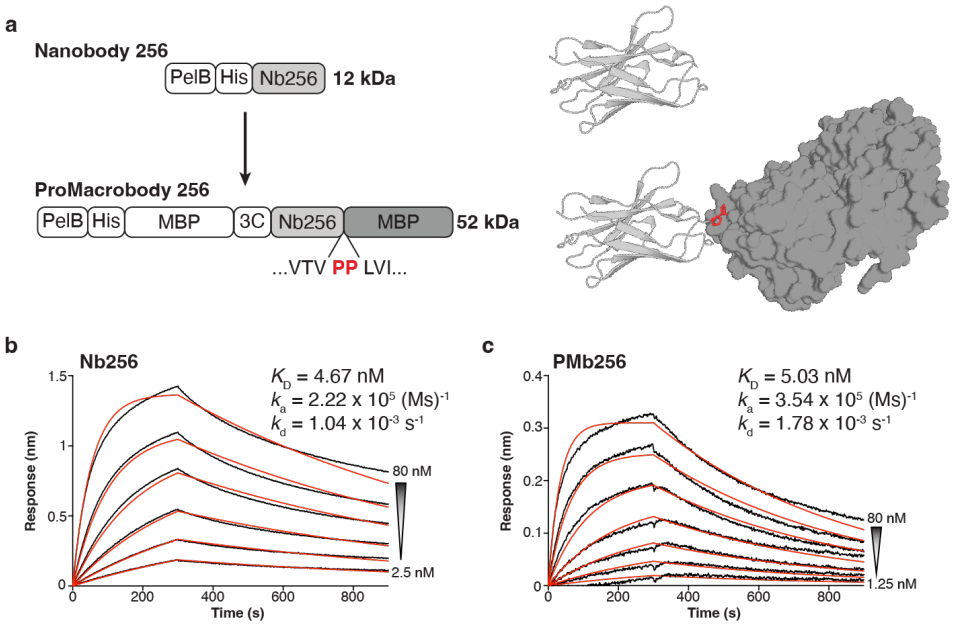
EXTENDED DATA FIGURES



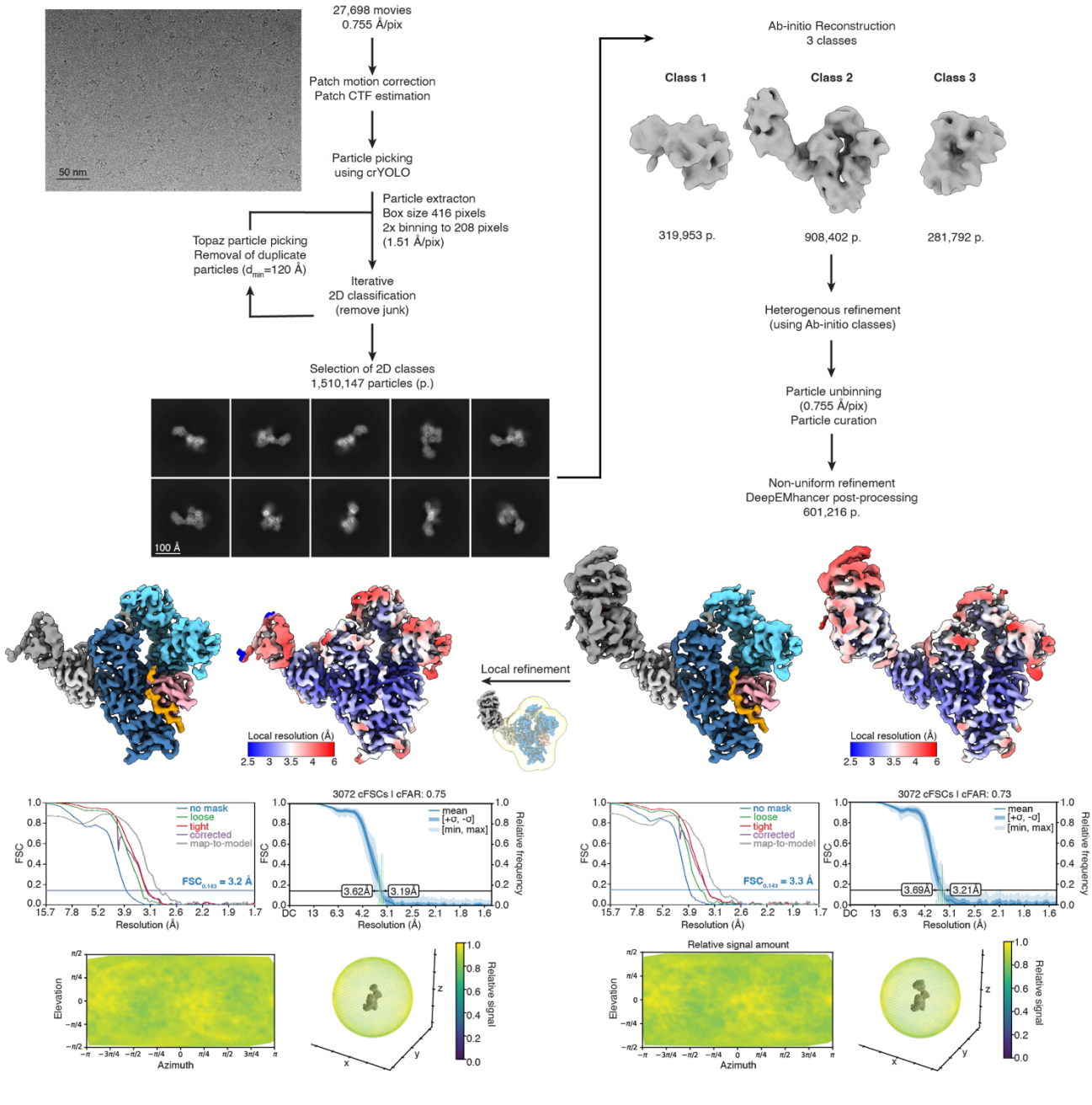
Extended Data Fig. 1. Recombinant Themis co-purifies with endogenous Grb2

a, Full-length C-terminal StrepII-tagged Themis (Themis^{FL}) was expressed and purified from HEK suspension cells. Purified recombinant Themis^{FL} was incubated at room temperature and samples were taken at indicated timepoints to check protein stability. Representative TGX gel and western blot analysis with an anti-StrepII tag antibody is shown. The predominant band at approximately 73 kDa corresponds to purified Themis^{FL}. Themis^{FL} loses its C-terminal StrepII tag and moves to a lower molecular weight (MW) over time. **b,** Themis was expressed and purified from HEK suspension cells. Representative TGX gel analysis of elution (E) fractions after affinity purification. The same gel was probed with specific anti-Themis and anti-Grb2 antibodies and detected with fluorescently tagged secondary antibodies. A fraction of endogenous cellular Grb2 was pulled down during Themis^{FL} purification from HEK suspension cells. **b,** SEC-MALLS analysis of co-expressed and purified Themis$^{1-563}$-Grb2 in complex with Nb256. The Themis$^{1-563}$-Grb2-Nb256 complex has a theoretical mass of 103 kDa.



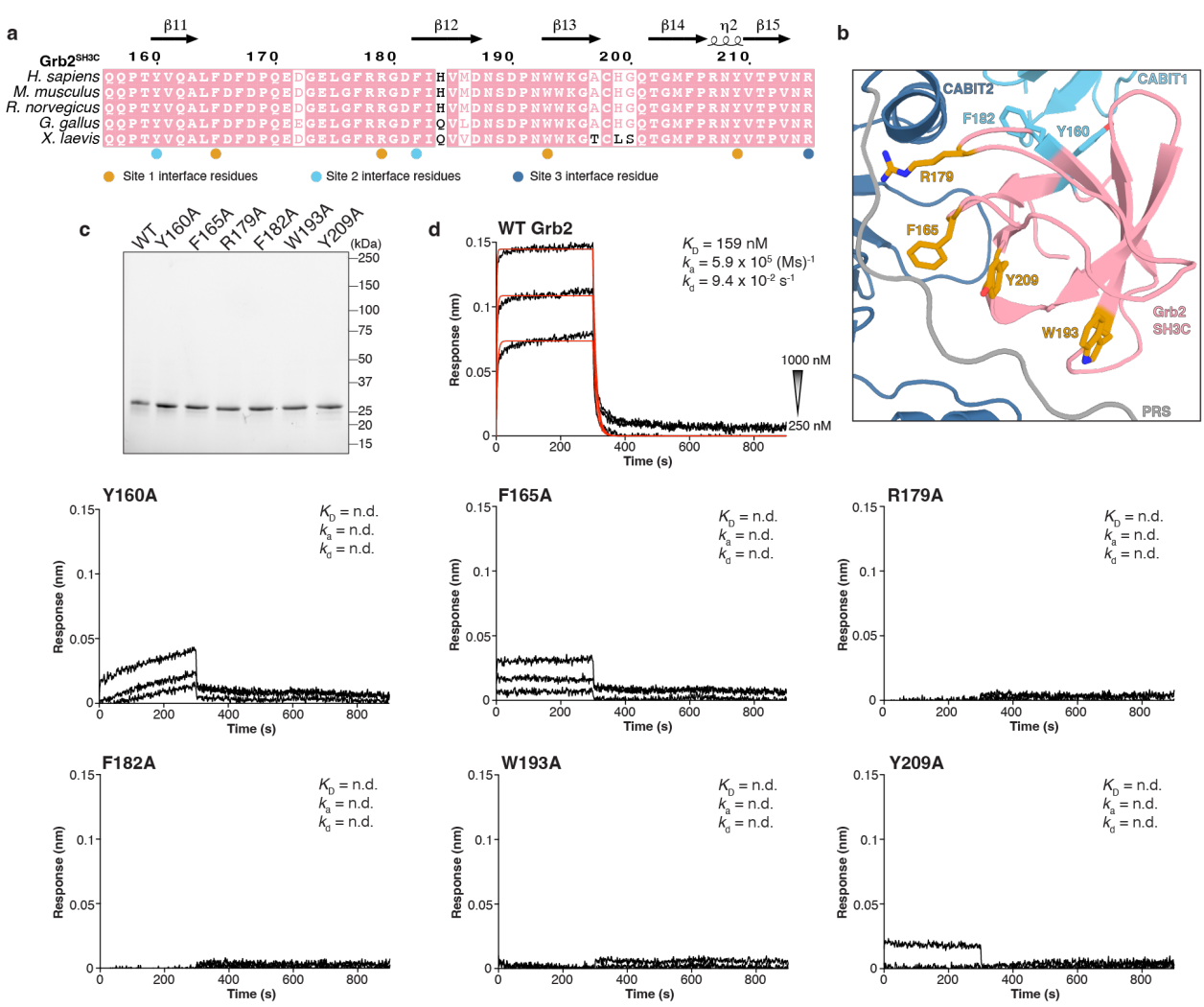
Extended Data Fig. 2. Conversion of Nb256 to PMb256

a, Comparison of the construct layout for the expression of Nb256 and PMb256. The original Nb construct contained a pelB leader sequence for periplasmic expression, followed by a hexahistidine (His6) tag and the Nb256 sequence. To generate PMb256, the Nb256 sequence was cloned into a pBXNPH3M expression vector containing a PelB leader sequence, followed by a decahistidine (His10) tag, MBP and 3C protease site. A truncated MBP moiety starting at residue Leu7 (LVI…) is encoded C-terminal to the Nb256 sequence, connected by a rigid double proline linker. Cartoon representations of Nb256 and PMb256 are shown. Structures were extracted from the cryo-EM model of the Themis$^{1-563}$-Grb2-PMb256 complex. The double proline linker is coloured in red and the MBP moiety is shown in surface representation. **b,** Representative BLI response curve to characterize binding of Themis1-563 to immobilized Nb256. **c,** Representative BLI response curve to characterize binding of Themis1-563 to immobilized PMb256. Response curves were fitted with a 1:1 model (red) to quantify the kinetics (k_a, k_d) and binding affinity (K_D).



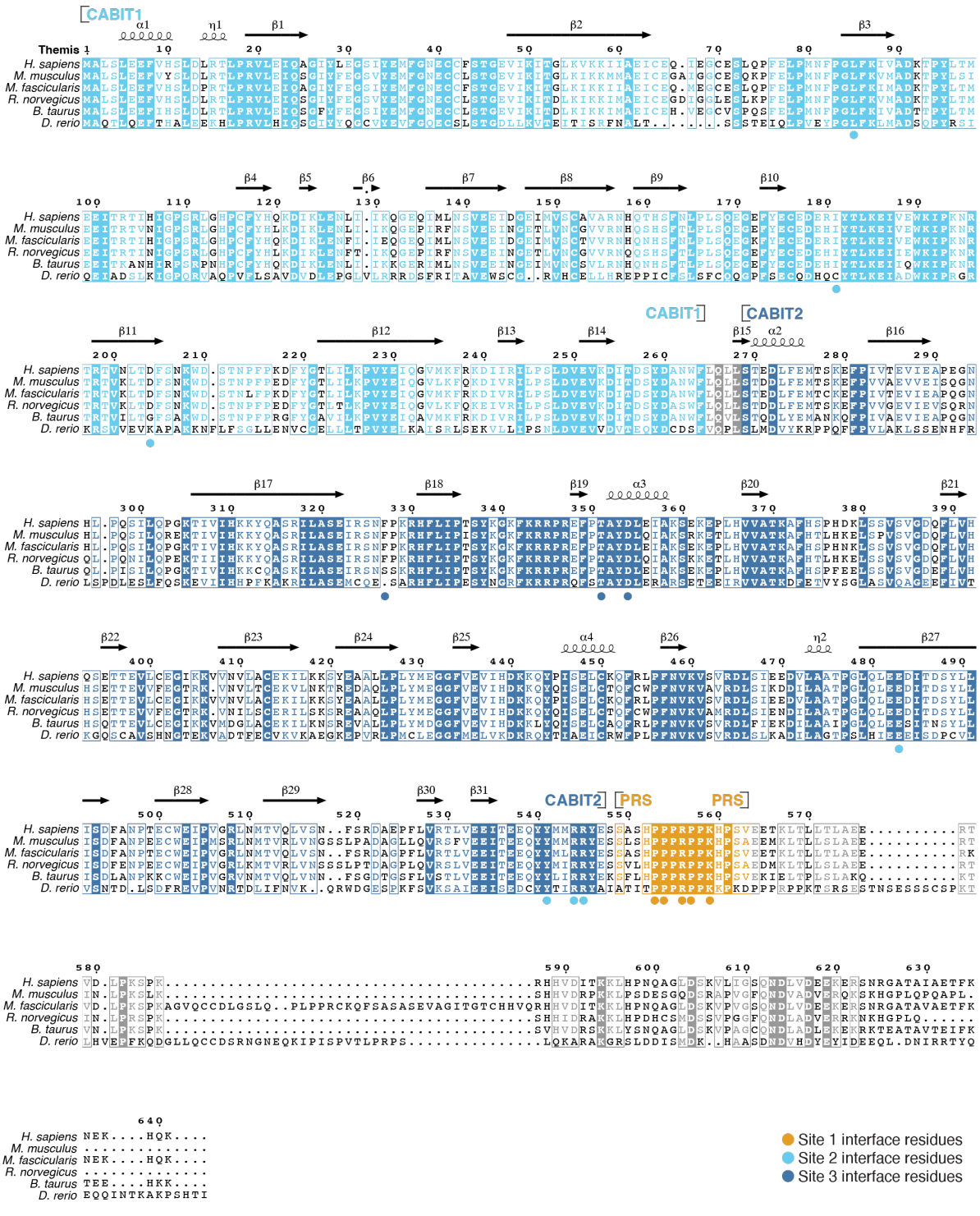
Extended Data Fig. 3. Cryo-EM data processing workflow for the Themis$^{1\text{-}563}$-Grb2-PMb256 complex

Data processing was performed in CryoSPARC v4.5.1[47,48] and map post-processing was performed using DeepEMhancer[53]. Gold standard Fourier Shell Correlation (FSC) curves are shown and the estimated resolution at FSC = 0.143 (blue line) is shown for the corrected FSC curves (purple line).



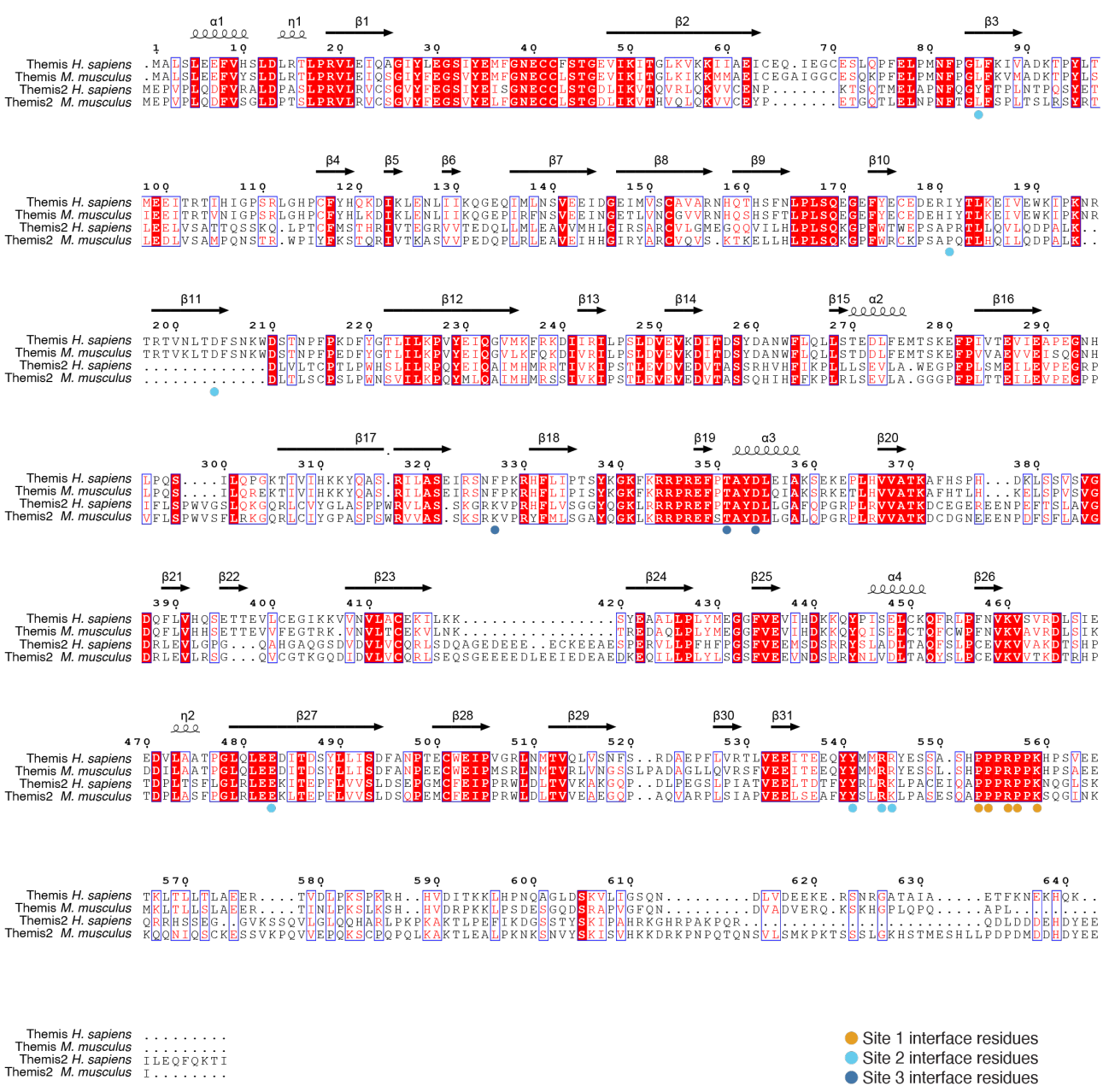
Extended Data Fig. 4. Functional interrogation of Grb2^{SH3C} mutants

a, Sequence alignment of various vertebrate Grb2^{SH3C} sequences using the ESPripT server (https://espript.ibcp.fr) and structural annotation according to secondary structure elements. Coloured symbols indicate residues participating in interaction sites 1-3. *H. sapiens* (Homo sapiens); *M. musculus* (Mus musculus); *R. norvegicus* (Rattus norvegius); *G. gallus* (Gallus gallus); *X. laevis* (Xenopus laevis). **b,** Zoom-in view of the Themis$^{1-563}$-Grb2^{SH3C} interaction interfaces. Site 1 and site 2 interface residues targeted by mutagenesis are shown in stick representation and coloured as in **a**. Grb2^{SH3C}, Themis^{PRS}, Themis^{CABIT1}, Themis^{CABIT2} are shown in pink, orange, cyan and dark blue, respectively. **c,** Representative SDS-PAGE analysis of purified Grb2^{SH3C} mutants used in BLI experiments. All mutant variants displayed similar biochemical behaviour to wild-type (WT) Grb2 **d,** BLI response curves for the interaction of immobilized WT or mutant Grb2FL variants with Themis$^{1-563}$. A 1:1 binding model was fitted (in red) to quantify the kinetics (k_a, k_d) and binding affinity (K_D). For all Grb2 mutants tested, kinetics and affinity could not be determined (n.d.) due to no binding or poor data fitting.



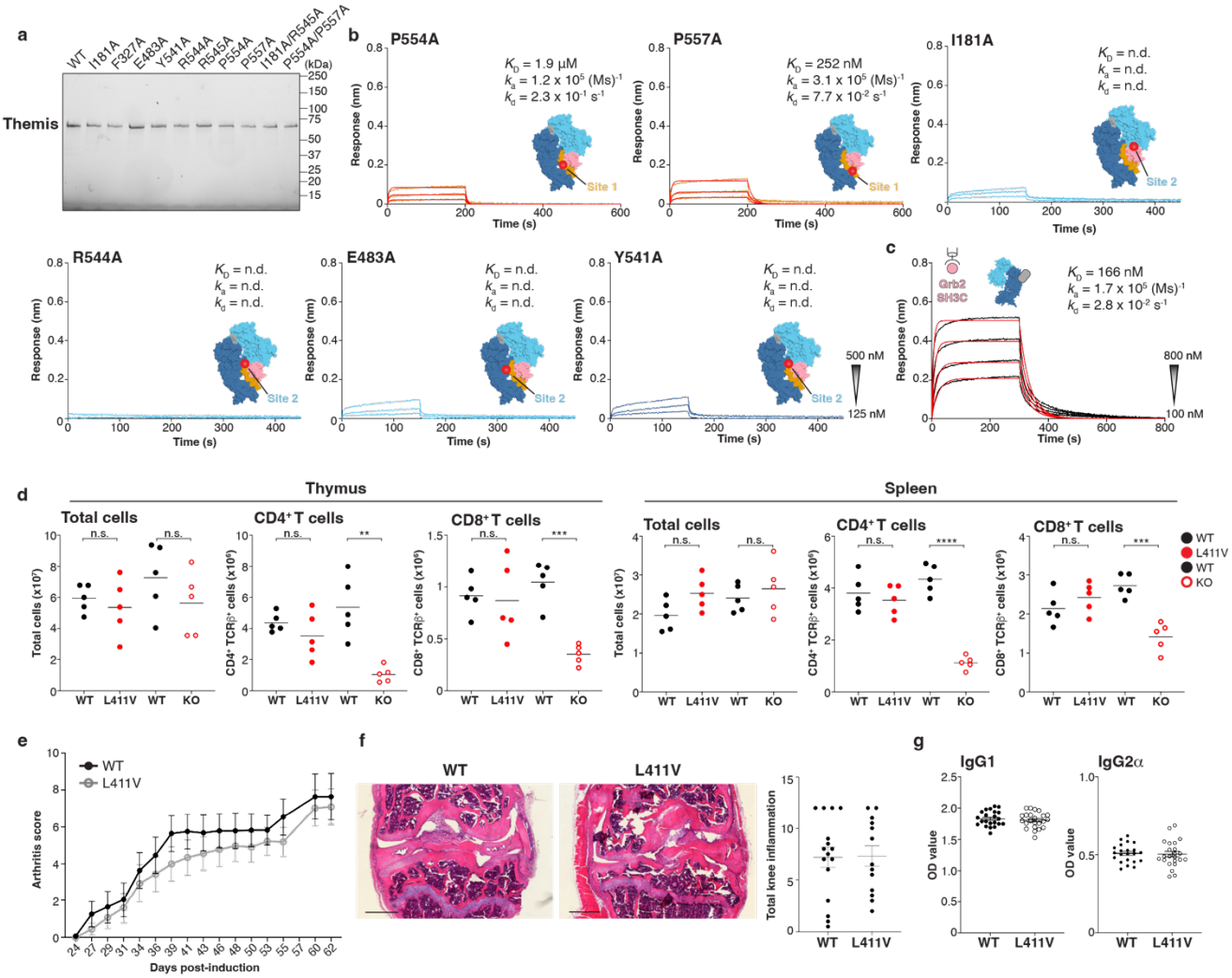
Extended Data Fig. 5. Conservation of Grb2^{SH3C} interfacing residues in Themis

Structurally annotated multiple sequence alignment of various vertebrate Themis sequences using the ESPripT server (https://espript.ibcp.fr). Coloured symbols indicate residues participating in interaction sites 1-3. CABIT1, CABIT2 and PRS domain delineations are indicated. *H. sapiens* (Homo sapiens); *M. musculus* (Mus musculus); *M. fascicularis* (Macaca fascicularis); *R. norvegicus* (Rattus norvegius); *B. taurus* (Bos taurus); *D. rerio* (Danio rerio).



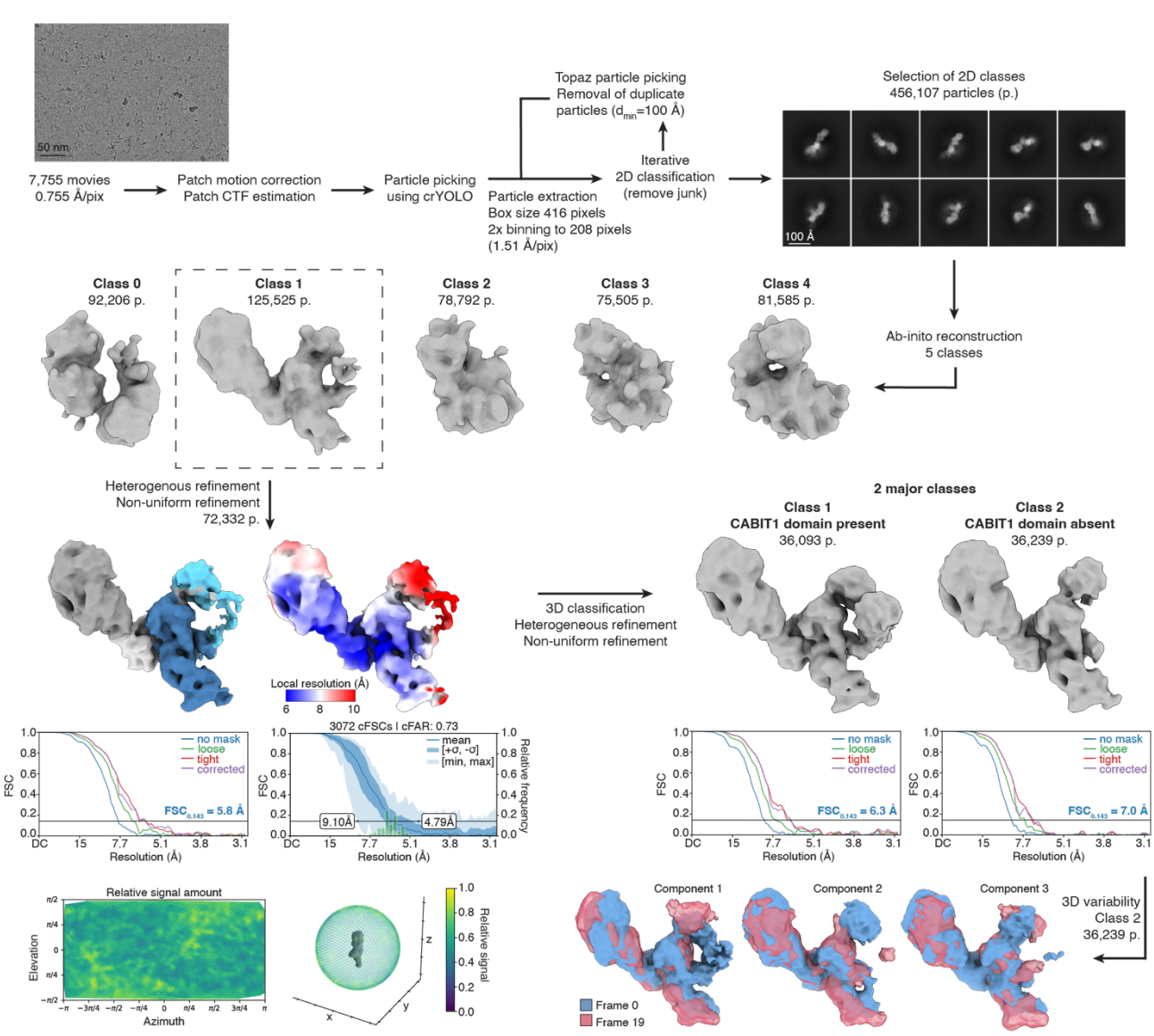
Extended Data Fig. 6. Sequence alignment of Themis and Themis2

Structurally annotated multiple sequence alignment of Themis and Themis2 sequences using the ESPripT server (https://espript.ibcp.fr). Coloured symbols indicate residues participating in interaction sites 1-3 with Grb2SH3C. *H. sapiens* (Homo sapiens); *M. musculus* (Mus musculus).



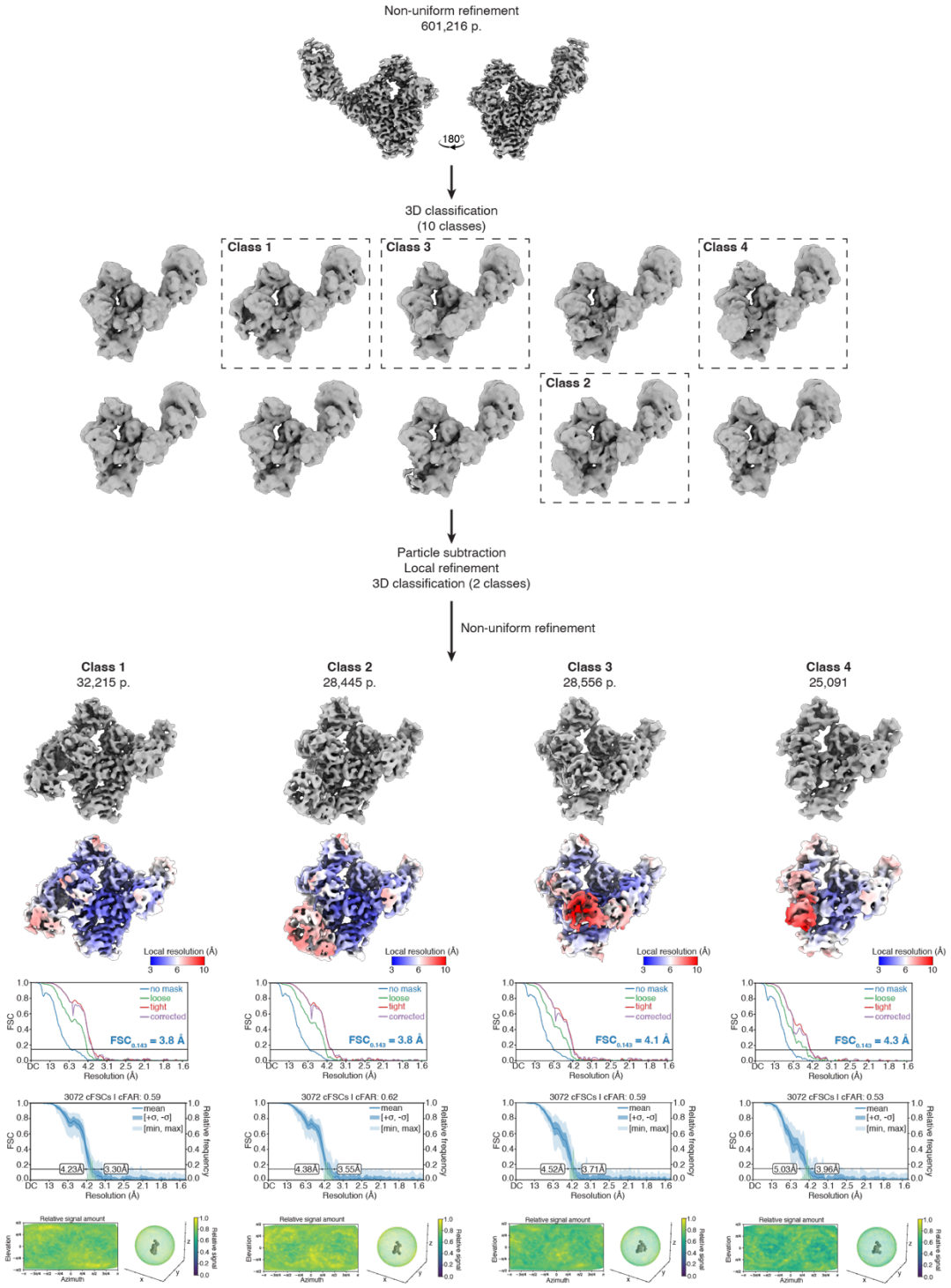
Extended Data Fig. 7. Functional interrogation of Themis mutants

a, SDS-PAGE analysis of purified Themis$^{1\text{-}563}$ mutants used in BLI experiments. All mutant variants displayed similar biochemical behaviour to wild-type (WT) Themis$^{1\text{-}563}$. **b,** BLI response curves for the interaction of immobilized WT Grb2^{FL} with Themis$^{1\text{-}563}$ mutants. A 1:1 binding model was fitted (in red) to quantify the kinetics (k_a, k_d) and binding affinity (K_D). Kinetics and affinity could not be determined (n.d.) for the indicated mutants due to no binding or poor data fitting. The location of each mutation in the Themis$^{1\text{-}563}$-Grb2^{SH3C} interface are indicated on a surface representation of the complex. **c,** BLI response curve for the interaction of immobilized GST-Grb2^{SH3C} with Themis$^{1\text{-}563}$ in complex with Nb256. A 1:1 binding model was fitted (in red) to quantify the kinetics (k_a, k_d) and binding affinity (K_D). For all BLI experiments, start and end concentrations of the 2-fold dilution series is shown as an inset. **d,** Total cellularity and quantification of CD4$^+$ and CD8$^+$ single positive T cell populations in the thymus and spleen of Themis^{KO} and Themis^{L411V} mice compared to wild-type (WT) littermates by flow cytometry. Data are from one experiment with five mice per group. Bars represent the mean of the data. n.s. non-significant, **p < 0.01, ***p < 0.001, ****p < 0.0001 by two-tailed t-test. **e,** Clinical arthritis scores were assessed in the ankle joints of mice after induction of CIA, in groups of male Themis^{L411V} transgenic mice (n=23) and wild-type (WT) littermates (n=23). Results are the mean ± SEM. **f,** Knee joints of male WT or Themis^{L411V} transgenic mice were examined for histological features of arthritis to confirm the clinical data. Hematoxylin and eosin staining was carried out to assess the amount of infiltrate and exudate in the femorotibial joints. Scale bars represent 500 µm. **g,** Serum antibody levels of IgG1 and IgG2α anti-collagen antibodies from WT and Themis^{L411V} mice were measured by ELISA.



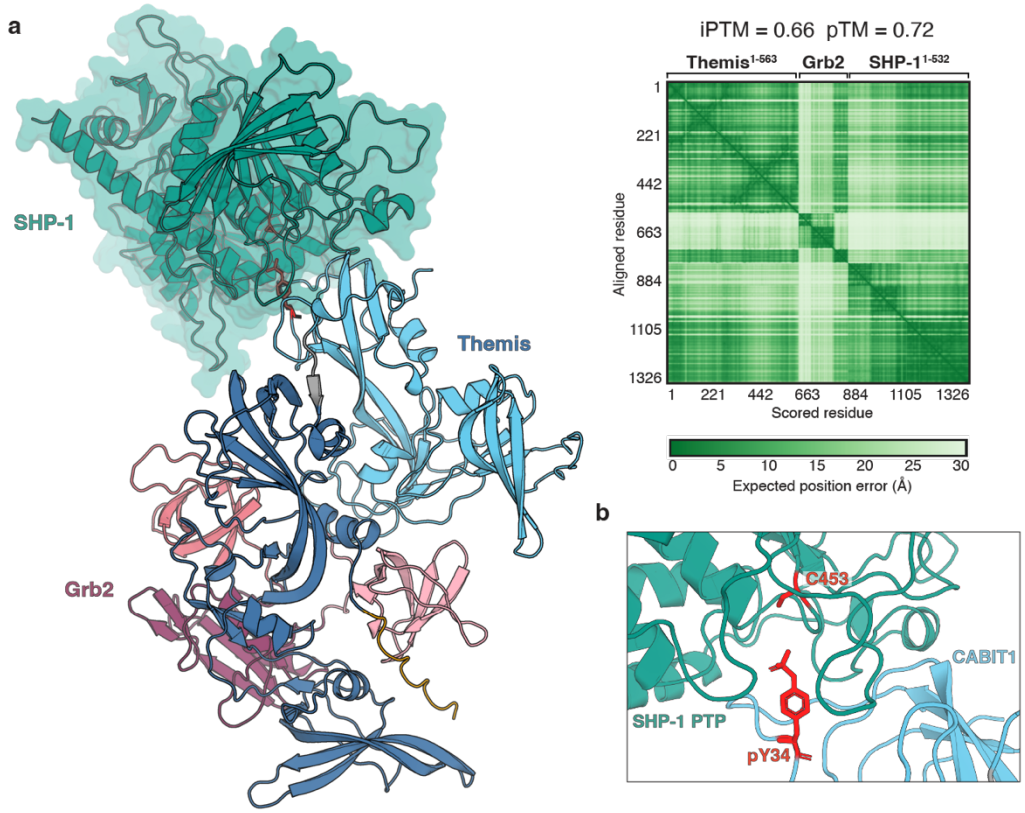
Extended Data Fig. 8. Cryo-EM data processing workflow for the Themis$^{1\text{-}563}$-PMb256 complex

Data processing was performed in CryoSPARC v4.5.1. Sharpened maps from non-uniform refinement in CryoSPARC are shown. Gold standard Fourier Shell Correlation (FSC) curves are shown and the estimated resolution at FSC = 0.143 (blue line) is shown for the corrected FSC curves (purple line).



Extended Data Fig. 9. 3D classification and data processing workflow to identify conformational heterogeneity of Grb2^{SH3N} and Grb2^{SH2} domains

Data processing was performed in CryoSPARC v4.5.1. Particles (601,216 p.) contributing to the Themis$^{1-563}$-Grb2-PMb256 map were further analysed in a 3D classification job without alignment. The resulting 4 classes exhibiting volume corresponding to the N-terminal domains of Grb2 were subjected to particle subtraction to remove the MBP moiety from the PMb followed by local refinement. Final particle numbers (p.) and sharpened maps from non-uniform refinement for each class are shown. Gold standard Fourier Shell Correlation (FSC) curves are shown and the estimated resolution at FSC = 0.143 (blue line) is shown for the corrected FSC curves (purple line).



Extended Data Fig. 10. AlphaFold3 predicts the formation of a Themis-Grb-SHP-1 ternary complex

a, AlphaFold3 prediction of Themis$^{1-563}$ in complex with Grb2^{FL} and SHP-1$^{1-532}$ using the Google Deepmind webserver (https://golgi.sandbox.google.com). Phosphorylation of Tyr34 (pY34) in Themis was included as a post-translational modification in the prediction for the complex. Truncated sequences for Themis and SHP-1 were used for prediction to exclude lower confidence scores caused by intrinsically disordered sequences at the C-termini of these proteins. Truncated SHP-1 lacking the flexible 61 residue C-terminal tail (residues 1-532) has previously been characterized using X-ray crystallography. Themis and Grb2^{FL} are coloured as before. SHP-1 is coloured in teal and depicted as a surface representation overlaying the cartoon representation. The iPTM and PTM scores and PAE plot are also included as confidence metrics. **b,** Zoom-in view of the predicted interface between Themis pY34 and the catalytic pocket of SHP-1 phosphatase (PTP) domain. Themis pY34 and the active site cysteine of SHP-1 (C453) are depicted in stick representation and coloured red.

Extended Data Table 1. Analysis of Themis$^{1-563}$:Grb2:PMb256 interaction interfaces

Interacting residues in the Themis$^{1-563}$:Grb2^{SH3C} and Themis$^{1-563}$:PMb256 interfaces as determined by the PISA server (https://www.ebi.ac.uk/pdbe/pisa/) and manual validation.

Themis$^{1-563}$:Grb2		
Themis$^{1-563}$	Grb2^{SH3C}	Distance (Å)
Hydrogen bonds		
ARG 556 [NH1]	GLU 174 [OE1]	3.07
ARG 111 [NH1]	ARG 179 [O]	3.60
ARG 111 [NH2]	ARG 179 [O]	3.74
ARG 111 [NH2]	ASP 181 [OD1]	3.59
ARG 180 [NH2]	ASP 181 [OD1]	3.05
TYR 541 [O]	GLN 162 [NE2]	3.09
GLU 547 [OE2]	ARG 179 [NH1]	2.96
GLU 483 [OE1]	ARG 179 [NH2]	2.75
PRO 557 [O]	TRP 193 [NE1]	3.56
PRO 558 [O]	TRP 193 [NE2]	3.58
ASP 354 [OD2]	ARG 207 [NH2]	2.45
TYR 353 [OH]	ASN 208 [ND2]	3.64
PRO 554 [O]	TYR 209 [OH]	3.13
TYR 541 [OH]	ASN 214 [N]	3.52
ASP 204 [OD1]	ARG 215 [NH2]	2.59
PHE 205 [O]	ARG 215 [NH2]	3.70
SER 325 [O]	ASN 216 [ND2]	3.55
Salt bridges		
LYS 559 [NZ]	GLU 171 [OE1]	3.44
ARG 556 [NH1]	GLU 174 [OE1]	3.07
ARG 556 [NH1]	GLU 174 [OE2]	3.65
ARG 111 [NH2]	ASP 181 [OD1]	3.59
ARG 180 [NH2]	ASP 181 [OD1]	3.05
GLU 547 [OE2]	ARG 179 [NH1]	2.96
GLU 547 [OE2]	ARG 179 [NH2]	3.68
GLU 483 [OE1]	ARG 179 [NH2]	2.75
ASP 354 [OD2]	ARG 207 [NH1]	3.53
ASP 354 [OD1]	ARG 207 [NH2]	3.63
ASP 354 [OD2]	ARG 207 [NH2]	2.45
ASP 204 [OD2]	ARG 215 [NH2]	3.48
ASP 204 [OD1]	ARG 215 [NH2]	2.59
Buried residues (numerical order)		
LEU 85	TYR 160	
ILE 181	PHE 165	
PHE 327	PHE 182	
THR 351		
ARG 544		
ARG 545		
PRO 553		

Themis$^{1-563}$:PMb256		
Themis$^{1-563}$	PMb256	Distance (Å)
Hydrogen bonds		
SER 462 [O]	TYR 32 [OH]	3.88
GLU 469 [O]	TYR 105 [N]	2.81
LEU 466 [O]	TYR 105 [OH]	3.50
GLU 469 [OE2]	TRP 108 [NE1]	3.72
ARG 464 [NH1]	ASP 99 [OD2]	2.52
ASP 471 [N]	VAL 103 [O]	3.24
Salt bridges		
GLU 469 [OE1]	ARG 45 [NH2]	3.94
GLU 470 [OE1]	ARG 104 [NH1]	3.99
ARG 464 [NH1]	ASP 99 [OD2]	2.52
Buried residues (numerical order)		
VAL 463	PHE 37	
ALA 474	LYS 43	
ASN 497	GLU 44	
LEU 527		