nature

https://doi.org/10.1038/s41586-022-05571-7

Accelerated Article Preview

# Antiviral signaling by a cyclic nucleotide activated CRISPR protease

Received: 6 December 2021

Accepted: 17 November 2022

Accelerated Article Preview

Cite this article as: Rouillon, C. et al. Antiviral signaling by a cyclic nucleotide activated CRISPR protease. *Nature* https://doi.org/10.1038/s41586-022-05571-7 (2022)

**Christophe Rouillon, Niels Schneberger, Haotian Chi, Katja Blumenstock, Stefano Da Vela, Katrin Ackermann, Jonas Moecking, Martin F. Peter, Wolfgang Boenigk, Reinhard Seifert, Bela E. Bode, Jonathan L. Schmid-Burgk, Dmitri Svergun, Matthias Geyer, Malcolm F. White & Gregor Hagelueken**

This is a PDF file of a peer-reviewed paper that has been accepted for publication. Although unedited, the content has been subjected to preliminary formatting. Nature is providing this early version of the typeset paper as a service to our authors and readers. The text and figures will undergo copyediting and a proof review before the paper is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers apply.

ACCELERATED ARTICLE PREVIEW

# Antiviral signaling by a cyclic nucleotide activated CRISPR protease

Christophe Rouillon[1,2,*,#], Niels Schneberger[1,#], Haotian Chi[3], Katja Blumenstock[4], Stefano Da Vela[5], Katrin Ackermann[6], Jonas Moecking[1], Martin F. Peter[1], Wolfgang Boenigk[2], Reinhard Seifert[2], Bela E. Bode[6], Jonathan L. Schmid-Burgk[4], Dmitri Svergun[5], Matthias Geyer[1], Malcolm F. White[3], Gregor Hagelueken[1,*]

[1] Institute of Structural Biology, University of Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

[2] Max Planck Institute for Neurobiology of Behavior – caesar, Ludwig-Erhard-Allee 2, 53175 Bonn, Germany

[3] School of Biology, University of St Andrews, North Haugh, St Andrews, KY16 9ST, UK.

[4] Institute of Clinical Chemistry and Clinical Pharmacology, University and University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

[5] EMBL Hamburg Unit, c/o DESY, Notkestr. 85, 22607, Hamburg, Germany

[6] EaStCHEM School of Chemistry, Biomedical Sciences Research Complex, and Centre of Magnetic Resonance, University of St Andrews North Haugh, St Andrews KY16 9ST, UK

[#]these authors contributed equally

[*]Corresponding author emails: back2crispr@gmail.com (C.R.), hagelueken@uni-bonn.de (G.H.)

ACCELERATED ARTICLE PREVIEW

**CRISPR defense systems such as the well-known DNA-targeting Cas9 and the RNA-targeting type III systems are widespread in prokaryotes[1,2]. The latter can orchestrate a complex antiviral response that is initiated by the synthesis of cyclic oligoadenylates (cOAs) upon foreign RNA recognition[3–5]. Among a large set of proteins that were linked to type III systems and predicted to bind cOAs[6,7], a CRISPR associated Lon protease (CalpL) stood out to us. The protein contains a sensor domain of the SAVED (SMODS-associated and fused to various effector domains) family[7], fused to a Lon protease effector domain. However, the mode of action of this effector was unknown. Here, we report the structure and function of CalpL and show that the soluble protein forms a stable tripartite complex with two further proteins, CalpT and CalpS, that are encoded in the same operon. Upon activation by $cA_4$, CalpL oligomerizes and specifically cleaves the MazF-homolog CalpT, releasing the extracytoplasmic function (ECF) sigma factor CalpS from the complex. This provides a direct connection between CRISPR-based foreign nucleic acid detection and transcriptional regulation. Furthermore, the presence of a $cA_4$-binding SAVED domain in a CRISPR effector reveals an unexpected link to the cyclic oligonucleotide-based antiphage signaling system (CBASS).**

ACCELERATED ARTICLE PREVIEW

**Main**

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) is a prokaryotic adaptative immune system that enables microorganisms to fend off attacks from mobile genetic elements such as phages, viruses, or plasmids[8]. The protein complex Cas1-Cas2 captures short DNAs from invaders and integrates them as "memories" into a CRISPR locus[9]. Transcripts of these "memories" are processed into small CRISPR RNAs (crRNAs) and integrated into large ribonucleoprotein (RNP) complexes, which can sense the presence of a matching foreign nucleic acid in the cell[10]. Once an invading sequence is detected, an antiviral response is triggered. Depending on the type of CRISPR system[1], this response can be markedly different, ranging from cleavage of the invading nucleic acid by the RNP as in the case of Cas9[11], to a complex multipronged defense strategy as found in type III CRISPR systems[12]. For the latter, the Cas10 subunit of the RNP has a cyclase activity that converts ATP into a recently discovered class of cyclic oligoadenylates (cOAs) upon viral RNA recognition[3–5]. The cOAs are constructed from 3 to 6, 3'-5' linked AMP units[13] and act as second messengers, typically by binding to proteins harboring a CARF (CRISPR-associated Rossmann-fold) domain[14]. There is a wide variety of CARF proteins linked to effector domains with functions ranging from RNA cleavage, supercoiled DNA nicking, dsDNA cleavage to transcription modulation[12,15–20]. The downstream effects of those cOA-activated proteins can lead to viral clearance, an abortive infection or a dormant state of the cell, enabling it to weather the phage attack[18,21].

Recently, two bioinformatic teams cataloged CARF-domain encoding genes that are likely linked to a functional type III system in bacterial and archaeal genomes[6,7]. Together, the studies revealed more than 100 such genes, including several membrane proteins and many proteins with currently unknown functions. Another study proposed that some of those type III-associated proteins contain a SAVED domain ('SMODS-associated and fused to various effectors domains'; SMODS being the acronym for 'second messenger oligonucleotide or dinucleotide synthetase'[22]) instead of CARF, reminiscent of the recently discovered CBASS system ('cyclic-oligonucleotide-based antiphage signaling systems'[23]). One of these proteins is CalpL (CRISPR associated Lon protease; initially termed Lon-CARF[6]), a 60 kDa protein with two predicted transmembrane helices, a Lon-protease domain and a SAVED4 domain[24].

Here we report the structure and function of the CalpL protein from the thermophilic bacterium *Sulfurihydrogenibium sp.* YO3AOP1. We find that CalpL is a soluble monomer and forms a 1:1:1 complex with CalpT and CalpS, encoded by adjacent genes in the locus. Once activated by $cA_4$, CalpL oligomerizes and proteolytically cleaves CalpT, releasing the $CalpT_{23}$/S complex, which has striking similarities to bacterial σ-factor/anti-σ-factor pairs.

ACCELERATED ARTICLE PREVIEW

**Structure of CalpL**

A synthetic, codon-optimized variant of the CalpL gene from *Sulfurihydrogenibium sp.* YO3AOP1 (UniProt ID B2V8L9) was expressed in *E. coli.* Although predicted to be a transmembrane protein[6,7], the protein was found in the soluble fraction of the cell lysate and behaved as a monomer during size exclusion chromatography (Extended Data Fig. 1a,b and below). CalpL was crystallised and the structure was solved at a resolution of 2.1 Å by single-wavelength anomalous dispersion phasing and refined to $R/R_{free}$ values of 19.3/22.5 (Fig. 1a, Extended Data Fig. 1c,d, Extended Data Table 1)[25,26].

The Lon protease domain consists of a four-stranded mixed β-sheet (β1-4), sandwiched between αD of the N-terminal domain and αG, H, J on the other side (Fig. 1a, Extended Data Fig. 1d). Various close structural homologs were identified and are listed in Extended Data Fig. 2. In CalpL, the catalytic Ser-Lys dyad, a hallmark of Lon proteases, is formed by S152 (loop β4-αH) and K193 (αJ) and lies at the end of a narrow channel that presumably binds the substrate peptide (Extended Data Fig. 2c). A superposition of the Lon-protease domain of CalpL with the acyl-enzyme intermediate state of the yellowfin ascites virus ATP-independent Lon protease[27] hints at the location of the P1 site in CalpL (Extended Data Fig. 2d). Structural modelling indicated that only amino acids with small hydrophobic side chains such as Ala or Gly can be accommodated in this site.

The C-terminal part of CalpL folds into a SAVED4 domain[24]. It consists of two pseudo-symmetric CARF-like domains with a pseudo-two-fold axis running between helices αP and αS (Fig. 1a). Interestingly, the TMHMM 2.0 server predicted that those helices and the directly preceding β-strands form transmembrane helices or at least membrane associated helices[6,28] (Extended Data Fig. 1b), which is clearly not the case. The SAVED domain has an extensive, positively charged cavity on its molecular surface, suited to bind a cOA ligand. The CARF- or SAVED-domains of the cOA-activated effector proteins Cap4, Cap5, and Can1 are structural homologs found in CBASS and CRISPR systems [17,23,29] (Extended Data Fig. 2b). Despite the low sequence identities (9-14% identical amino acids) the fold of the CARF-like domains is conserved. The position of the effector domain relative to the SAVED or CARF domains is, however, entirely different between the four structures (Extended Data Fig. 2e-g).

Surface plasmon resonance experiments showed that of the four tested cOAs CalpL selectively binds $cA_4$ with a dissociation constant of ~1 nM (Fig. 1b). We determined a 2.2 Å crystal structure of the CalpL/$cA_4$ complex and found the ligand bound to the SAVED domain at the expected position (Fig. 1b,c). As for the SAVED domain itself, the bound $cA_4$ molecule adopts a pseudo two-fold symmetry (Extended Data Fig. 3a). The cyclic tetra adenylate is

ACCELERATED ARTICLE PREVIEW

involved in a large number of polar and hydrophobic interactions with the SAVED domain, such that three of the four phosphate groups of the ligand are solvent exposed (Extended Data Fig. 3a). A comparison between the apo- and the $cA_4$-complex structure reveals no major conformational changes, apart from small shifts in the loops surrounding the $cA_4$ ligand (Extended Data Fig. 3b). A comparison with the $cA_3$ bound Cap4 structure (Extended Data Fig. 3c) illustrates how the loops on top of the SAVED domains shape the binding site for the cognate ligand in each structure.

The N-terminal domain of CalpL forms a bundle of six α-helices (αA-F, Fig. 1a, Extended Data Fig. 1d) with weak structural similarity to the N-terminal protein-protein interaction domain of Katanin p60-N in the p60p80-CAMSAP complex (Extended Data Fig. 2b)[30].

**CalpL specifically cleaves CalpT**

We used the WebFLAGs server[31] to study the gene neighborhood of CalpL homologs and noticed a small 812 bp open reading frame (271 amino acids, 31.8 kDa, UniProt ID B2V8L8) with no annotated function upstream of the *calpL* gene (Fig. 2a). We analyzed its sequence with HHPRED[32] and found homologies to the MazF toxin in the N-terminal half of the protein and weak homologies to DUF2080, a "domain of unknown function", in the C-terminal half (Fig. 2a). We predicted the structure with AlphaFold2[33]. The software produced a model of a two-domain protein with a ~23 kDa and a ~10 kDa domain connected by an apparently flexible linker (Fig. 2b). The structural model was submitted to the DALI server[34], revealing structural similarities to MazF-like toxins (N-terminal fragment) and various immunoglobulin fold containing proteins (C-terminal fragment). Interestingly, the predicted structure appears as a structural mimic of the MazEF complex with helices αA, αD, and αE blocking the region that binds to the ssRNA target of MazF in a similar fashion to MazE (Extended Data Fig. 4a,b)[35]. We investigated whether this protein (named "CalpT" for "target") is cleaved by the CalpL protease. The gene was expressed in *E. coli* and the protein purified to near homogeneity (Extended Data Fig. 4c). CalpL, CalpT, and different cOAs (3, 4, 5, 6) were mixed at 1:1:1.5 molar ratios and incubated at 60 °C for one hour. Strikingly, we found that in the presence of $cA_4$, CalpT was cleaved by CalpL. SDS-PAGE analysis revealed two distinct cleavage products with molecular weights of 23 kDa and 10 kDa, respectively, suggesting a single cleavage site (Fig. 2c). The activity for the other cOAs was significantly lower. We repeated the experiment with an S152A variant of CalpL, which lacks the nucleophilic serine needed for its peptidase activity. Since this variant showed no protease activity, the CalpL protease active site is responsible for the observed proteolytic activity (Fig. 2c).

ACCELERATED ARTICLE PREVIEW

The peptide sequences of the two cleavage fragments were determined with peptide mass fingerprinting (Extended Data Fig. 4d). This analysis confirmed that the 23 kDa ($CalpT_{23}$) fragment corresponds to the N-terminal two-thirds of the CalpT protein and the 10 kDa fragment ($CalpT_{10}$) to the C-terminal one-third. Based on this result and considering the predicted structure (Fig. 2b), we mapped the location of the cleavage site to the stretch of residues between amino acids ~170-200 of CalpT. As mentioned above, our CalpL structure suggested that only peptides containing an alanine or glycine as the P1 residue will fit into the active site of CalpL. We therefore created glutamic acid mutants of all four alanine residues in the cleavage region: A172, A182, A195, and A201 (Fig. 2b, magenta spheres; the stretch of residues does not contain any glycine). Peptidase assays with all four CalpT variants were conducted and only the A195E mutation abolished the cleavage completely (Extended Data Fig. 4e). At this position, the amino acid sequence reads $V_{190}$LRHVA|ST, where A195 is most likely the P1 residue. Notably, A195 is conserved amongst CalpT homologs (Fig. 2d). The peptide fingerprint data in Extended Data Fig. 4d also supports this conclusion, as for $CalpT_{23}$, the peptide coverage extended almost exactly to the identified cleavage site. We did not observe any non-tryptic peptides that corresponded to the identified cleavage site. Thus, the exact molecular weights of the CalpT cleavage products are 23.0 kDa ($CalpT_{23}$) and 8.7 kDa ($CalpT_{10}$), fitting to the sizes observed in SDS PAGE analysis (Fig. 2c).

**CalpL and CalpT form a 1:1 complex**

To test whether CalpL and CalpT form a stable complex, we analysed the individual proteins and their equimolar mixtures by SEC-MALS (Fig. 3a). CalpL alone eluted in a single peak at 17.1 ml and the CalpT protein eluted at 17.9 ml, both at the expected molecular weights for the monomeric proteins. The 1:1 mixture of the two proteins resulted in a single elution peak at 16.2 ml. The $MW_{MALS}$ of the complex was 82.4 kDa, suggestive of a 1:1 complex of the CalpL and CalpT proteins (52 + 30.5 kDa). We used surface plasmon resonance (SPR) to quantify the interaction strength between CalpL and -T and found that the two proteins form a very strong complex with a sub-nanomolar $K_D$ (Extended Data Fig. 5a). Interestingly, a similar affinity was observed for a construct where $CalpT_{10}$ (including the cleavage site) was fused to a VHH domain targeting an unrelated protein (Extended Data Fig. 5b,c). The artificial construct was readily cleaved by the protease upon activation by $cA_4$ (Extended Data Fig. 5d). Thus, the $CalpT_{23}$ fragment plays no important role either in the formation of the CalpL/T complex or in the cleavage process. A second artificial construct, where the $CalpT_{10}$ moiety was also replaced

ACCELERATED ARTICLE PREVIEW

by an unrelated VHH was not cleaved (Extended Data Fig. 5d). Hence, the CalpT$_{10}$ subunit is required for cleavage.

To follow the fate of the complex after cleavage, we repeated the experiment in presence of an excess (1:1.1) of cA$_4$ (Fig. 3a, violet). Here, we observed three peaks corresponding to CalpL/T$_{10}$, CalpT$_{23}$ and cA$_4$. Correspondingly, for the inactive CalpL S152A variant, the CalpL/T complex was observed but the addition of cA$_4$ did not lead to the observed split into three peaks (Extended Data Fig. 5e). We also checked whether the four cleavage-site variants of CalpT could still form a complex with CalpL (Extended Data Fig. 5f). Whereas A172E, A182E, and the P1 site variant A195E did, the A201E variant did not. The glutamate at this position apparently weakened the interaction, explaining the reduced cleavage efficiency of this mutant (Extended Data Fig. 4e).

We isolated the CalpL/T$_{10}$ complex for crystallization and determined its structure at 3.3 Å resolution by molecular replacement, using the CalpL crystal structure and the AlphaFold model of CalpT$_{10}$ as search models (Fig. 3a, Extended Data Fig. 5g, Extended Data Table 1). Indeed, the CalpT$_{10}$ fragment binds to the N-terminal domain of CalpL and, as indicated by the mutation analysis, A201 of CalpT is part of the interface in addition to the CalpL hydrophobic residues W28, L6, V14, L18, E20, E13, K8, H2 and CalpT residues K200, Y210, Y203, E222 (Extended Data Fig. 5g). A small-angle X-ray scattering (SAXS) experiment was performed to exclude a crystal packing artefact, by measuring the SAXS profiles of CalpL and CalpL/T$_{10}$ complex by SEC-SAXS (Extended Data Table 2, experimental session I). The two profiles were fitted simultaneously using the multi-phase *ab initio* shape reconstruction program MONSA. The *ab initio* model was in excellent agreement with the SAXS data and compares well to our crystal structure (Extended Data Fig. 5h), thus confirming the arrangement of the two subunits in solution.

**cA$_4$ induced oligomerization of CalpL/T**

Intriguingly, while the C-terminal part of the CalpL cleavage site (T197 of CalpT) is visible in the complex crystal structure, it is more than 35 Å away from the protease active site (Fig. 3a), indicating that a cA$_4$ induced structural rearrangement of CalpL must occur to allow cleavage of CalpT. Recent studies on other SAVED-domain containing CBASS effectors demonstrated that cOA binding induces an oligomerization, which then activates the effector[23,29,36]. Dynamic light scattering (DLS) and SAXS experiments showed such a cA$_4$- and protein concentration-dependent oligomerization of CalpL (Fig. 3b, Extended Data Fig. 6a,b). A representative *ab initio* model for monomeric CalpL in the presence of cA$_4$ was obtained by SEC-SAXS

ACCELERATED ARTICLE PREVIEW

(Extended Data Table 2), resulting in an elongated and slightly bent model featuring two main lobes connected by a slightly thinner region at SAXS resolution. Using this monomeric unit, the concentration series in the range 2-5 mg/mL was modeled as a dimerizing mixture by a global SASREFMX fitting without imposing symmetry elements (as non-identical binding interfaces are to be expected in the presence of the $cA_4$ ligand). SAXS modelling of the dimerizing mixture produced stable solutions (normalized spatial discrepancy (NSD) ~0.95) featuring elongated shapes that were large enough to accommodate two CalpL molecules (Extended Data Fig. 6c).

We noticed a distinct positively charged patch on the face opposite of the $cA_4$ binding site, where R361, R338 and K364 coordinate a sulfate ion in the $cA_4$ complex structure (Fig. 1c). Due to this structural feature, binding of the negatively charged $cA_4$ molecule would result in a charge complementarity between the top and bottom sides of the SAVED domain. This supports an arrangement where two or more CalpL molecules would form stacks, with $cA_4$ sandwiched in between, similar to the architecture observed for other SAVED domain oligomers[23,29,36,37]. The SAXS *ab initio* model in Extended Data Fig. 6c would best agree with a staggered arrangement of the CalpL monomers. An attractive model for the activation of CalpL by $cA_4$ would thus be an in-trans cleavage reaction in the observed oligomers. To test this, we performed a cleavage assay, were mixtures of preformed CalpL/T complexes, for instance CalpL/T (wt/wt), CalpL/T (S152A/wt), CalpL/T (wt/A195E), were tested for $cA_4$ induced cleavage. A 1:1 mixture of CalpL/T (S152A/wt) with CalpL/T (wt/A195E), *i.e.*, two complexes that are not capable of in-cis cleavage, led to 50% cleaved CalpT (Fig. 3c). The remaining 50% could not be cleaved due to the A195E mutation, but this could be titrated by changing the ratio of the two complexes (Fig. 3c). Further, a mixture of CalpL/T (S152A/wt) with uncomplexed CalpL led to complete cleavage of CalpT. To support the idea of oligomerization induced cleavage, we introduced mutations to the backside of the SAVED domain, aiming to disturb the presumed oligomerization interface (Fig. 3d and 1c). While mutant R361E had a ~50% reduced activity, R338E had no cleavage activity. We also found that the R493C mutant used for the SPR experiments had a ~50% reduced cleavage activity.

While our data show that in-trans cleavage occurs, we can currently not distinguish, whether cleavage occurs inside one particular CalpL/T oligomer or between two oligomers. For the latter, three CalpL/T units would have to assemble for the VA/ST sequence in CalpT to be able to reach the protease active site of a CalpL molecule in the oligomer. A high-resolution structure of such a CalpL/T stack will be necessary to unravel the molecular details of the activation.

ACCELERATED ARTICLE PREVIEW

**CalpL/T bind the ECF σ factor CalpS**

Our initial assumption that CalpT will have a MazF-like nuclease activity could not be confirmed experimentally. We could neither identify any signs of an RNase activity biochemically (Extended Data Fig. 7a-c), nor with RNase-seq of random libraries (Extended Data Fig. 7d,e). Moreover, expression of $CalpT_{23}$ constructs in *E. coli* were tolerated by the cells (not shown). We also looked at a dimerization of the $CalpT_{23}$ fragment, similar to the active dimeric MazF enzyme. To test this, we spin labelled the CalpL/CalpT complex at position 119 of CalpT (Extended Data Fig. 7f,g) and measured the interspin distance in the presence and absence of $cA_4$. According to the EPR data, the $cA_4$ induced cleavage did not lead to changes of the conformational state of the $CalpT_{23}$ (Extended Data Fig. 7h-j).

All this turned our attention towards CalpS, a third conserved protein encoded by the operon (224 amino acids, 26.5 kDa, UniProt ID B2V8L7). The protein has strong sequence similarities to ECF family σ factors, which tailor transcription in diverse stress conditions[32,38,39] (Extended Data Fig. 8a). Interestingly, AlphaFold2 supported a heterotrimeric complex between CalpL, -T and -S, which was consistent with our finding that the $CalpT_{23}$ domain is not involved in CalpL/T complex formation (Fig. 4a). The predicted CalpT/S interface has a combined buried surface area of ~4000 $Å^2$ [40], high confidence scores in the interface area and convincing sidechain interactions. In the prediction, CalpT binds to both the $σ_2$- and $σ_4$-domains of the σ-factor and blocks most of the -10-region interface (Extended Data Fig. 8bc). The same interface is targeted by so called anti-σ-factors, preventing the interaction of the σ-factor/RNA-polymerase (RNAP) complex with its cognate promotor (Extended Data Fig. 8d)[38,39]. Note that in the predicted CalpT/S complex, the $σ_2$- and $σ_4$ domains are tied together in a way that would not allow the σ-factor to bind to the RNAP (Extended Data Fig. 8e).

To put the existence of this complex to the test, we cloned the *calpS* gene and co-expressed the His-CalpT/S proteins in *E. coli*. As predicted, the two proteins formed a stable complex that could be isolated by gel filtration (Fig. 4b). Furthermore, addition of CalpL to the CalpT/S complex resulted in a ternary complex, that disintegrated into $CalpL/T_{10}$ and $CalpT_{23}$/S upon addition of $cA_4$ (Fig. 4bcd).

We noticed that expression of CalpS alone (instead of coexpression with CalpT) led to a copurification of the protein with the α- and β subunits of the DNA-directed RNA polymerase of *E. coli* (44% sequence identity between β subunits of *Sulfurihydrogenibium sp.* and *E. coli* RNAP). This corroborates the prediction that CalpS is a σ-factor and that CalpT inhibits its interaction with the RNAP (Fig. 4e). Thus, CalpT has striking functional similarities to anti-σ-

ACCELERATED ARTICLE PREVIEW

factor proteins and literally links the $cA_4$ sensor CalpL to the transcription machinery of the cell.

## Discussion

Protease signaling cascades are a common scheme in evolution that are often employed in emergency situations. Prokaryotic type-II toxin-antitoxin (TA) systems, for instance, are activated by degradation of the antitoxin by ATP-dependent Lon proteases[41]. The innate immune system of higher organisms also employs proteases such as caspases to initiate and amplify fast responses to external threats. Another well-known example are the cascades of proteases that control the clotting of blood[42]. Our work shows that the CalpL/T/S cascade amalgamates aspects of different defense systems such as CRISPR, CBASS, toxin/antitoxin systems, and σ/anti-σ-factors into a $cA_4$-controlled "fast response" signaling cascade.

Our data are summarized in the model sketched in Fig. 5, where in its inactive state, CalpL will be present in a 1:1:1 complex with CalpT and CalpS. Upon detection of a foreign RNA by the type III effector complex, $cA_4$ will be synthesized by its Cas10 subunit and the second messenger will bind to the SAVED domain of CalpL with nanomolar affinity. This drastically changes the surface electrostatics of the SAVED domain, enabling oligomerization of CalpL, as observed by SAXS and DLS (Fig. 3b and Extended Data Fig. 6). This activation mechanism has now emerged as a common theme in the SAVED-domain based effectors of CBASS defense systems[23,29,36] but has not been observed in CRISPR systems. The CalpL oligomerization triggers an in-trans cleavage of CalpT (Fig. 3c,d), releasing the $CalpT_{23}$/S subcomplex in a strictly $cA_4$ dependent manner (Fig. 4bcd). Some CARF domain proteins are known to auto deactivate by degrading cOA species [12], but this has not been observed for SAVED domains and has not yet been investigated for CalpL.

CalpS is member of the ECF family of ECF σ factors, which play a role in the sensing of extracellular stress events, such as cell envelope- or oxidative stress[38,39]. In striking resemblance to the Calp cascade, the release of such anti-σ factors is orchestrated by a sequence of proteolytic events called regulated intermembrane proteolysis (RIP)[43]. The activation of $\sigma^E$ from *E. coli*, for instance, proceeds via proteolytic cleavage of the membrane bound anti sigma factor RseA, releasing a soluble σ-factor/anti-σ-factor complex. The anti-σ-factor is subsequently degraded by ATP-dependent ClpXP proteases. [38,44]. Following the established paradigms, one might speculate that further proteolysis of $CalpT_{23}$ releases the sigma factor CalpS to allow transcriptional response.

ACCELERATED ARTICLE PREVIEW

Recently, two studies have revealed that type III-E CRISPR systems also function by activating the protease Csx29 (also known as TRP-CHAT) [45,46]. There are notable differences, as the protease (from the Caspase family) is completely unrelated to CalpL and is not activated by cOA. Csx29, part of the Type III-E effector complex, is activated by conformational changes upon foreign RNA detection and cleaves an uncharacterized protein (Csx30) encoded in the operon [47]. Interestingly, Csx30 binds a σ-factor also homologous to the ECF family (termed CASP-σ). Furthermore, it was shown that the Csx30 protein inhibits CASP-σ and this inhibition is relieved by Csx29 mediated proteolytic cleavage. CASP-σ has a high affinity for a DNA sequence that is found in the promotor of Cas1-2, proteins of CRISPR adaptation for the acquisition of new viral memories [48]. Since the CRISPR effectors and the two proteases are completely unrelated, this appears to be a striking example of convergent evolution.

Here, we have uncovered a cOA-mediated signaling cascade from viral RNA detection to the proteolytic release of a σ-factor that binds RNA polymerase. Notably, the Cas10 proteins associated with Calp operons lack an HD nuclease domain and auxiliary cOA activated nucleases such as Csx1 are rarely found, suggesting transcriptional changes sufficient for CRISPR antiviral immunity in organisms such as *Sulfurihydrogenibium.* It will be exciting to find out the DNA targets of CalpS to understand better how the Calp cascade shapes the antiviral response, buying the organism enough time to survive a viral attack[18,21].

ACCELERATED ARTICLE PREVIEW

**Main text references**


1. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Classification and nomenclature of CRISPR-Cas systems: where from here. *The CRISPR Journal* 1, 325-336 (2018).
2. Zhu, Y., Klompe, S. E., Vlot, M., van der Oost, J. & Staals, R. H. J. Shooting the messenger: RNA-targetting CRISPR-Cas systems. *Bioscience reports* 38, BSR20170788 (2018).
3. Kazlauskiene, M., Kostiuk, G., Venclovas, Č., Tamulaitis, G. & Siksnys, V. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* 357, 605-609 (2017).
4. Niewoehner, O., Garcia-Doval, C., Rostøl, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A. & Jinek, M. Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* 548, 543-548 (2017).
5. Rouillon, C., Athukoralage, J. S., Graham, S., Grüschow, S. & White, M. F. Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* 7, e36734 (2018).
6. Shmakov, S. A., Makarova, K. S., Wolf, Y. I., Severinov, K. V. & Koonin, E. V. Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proc Natl Acad Sci U S A* 115, E5307-E5316 (2018).
7. Shah, S. A., Alkhnbashi, O. S., Behler, J., Han, W., She, Q., Hess, W. R., Garrett, R. A. & Backofen, R. Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families. *RNA Biol* 16, 530-542 (2019).
8. Gasiunas, G., Sinkunas, T. & Siksnys, V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cell Mol Life Sci* 71, 449-465 (2014).
9. Sasnauskas, G. & Siksnys, V. CRISPR adaptation from a structural perspective. *Curr Opin Struct Biol* 65, 17-25 (2020).
10. Jiang, F. & Doudna, J. A. CRISPR–Cas9 Structures and Mechanisms. *Annual Review of Biophysics* 46, 505-529 (2017).
11. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816-821 (2012).
12. Athukoralage, J. S. & White, M. F. Cyclic oligoadenylate signalling and regulation by ring nucleases during type III CRISPR defence. *RNA* rna.078739.121 (2021).
13. Jia, N., Jones, R., Sukenick, G. & Patel, D. J. Second messenger cA4 formation within the composite Csm1 Palm pocket of type III-A CRISPR-Cas Csm complex and its release path. *Molecular cell* 75, 933-943. e6 (2019).
14. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V. & Aravind, L. CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Frontiers in genetics* 5, 102 (2014).
15. Lau, R. K., Ye, Q., Birkholz, E. A., Berg, K. R., Patel, L., Mathews, I. T., Watrous, J. D., Ego, K., Whiteley, A. T. & Lowey, B. Structure and mechanism of a cyclic trinucleotide-activated bacterial endonuclease mediating bacteriophage immunity. *Molecular cell* 77, 723-733. e6 (2020).
16. Lintner, N. G., Frankel, K. A., Tsutakawa, S. E., Alsbury, D. L., Copié, V., Young, M. J., Tainer, J. A. & Lawrence, C. M. The structure of the CRISPR-associated protein Csa3 provides insight into the regulation of the CRISPR/Cas system. *Journal of molecular biology* 405, 939-955 (2011).
17. McMahon, S. A., Zhu, W., Graham, S., Rambo, R., White, M. F. & Gloster, T. M. Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. *Nat Commun* 11, 500 (2020).


ACCELERATED ARTICLE PREVIEW

18. Rostøl, J. T., Xie, W., Kuryavyi, V., Maguin, P., Kao, K., Froom, R., Patel, D. J. & Marraffini, L. A. The Card1 nuclease provides defence during type III CRISPR immunity. *Nature* 590, 624-629 (2021).
19. Garcia-Doval, C., Schwede, F., Berk, C., Rostøl, J. T., Niewoehner, O., Tejero, O., Hall, J., Marraffini, L. A. & Jinek, M. Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. *Nature communications* 11, 1-9 (2020).
20. Lawrence, C. M., Charbonneau, A. & Gauvin, C. Cyclic Tetra-Adenylate (cA4) Activates CRISPR Associated Transcription Factor Csa3, Providing Feedback Activation of Protospacer Acquisition and crRNA Expression. *The FASEB Journal* 34, 1-1 (2020).
21. Meeske, A. J., Nakandakari-Higa, S. & Marraffini, L. A. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature* 570, 241-245 (2019).
22. Burroughs, A. M., Zhang, D., Schäffer, D. E., Iyer, L. M. & Aravind, L. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res* 43, 10633-10654 (2015).
23. Lowey, B., Whiteley, A. T., Keszei, A. F. A., Morehouse, B. R., Mathews, I. T., Antine, S. P., Cabrera, V. J., Kashin, D., Niemann, P., Jain, M., Schwede, F., Mekalanos, J. J., Shao, S., Lee, A. S. Y. & Kranzusch, P. J. CBASS Immunity Uses CARF-Related Effectors to Sense 3'-5'- and 2'-5'-Linked Cyclic Oligonucleotide Signals and Protect Bacteria from Phage Infection. *Cell* 182, 38-49.e17 (2020).
24. Makarova, K. S., Timinskas, A., Wolf, Y. I., Gussow, A. B., Siksnys, V., Venclovas, Č. & Koonin, E. V. Evolutionary and functional classification of the CARF domain superfamily, key sensors in prokaryotic antivirus defense. *Nucleic acids research* 48, 8828-8847 (2020).
25. Zwart, P. H., Afonine, P. V., Grosse-Kunstleve, R. W., Hung, L.-W., Ioerger, T. R., McCoy, A. J., McKee, E., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K., Storoni, L. C., Terwilliger, T. C. & Adams, P. D. Automated structure solution with the PHENIX suite. *Methods in Molecular Biology (Clifton, N.J.)* 426, 419-435 (2008).
26. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D* 66, 12-21 (2010).
27. Chung, I. Y. & Paetzel, M. Crystal structures of yellowtail ascites virus VP4 protease: trapping an internal cleavage site trans acyl-enzyme complex in a native Ser/Lys dyad active site. *J Biol Chem* 288, 13068-13081 (2013).
28. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes. *Journal of Molecular Biology* 305, 567-580 (2001).
29. Fatma, S., Chakravarti, A., Zeng, X. & Huang, R. H. Molecular mechanisms of the CdnG-Cap5 antiphage defense system employing 3′, 2′-cGAMP as the second messenger. *Nature Communications* (2021).
30. Jiang, K., Faltova, L., Hua, S., Capitani, G., Prota, A. E., Landgraf, C., Volkmer, R., Kammerer, R. A., Steinmetz, M. O. & Akhmanova, A. Structural Basis of Formation of the Microtubule Minus-End-Regulating CAMSAP-Katanin Complex. *Structure* 26, 375-382.e4 (2018).
31. Saha, C. K., Sanches Pires, R., Brolin, H., Delannoy, M. & Atkinson, G. C. FlaGs and webFlaGs: discovering novel biology through the analysis of gene neighbourhood conservation. *Bioinformatics* 37, 1312-1314 (2021).

ACCELERATED ARTICLE PREVIEW

32. Zimmermann, L., Stephens, A., Nam, S.-Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A. N. & Alva, V. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *Journal of molecular biology* 430, 2237-2243 (2018).
33. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P. & Hassabis, D. Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583-589 (2021).
34. Holm, L. & Rosenström, P. Dali server: conservation mapping in 3D. 38, W545-9 (2010).
35. Simanshu, D. K., Yamaguchi, Y., Park, J.-H., Inouye, M. & Patel, D. J. Structural basis of mRNA recognition and cleavage by toxin MazF and its regulation by antitoxin MazE in Bacillus subtilis. *Molecular cell* 52, 447-458 (2013).
36. Hogrel, G., Guild, A., Graham, S., Rickman, H., Grushow, S., Bertrand, Q., Spagnolo, L. & White, M. F. Cyclic nucleotide-induced superhelical structure activates a bacterial TIR immune effector. *bioRxiv* (2022).
37. Hogrel, G., Guild, A., Graham, S., Rickman, H., Grüschow, S., Bertrand, Q., Spagnolo, L. & White, M. F. Cyclic nucleotide-induced helical structure activates a TIR immune effector. *Nature* 1-5 (2022).
38. Paget, M. S. Bacterial sigma factors and anti-sigma factors: structure, function and distribution. *Biomolecules* 5, 1245-1265 (2015).
39. Sineva, E., Savkina, M. & Ades, S. E. Themes and variations in gene regulation by extracytoplasmic function (ECF) sigma factors. *Current opinion in microbiology* 36, 128-137 (2017).
40. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* 372, 774-797 (2007).
41. Schuster, C. F. & Bertram, R. Toxin–antitoxin systems are ubiquitous and versatile modulators of prokaryotic cell fate. *FEMS microbiology letters* 340, 73-85 (2013).
42. Walsh, P. N. & Ahmad, S. S. Proteases in blood clotting. *Essays in biochemistry* 38, 95-112 (2002).
43. Brown, M. S., Ye, J., Rawson, R. B. & Goldstein, J. L. Regulated intramembrane proteolysis: a control mechanism conserved from bacteria to humans. *Cell* 100, 391-398 (2000).
44. Fei, X., Bell, T. A., Barkow, S. R., Baker, T. A. & Sauer, R. T. Structural basis of ClpXP recognition and unfolding of ssrA-tagged substrates. *Elife* 9, e61496 (2020).
45. Hu, C., van Beljouw, S. P. B., Nam, K. H., Schuler, G., Ding, F., Cui, Y., Rodríguez-Molina, A., Haagsma, A. C., Valk, M. & Pabst, M. Craspase is a CRISPR RNA-guided, RNA-activated protease. *Science* eadd5064 (2022).
46. van Beljouw, S. P. B., Haagsma, A. C., Rodríguez-Molina, A., van den Berg, D. F., Vink, J. N. A. & Brouns, S. J. J. The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. *Science* 373, 1349-1353 (2021).
47. Kato, K., Okazaki, S., Schmitt-Ulms, C., Jiang, K., Zhou, W., Ishikawa, J., Isayama, Y., Adachi, S., Nishizawa, T. & Makarova, K. S. RNA-triggered protein cleavage and cell growth arrest by the type III-E CRISPR nuclease-protease. *Science* eadd7347 (2022).
48. Strecker, J., Demircioglu, F. E., Li, D., Faure, G., Wilkinson, M. E., Gootenberg, J. S., Abudayyeh, O. O., Nishimasu, H., Macrae, R. K. & Zhang, F. RNA-activated protein cleavage with a CRISPR-associated endopeptidase. *Science* eadd7450 (2022).


ACCELERATED ARTICLE PREVIEW

## Figure legends

**Fig. 1 | Structure of apo and $cA_4$ bound CalpL. a**, Overall structure of CalpL in the apo state. The structure is shown as a cartoon model and the individual domains are labelled and color-coded. The N- and C-termini, as well as the protease active site residues, are marked by spheres. The positions of key structural elements are indicated. **b**, Single cycle kinetics SPR measurements of different cOAs binding to immobilized CalpL. The experiment was performed multiple times (n=3 technical replicates) for $cA_4$ and once for the other cOAs. **c**, The CalpL/$cA_4$ complex structure. CalpL is shown as a surface model and the electrostatic potential is mapped onto the structure (blue - positive, red – negative). The bound $cA_4$ molecule is shown as spheres. Turning the structure by 180° reveals a positively charged patch opposite of the SAVED domain.

**Fig. 2 | CalpL is activated by $cA_4$ and cleaves CalpT. a**, The WebFLAGs server [31] was used to investigate the genomic neighborhood of CalpL (green). The primary structure of CalpT (red) is shown on top. Regions with homologies found by HHPRED [32] are marked. **b**, A structural prediction (AlphaFold2, [33]) of CalpT. The protein is shown as cartoon and colored according to the prediction confidence (pLDDT [49], predicted local distance difference test). **c**, SDS-PAGE analysis of CalpL induced cleavage of CalpT. The experiment was repeated multiple times (n =2 biological replicates and n>>3 technical replicates) **d**, Sequence alignment [50] showing that the identified P1 site at A195 is conserved among CalpT homologs. For gel source data, see Supplementary Figure 1.

**Fig. 3 | CalpL and -T form a stable complex and $cA_4$-induced oligomerization of CalpL. a**, SEC-MALS traces (solid lines: $UV_{280}$, dashed lines: $MW_{MALS}$) of proteolysis reactions with different combinations of CalpL wt, CalpT wt, and cOA. The SEC-MALS experiment was performed multiple times with slight variations of buffer and concentrations (n = 3 technical replicates). The schematic indicates the molecular species behind the individual peaks. Inset: Crystal structure of the CalpL/$T_{10}$ complex. The distance of 36 Å between the P-1 position (S196) and the protease active site is indicated. **b**, Top: Concentration-normalized small angle X-ray scattering curves recorded at four different concentrations of CalpL. For each experiment, thirty sample intensity frames and sixty buffer intensity frames were collected and averaged. For each data set and angular point the errors were computed following the Poisson statistics. The data points represent the average intensity difference (sample-buffer) and the error bars represent the standard deviation. The experiment was performed once for each concentration. Bottom: Molecular weights from forward scattering I(0) calculated from the SAXS curves plotted vs the concentration. The apparent molecular weight of the protein in the presence of $cA_4$ increases with the concentration. **c**, Protease assays with preformed CalpL/T complexes as indicated in the figure. The experiment was performed three times (n = 3 technical replicates). **d**, Protease assays with CalpL mutants in the positive patch on the backside of the SAVED domain (Fig. 1). The experiment was performed twice (n = 2 technical replicates). For gel source data, see Supplementary Figure 1.

**Fig. 4 | CalpT links $cA_4$ detection to the transcription machinery of the cell. a**, Structural information about the CalpL/T/S complex as obtained by crystallography, SAXS and AlphaFold2. **b-d**, CalpT and S form a complex that is stable during gelfiltration and disintegrates into CalpL/$T_{10}$ and $CalpT_{23}$/S. The experiments were performed twice (n=2 technical replicates). **e**, His-CalpS and CalpT can be copurified from *E. coli*. Overexpression of CalpS alone leads to copurification of the α- and β-subunits of the *E. coli* RNAP. The complex formations were replicated multiple times in two independent laboratories (n>3 biological replicates). For gel source data, see Supplementary Figure 1.

ACCELERATED ARTICLE PREVIEW

**Fig. 5 | Model for CalpL/T/S mediated antiviral defense. a**, The *calpL*, *calpT* and *calpS* genes are located in close proximity to the type III-B CRISPR genes of *Sulfurihydrogenibium sp.* YO3AOP1 (modified from [6]). **b**, Once activated, the Cas10 subunit of the RNP synthesizes $cA_4$ from ATP. The second messenger binds to preformed CalpL/T/S complexes. Oligomerization leads to proteolytic cleavage of CalpT, releasing the $CalpT_{23}$/S fragment. $CalpT_{23}$ is likely degraded by proteases, allowing CalpS to bind to the RNA polymerase.

ACCELERATED ARTICLE PREVIEW

## Methods

### Expression and purification of CalpL

The codon-optimized gene for CRISPR–Lon was cloned into a pET11a vector with an N-terminal 10xHis-TEV tag. Site-directed mutagenesis was performed according to a protocol by Liu et al.[51]. All CalpL constructs were expressed in lysogeny broth (LB) medium. *E. coli* BL21(DE3) cells were grown at 37 °C until an $OD_{600}$ of 0.6-0.8 was reached. Then, protein expression was started by induction with 0.4 mM IPTG, and the cell suspension was incubated at 30 °C for 4.5 h with shaking. Cells were harvested by centrifugation at 4,000*rcf for 25 min. at 20 °C and resuspended in lysis buffer (20 mM Tris, 50 mM NaCl, pH 8.0). The cells were lysed with a sonicator and cell debris was removed by centrifugation at 48,000*rcf for 45 min. at 4 °C. For protein purification, $Ni^{2+}$-affinity chromatography (20 mM Tris, 50 mM NaCl, pH 8.0; 500 mM imidazole was included for elution) was followed by size-exclusion chromatography (20 mM Tris, 50 mM NaCl, pH 8.0) using a Superdex 200 16/600 column. After that, the His-tag was cleaved off by overnight incubation at 4 °C with a 1:50 molar ratio of protein to TEV protease (20 mM Tris, 50 mM NaCl, pH 8.0). A second $Ni^{2+}$-affinity chromatography was used to remove the TEV protease and uncleaved protein. The purity of the protein was checked by SDS-PAGE after each purification step. After successful purification, the proteins were concentrated, flash-frozen in liquid nitrogen, and stored at -80 °C in 20 mM Tris, 50 mM NaCl, pH 8.0. The selenomethionine derivative of CalpL was prepared using *E. coli* B834 cells and the "SelenoMethionine Medium Complete" kit from Molecular Dimensions according to the instructions. Protein expression and purification were done in the same way as for the native protein.

### Expression and purification of CalpT

The codon-optimized synthetic gene (BioCat) for CalpT (UNIPROT-ID: B2V8L8), including an N-terminal 10x His-TEV tag was cloned into a pET11a vector. Protein expression was done using the same expression strain and the same conditions as for CalpL. Cells were harvested by centrifugation at 4,000*rcf for 25 min. at 20 °C and resuspended in lysis buffer (25 mM Tris, 500 mM NaCl, 10% glycerol, 1 mM DTT, pH 8.0). The cells were lysed with a sonicator and cell debris was removed by centrifugation at 48,000*rcf for 45 min. at 20 °C. For protein purification, $Ni^{2+}$-affinity chromatography (25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0; 1 M imidazole was included for elution) was followed by size-exclusion chromatography (25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0) using a Superdex 75 16/600 column. After that, the His-tag was cleaved off by overnight incubation at

ACCELERATED ARTICLE PREVIEW

4 °C with a 20:1 ratio (m/m) of protein to TEV protease (25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0). A second $Ni^{2+}$-affinity chromatography was used to separate the TEV protease and uncleaved protein. The purity of the protein was checked by SDS-PAGE after each purification step. After successful purification, the proteins were concentrated, flash-frozen in liquid nitrogen, and stored at -80 °C in 25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0.

**Expression and purification of CalpS**

The codon-optimized gene of CalpS was purchased from Integrated DNA Technologies (IDT, Coralville, Iowa, USA) as a G-Bock with flanking restriction sites for cloning. *SF* was cloned into NcoI and BamHI restriction sites of vector pEV5HisTEV[52], allowing expressed proteins with an N-terminal 8x His-TEV tag. For expression, *E. coli* C43(DE3) cells with sequencing-verified construct were incubated at 37 °C with shaking at 180 rpm until $OD_{600}$ of the cells was between 0.6 and 0.8. Then, the cell culture was grown at 16 °C overnight after inducing with 0.2 mM IPTG. The cell pellet was collected by centrifugation at 4000 rpm (Beckman Coulter Avanti JXN-26; JLA8.1 rotor) at 4 °C for 15 min. For purification, cell pellet was resuspended into buffer A (50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 20 mM imidazole, and 10% glycerol) and lysed by sonication. The cleared cell lysate was loaded onto a 5 mL HisTrap FF column (GE Healthcare) equilibrated with buffer A. The His-tagged SF was eluted in a linear gradient with buffer B (50 mM Tris-HCl pH 7.5, 0.5 M NaCl, 0.5 M imidazole, and 10% glycerol). The his-tag was then removed by incubating with TVE protease at room temperature overnight before recovering TEV-cleaved SF through a HisTrap column again. Size-exclusion chromatography was finally used to purify the SF in SEC buffer (20 mM Tris–HCl, 0.25 M NaCl, 1 mM DTT, 10% glycerol, pH 7.5). The purity of SF was evaluated on the SDS-PAGE at each purification step. Concentrated SF was flash-frozen in liquid nitrogen and stored at -70 °C.

**Co-expression and co-purification of sigma factor (SF) and CalpT**

For expression of His-tagged CalpS with CalpT, the fragment of CalpT flanking NdeI and XhoI sites was cloned into MCS-2 of vector pCDFDuet™-1 (Novagen, Merck Millipore). The constructs pEV5HisTEV-SF and pCDFDuet-CalpT were co-transformed into *E. coli* C43(DE3) cells. The cell was induced by 0.2 mM IPTG after reaching OD600 of 0.6-0.8 and grown overnight at 16 °C with shaking at 180 rpm.

For expression of His-tagged CalpT with CalpS, a G-Block of SF was constructed into MCS-1 (NcoI and BamHI) of vector pCDFDuet™-1. *E. coli* C43(DE3) cells were transformed with

ACCELERATED ARTICLE PREVIEW

constructs pET11a-CalpT and pCDFDuet-SF and grown at 37°C with shaking at 180 rpm. Cell was induced by 0.2 mM IPTG once OD600 of 0.6-0.8 and cultivated at 16 °C overnight, then purified as described above.

**Protease assay**

For protease activity assays CalpL and CalpT were used at a final concentration of c = 4.64 µM each. The different cOAs were used at a final concentration of c = 5.11 µM. The protein solutions were prepared in 20 mM Tris, 50 mM NaCl, pH 8.0 and incubated for 1 hr at 60 °C. Subsequently, the cOA was added and the mixture was incubated for another 1 hr at 60 °C. For SDS-PAGE 3 µl of 4x SDS-loading buffer was added to 9 µl of the sample, the mixture was heated for 5 min at 94 °C and 10 µl were loaded to a 15% polyacrylamide gel, which was run at 250 V for 40 min.

**Size exclusion chromatography analysis (SEC)**

To determine the interaction of the complex of CalpS and CalpT with CalpL, the SEC runs were carried out on a Superose6 increase 10/300 chromatography column (GE Healthcare) equilibrated with SEC buffer (20 mM Tris, 0.25 M NaCl, 1 mM DTT, 10% glycerol, pH 8.0). The injected volume of tested sample solution was 200 ul at flow rate of 0.5 ml $min^{-1}$. The final concentrations were set to c(CalpL) = 63.3 µmol*$l^{-1}$, c(CalpT/S) = 115.8 µmol*$l^{-1}$, and c($cA_4$) = 60 µmol*$l^{-1}$ diluted by using SEC buffer. All samples were incubated at 60 °C for 60 min before cooling down to room temperature and loading onto column.

**Pull-down assay**

The magnetic nickel beads-based immobilized metal affinity chromatography (IMAC) was performed to detect releasing of CalpS from CalpL/T/S complex. The complex of His-tagged CalpS and CalpT was incubated with CalpL in binding buffer (20 mM Tris-HCl, pH 7.5, 60 mM NaCl, 0.01% Tween™-20) at 60 °C for 1 hr in presence or absence of $cA_4$. After cooling down to room temperature, the sample solution was mixed with pre-equilibrated beads (Magne, His Ni particle, Promega) with binding buffer on a roller for 20 min at 4 °C. The beads were washed tree times with 300 ul wash buffer (20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 10 mM imidazole, 0.01% Tween™-20) before eluted twice using 25 ul elution buffer (20 mM Tris-HCl, pH 8.0, 120 mM NaCl, 300 mM imidazole, 0.01% Tween™-20). The samples from first elution and 20% input were analyzed on the SDS-PAGE. The final concentrations were set to

ACCELERATED ARTICLE PREVIEW

c(CalpT/S) = 0.208mg/ml, c(CalpL) = 0.127 mg/ml, and c($cA_4$) = 2.5 µmol*$l^{-1}$ diluted by binding buffer.

**Analytical gel filtration and SEC-MALS analysis**

To investigate the complex formation of CalpS with CalpT and CalpL, analytical gel filtration was carried out on a SD 200 increase 10/300 column. When purifying CalpS, the complex consisting of CalpS and DNA-directed RNA polymerase subunits alpha and beta, eluted in a defined 50 mAU peak on a SD 200 16/600 gel filtration column. This peak was pooled and concentrated to V ≈ 1.5 ml. Thereafter, 440 µl of the complex were incubated with 60 µl of CalpT (350 µM), 60 µl of CalpT-CalpL S152A complex (195 µM) and 60 µl of CalpS buffer (20 mM Tris–HCl, 0.25 M NaCl, 1 mM DTT, 10% glycerol, pH 7.5), respectively. After centrifugation at 15000 rcf and 10 °C for 10 min, each sample was loaded onto a SD 200 increase 10/300 column for size-exclusion chromatography in CalpS buffer.

For determination of interactions between CalpL and CalpT, SEC-MALS runs were performed at room temperature on an Agilent 1260 Infinity II Prime Bio LC System coupled with a Wyatt miniDAWN® MALS detector, a Optilab rEX refractive index detector and a Superose6 increase 10/300 chromatography column (GE Healthcare) equilibrated with 25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0. Data acquisition and evaluation were carried out using ASTRA 8 software (Wyatt Technologies). The flow rate was set to 0.5 ml $min^{-1}$ and an injection volume of 50 µl was used for the experiments. Final concentrations were set to c(CalpL) = 51 µmol $l^{-1}$, c(CalpT) = 51 µmol $l^{-1}$, and c(cA4) = 60 µmol $l^{-1}$ by dilution with 25 mM Tris, 500 mM NaCl, 1 mM DTT, 10% glycerol, pH 8.0. The proteins were incubated for 40 min at 60 °C, $cA_4$ was added followed by an additional 20 min incubation at 60 °C. The samples were centrifuged at 15.000*g for 10 min. before injection.

**Mass spectrometry**

The gel bands were excised and cut into 1 $mm^3$ cubes. The samples were destained with 2x rinses each of ethanol, acetonitrile and 25mM ammonium bicarbonate, then subjected to reduction with 10 mM dithiorethritol, followed by alkylation with 20 mM iodoacetamide. The gel pieces with shrunk with acetonitrile and then soaked in 25 mM AmBic with 2 ng/µl trypsin and left to digest overnight at 37 °C. The peptides were soaked from the gel with 1% formic acid and concentrated to 20 µl in a speedvac. Between 1-7 µl of the sample, dependent on original gel coomassie staining, was loaded onto a Eksigent 2D ultra nano HPLC with Sceix 5600+ mass spectrometer. The Thermoscientific Acclaim Pepmap 100 trap (20 mm x 75 µm)

ACCELERATED ARTICLE PREVIEW

and column (150 mm x 75 µm) were in trap elute configuration with a flow of 5 µl/min and 300 nl/min respectively. The peptides were loaded onto the trap and washed for 5 minutes at 100% loading buffer (100% water, 0.05% TFA) before the trap was switched in line with the column and the peptides eluted with a linear gradient over 20 minutes of 98% A to 98% B where A is 100% water with 0.1% formic acid and B is 80% acetonitrile, 20% water, 0.1% formic acid. The eluent was sprayed directly into the nanosource of the mass spectrometer. MS data was collected from 400-1250 m/z in positive ionisation for 150 msec. Data dependant acquisition mode was utilized to collect MSMS data from 100-2000 m/z on the 20 strongest peptides with 2-5+ charge states. The peak list was extracted from the .wiff file using MSconvert and the .mgf file searched against an inhouse database of 7000 protein sequences to which the sequences of the proteins of interest were added. The following settings were used in the mascot search, trypsin, and semi trypsin as digest enzymes, fixed modification of carbamidomethyl (c) and variable modifcation of oxidation (M). MS tolerance was set at 20 ppm and MSMS at 0.1 Da.

**Surface plasmon resonance of cOA and CalpT binding to CalpL**

All surface plasmon resonance experiments were run on a Biacore™ 8K instrument (GE healthcare life sciences), using a streptavidin-functionalized sensor chip (Serie S Sensor Chip SA, GE healthcare life sciences). Data was recorded at a rate of 10 Hz and 25°C flow cell temperature. The running buffer contained 25 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% Glycerol, 0.05% TWEEN20. After three initial injections of 1 M NaCl in 50 mM NaOH (10 µL/min, 60 s), the biotinylated CalpL construct, R493C-biotin, was immobilized on the chip (86 nM, 5 µL/min, 180 s). Binding of CoA and CalpT was measured as single cycle kinetics. For the cOAs, a series of seven different concentrations (0.086, 0.26, 0.78, 2.33, 7, 21, 63 nM) were injected at a flow rate of 30 µL/min (contact time: 120s, dissociation time: 600 s). For CalpT and NIS038, a series of seven different concentrations (0.0625, 0.25, 1, 4, 16, 64, 256 nM) were injected applying the same parameters as above. The recorded data were double referenced by reference flow cell and blank cycle subtraction and data was analysed and fitted using the Biacore Insight Evaluation Software.

**X-ray crystallography**

Pure CalpL protein was concentrated to 20 mg/ml and crystallized at 20 °C using a Gryphon pipetting robot (Art Robbins) and commercial crystallization screens (Molecular Dimensions) using sitting drop plates. Hexagonal crystals appeared after one day in condition D7 of the

ACCELERATED ARTICLE PREVIEW

JCSG+ screen. Several rounds of optimization in sitting- and hanging drop plates were performed to achieve well-diffracting crystals. The final crystallization condition was 0.1 M Tris-Cl pH 8.0, 38.8% PEG400, 0.29 M $Li_2SO_4$. The SeMet derivative (see above) was crystallized under similar conditions and yielded identical crystals. The crystals were harvested without further cryo-protection and a diffraction dataset was recorded at beamline P13 ($\lambda = 0.9795$) operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany)[53]. The diffraction data were automatically processed with XDS[54]. The structure was solved using phenix.autosol and refined with phenix.refine[55]. Further model building was performed in Coot[56] and figures were prepared with PyMOL (www.pymol.org). The geometry of the model was checked with MolProbity[57]. The molprobity score was 1.43, the clashscore 4.53 and the Ramachandran statistics (outliers/favored) were 0.0/97.0%.

The same crystallization condition was used to obtain CalpL crystals for soaking with $cA_4$. After harvesting the crystals, they were incubated for approximately 3 minutes in a solution of mother liquor supplemented with 5% PEG 400 and 5 mM $cA_4$. Diffraction experiments were done at beamline P13 ($\lambda = 0.\ 0.97626$) at the PETRA III storage ring (DESY, Hamburg, Germany)[53]. A full dataset was recorded and automatically processed with XDS[54]. Molecular replacement was run with PHASER[58] using CalpL as search model. Further refinement of the structure was done with phenix.refine[55]. Using Coot[56], cyclic tetraadenylate could be fitted perfectly into a defined difference electron density inside the $cA_4$ binding pocket. The molprobity score was 1.32, the clashscore 2.85 and the Ramachandran statistics (outliers/favored) were 0.4/96.6%.

To obtain the crystal structure of the CalpL/$T_{10}$ complex, CalpL and CalpT protein solutions were mixed at 1:1 molar ratio (155 µM each) and incubated at 50 °C for 40 min. After incubation, $cA_4$ was added to a final concentration of 175 µM followed by 20 min of incubation at 50 °C. A total volume of 400 µl was loaded to a SD 200 increase 10/300 column and size-exclusion chromatography was conducted using CalpL buffer. The complex eluted in one single peak which was pooled and concentrated to approximately 30 mg/ml. Sitting drop crystallization plates were set up as described above. Crystals were obtained after several days in condition E2 of the JCSG+ screen (2 M ammonium sulfate, 0.1 M ammonium cacodylate, 0.2 M NaCl at pH 6.5). The crystals were harvested with 35% glycerol for cryo-protection and a diffraction dataset was recorded at beamline P13 ($\lambda = 0.\ 0.9762$) operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany)[53]. Automatic data processing was achieved using XDS[54]. CalpL was used as search model for molecular replacement with PHASER [58]. Structure refinement and model building was done with phenix.refine[55] and

ACCELERATED ARTICLE PREVIEW

Coot[56], respectively. The molprobity score was 2.36, the clashscore 11.93 and the Ramachandran statistics (outliers/favored) were 0.9/92.3%.

Geometric parameters of all described structures were checked with MolProbity[57]. All figures were prepared with PyMOL (www.pymol.org).

**Small-angle X-ray scattering**

Small-angle X-ray scattering (SAXS) experiments were conducted at the P12[59] beamline of the Petra III synchrotron (Hamburg, Germany), in two separate experimental sessions. The parameters for the data collections are reported in Extended Data Table 2. The scattering intensity, I(s) was collected as a function of the momentum transfer defined as $s=(4\pi\sin\theta)/\lambda$, where $2\theta$ is the scattering angle and $\lambda$ the X-ray wavelength employed.

The buffer employed for all SAXS experiments, also employed for SEC and for background subtraction, contained 20 mM Tris pH=8.0, 50 mM NaCl. CalpL, CalpL/$T_{10}$ and CalpL monomer SAXS curves were collected by SEC-SAXS at room temperature[60], employing a SD200 5/150 increase SEC column (GE Healthcare) online to the SAXS flow capillary. For each SEC-SAXS run, 15 min elutions at 0.3 mL/min flow rate were performed, collecting 900x 1 s exposures on the eluate.

Concentration series of CalpL in the range 1-5 mg/mL with and without a 1.2-fold molar excess of $cA_4$ were collected in batch mode after centrifugation 30 min at 30000 xg (5 °C), and 30 exposures of 0.1 s were collected while flowing 35 µL of solution through the 1 mm quartz capillary. The apparent molecular weights for these measurements are obtained from the forward scattering, I(0), using the I(0) of a SAXS curve from bovine serum albumin at 1.9 mg/mL in a HEPES buffer as secondary standard.

The primary data reduction was performed with the program SASFLOW[61], including automatic selection of the exposures to monitor for radiation damage prior to data averaging. The reduced data inspected and processed to obtain the overall protein parameters using PRIMUS[62] and the programs of the ATSAS suite[63]. Comparison of experimental SAXS curves with the crystal structures was performed with CRYSOL[64] *Ab initio* modeling of low-resolution protein structures was performed with DAMMIF[65], and MONSA[66] for multi-phase modeling, repeating 10 modeling runs. The resulting models were compared and averaged using DAMAVER[67], providing a normalized spatial discrepancy (NSD) value[68] that reflects the stability of the structural reconstructions (stable reconstructions have NSD<1). SAXS curves of CalpL at 2, 3, 5 mg/mL in the presence of cA4 were globally fitted as a mixture of monomer and dimer, using the program SASREFMX[69]. The mixture results in an overall scattering intensity which is the

ACCELERATED ARTICLE PREVIEW

linear combination of the scattering intensity of monomer and dimer species at varying volume fractions along the concentration series.

**Pulsed EPR experiments**

For site-specific spin labelling, a single cysteine mutant E119C of CalpT of was expressed and purified as described for wild-type CalpT. After purification, 250 µl of a 315 µM CalpT E119C solution was bound to $Ni^{2+}$-NTA beads. These were washed with 10 ml reducing buffer (25 mM Tris, pH = 8, 250 mM NaCl, 1 mM TCEP, 10% glycerol) and 20 ml wash buffer (25 mM Tris, pH = 8, 250 mM NaCl, 10% glycerol). Thereafter, the protein was eluted in buffer containing MTSSL (25 mM Tris, pH = 8, 250 mM NaCl, 1 M imidazole, 10% glycerol, 0.6 mM MTSSL). A PD10 desalting column was used to remove imidazole and free spin label, as well as for buffer exchange of CalpT E119R1, wild-type CalpL and CalpL S152A. All final buffers contained $D_2O$ instead of $H_2O$ and no reducing agent. An activity assay was done as described using the prepared samples with a final protein concentration of 55 µM and deuterated buffer (25 mM Tris, pH = 8, 250 mM NaCl, 5% glycerol). The success was checked on SDS-PAGE. All samples were flash frozen to $N_{2\,(l)}$ and sent on dry ice for the measurement. The labelling efficiency was determined to be 103% by cw-EPR spectroscopy (average of two measurements).

For pulse EPR measurements, samples of spin-labeled CalpL/T E119R1 in presence or absence of 1 molar equivalent $cA_4$ were mixed with 45% (v/v) deuterated ethylene glycol to yield 27.5 µM CalpL/T E119R1 in 65 µl final volume. Samples were transferred to 3 mm EPR quartz tubes, flash-frozen and stored in liquid nitrogen until use.

Pulsed electron-electron double resonance (PELDOR)[70] distance measurements were obtained at Q-band frequency (34 GHz) on a Bruker ELEXSYS E580 spectrometer with 3 mm cylindrical resonator (ER 5106QT2-2w, TE012 mode) using a pulse travelling wave tube (TWT) amplifier (Applied Systems Engineering) with nominal output of 150 W and an arbitrary waveform generator for rectangular pulses.

PELDOR experiments were performed with the 4-pulse DEER[71,72] pulse sequence ($\pi/2(\nu_A)$ – $\tau_1$ – $\pi(\nu_A)$ – ($\tau_1$ + t) – $\pi(\nu_B)$ – ($\tau_2$ - t) – $\pi(\nu_A)$ – $\tau_2$ – echo) at 50 K, with a frequency offset (pump – detection frequency) of +80 MHz (~3 mT). Shot repetition time (SRT) was set to 2.5 ms; $\tau_1$ was set to 380 ns, and $\tau_2$ was set to 5000 ns. Pulse lengths were 16 and 32 ns for $\pi/2$ and $\pi$ detection, and 12 ns for the inversion $\pi$ pump pulse. Unwanted echoes were suppressed with a 16-step phase cycle and nuclear modulation was averaged by adding 16 traces with $\tau_1$

ACCELERATED ARTICLE PREVIEW

incremented by 8 ns. The pump pulse was placed on the resonance frequency of the resonator and applied to the maximum of the nitroxide field-swept spectrum.

PELDOR data were analyzed using the ComparativeDeerAnalyzer version 2.0[73,74] within DeerAnalysis2022[75]; shown are the respective consensus fits and distance distributions.

**Ribonuclease assay**

Ribonuclease activity of cleaved CalpT (23-kD fragment) was assayed by incubating full-length CalpT with CalpL and five different fluorescent-labelled RNA substrates, which were synthesised with the fluorescent dye (6-FAM) attached at 5' end or at 3' end (purchased from Integrated DNA Technologies (IDT), Extended Data Fig. 7c). The mixture of CalpL (5.5 µM) and CalpT (5.5 µM) was incubated at 60 °C in 20 mM Tris-HCl, pH 8.0, 50 mM NaCl and 1 mM EDTA for 15 min, $cA_4$ (10 µM) was then added and the mixture was incubated for another 15 min at 60 °C, followed by adding one of the above RNA substrates into the mixture, incubating for an additional 30 min at 60°C. Finally, 6 µl of the sample was analyzed on SDS-PAGE (NuPAGE Bis-Tris Gel, Thermo Fisher Scientific) by heating at 95 °C for 5 min with 2 µL of SDS-PAGE loading dye (Thermo Fisher Scientific; NuPAGE Sample Reducing Agent and LDS Sample Buffer). The remaining 14 µl of the sample were loaded to 20% acrylamide, 7 M urea, 1×TBE denaturing gel, which was run at 30W, 45 °C for 2 hr. The gel was finally imaged by Typhoon FLA 7000 imager (GE Healthcare) at a wavelength of 532 nm (pmt 600~700).

**Ribonuclease target motif profiling (RNase-Seq)**

To investigate MazF activity, 20 U mRNA Interferase™ -MazF (TaKaRa, Cat# 2415A) were incubated with 400 ng of a single-stranded RNA library containing 10 random bases in 1X MazF Buffer (200 mM sodium phosphate, pH 7.5, and 0.05% Tween-20) at 37 °C. After 10 min or 2 hr, reactions were stopped by placing the samples on ice. Ribonuclease activity of CalpT (23 kDa fragment) was investigated by incubating a solution containing 4.64 µM CalpL and 5.57 µM CalpT in 20 mM Tris pH 8.0, 50 mM NaCl at 37 °C. After 30 min, 400 ng of a ssRNA library was added together with $cOA_4$ to a final concentration of 5.57 µM. The reaction was incubated for 10 min or 2 hr at 37 °C. To prepare next generation sequencing libraries, RT-PCR was performed on 1 µl of each sample in a 10 µl reaction containing 1x KAPA HiFi HotStart ReadyMix (pre-heated for 5 min at 98 °C; Roche, Cat# KK2602), 15 U WarmStart RTx Reverse Transcriptase (NEB, Cat# M0380) and 0.5 µM staggered MiSeq gRNA primer mix using the following temperature conditions: 15 min 65 °C, 3 min 72 °C, 30 sec 98 °C, 20

ACCELERATED ARTICLE PREVIEW

cycles of: Denaturation (10 sec at 98 °C), annealing (20 sec at 65 °C) and extension (1 min at 72 °C); Final extension: 5 min at 72 °C. Barcodes and Illumina-compatible constant handles were added using a secondary NEBNext (NEB) PCR. Samples were pooled and column-purified using QIAprep Spin columns. The final library was quantified using a NanoDrop photospectrometer and sequenced on an Illumina MiSeq using the v2 chemistry.
The primers used for ribonuclease target motif profiling are listed in Extended Data Fig. 7e.

**Structural predictions with AlphaFold2**

The source code of the AlphaFold2 algorithm was downloaded from https://github.com/AlphaFold and installed as described https://github.com/AlphaFold. The algorithm was run locally using the CASP14 preset or via ColabFold[76].

**Statistics and reproducibility**

Information concerning statistics and reproducibility for the experiments shown in this study are given in the figure legends of the corresponding experiments. The key findings of this study (enzymatic activities and formation of macromolecular complexes) have been reproduced in two laboratories (G.H. at the Institute of Structural Biology of the University of Bonn, Germany and M.F.W. at the Biomedical Sciences Research Complex of the University of St Andrews, Scotland, UK).

ACCELERATED ARTICLE PREVIEW

## Additional references


49. Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., Bridgland, A., Cowie, A., Meyer, C. & Laydon, A. Highly accurate protein structure prediction for the human proteome. *Nature* 596, 590-596 (2021).
50. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic acids research* 42, W320-W324 (2014).
51. Liu, H. & Naismith, J. H. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology* 8, 91 (2008).
52. Rouillon, C., Athukoralage, J. S., Graham, S., Grüschow, S. & White, M. F. Investigation of the cyclic oligoadenylate signaling pathway of type III CRISPR systems. *Methods Enzymol* 616, 191-218 (2019).
53. Cianci, M., Bourenkov, G., Pompidor, G., Karpics, I., Kallio, J., Bento, I., Roessle, M., Cipriani, F., Fiedler, S. & Schneider, T. R. P13, the EMBL macromolecular crystallography beamline at the low-emittance PETRA III ring for high-and low-energy phasing with variable beam focusing. *Journal of synchrotron radiation* 24, 323-332 (2017).
54. Kabsch, W. Automatic-Indexing of Rotation Diffraction Patterns. 21, 67-71 (1988).
55. Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S. & McCoy, A. J. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology* 75, 861-877 (2019).
56. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* 60, 2126-2132 (2004).
57. Williams, C. J., Headd, J. J., Moriarty, N. W., Prisant, M. G., Videau, L. L., Deis, L. N., Verma, V., Keedy, D. A., Hintze, B. J. & Chen, V. B. MolProbity: More and better reference data for improved all-atom structure validation. *Protein Science* 27, 293-315 (2018).
58. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. Phasercrystallographic software. *Journal of Applied Crystallography* 40, 658-674 (2007).
59. Blanchet, C. E., Spilotros, A., Schwemmer, F., Graewert, M. A., Kikhney, A., Jeffries, C. M., Franke, D., Mark, D., Zengerle, R., Cipriani, F., Fiedler, S., Roessle, M., Svergun, D. I. & IUCr. Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *Journal of Applied Crystallography* 48, 431-443 (2015).
60. Graewert, M. A., Da Vela, S., Gräwert, T. W., Molodenskiy, D. S., Blanchet, C. E., Svergun, D. I. & Jeffries, C. M. Adding size exclusion chromatography (SEC) and light scattering (LS) devices to obtain high-quality small angle X-ray scattering (SAXS) data. *Crystals* 10, 975 (2020).
61. Franke, D., Kikhney, A. G. & Svergun, D. I. Automated acquisition and analysis of small angle X-ray scattering data. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 689, 52-59 (2012).
62. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. 36, 1277-1282 (2003).
63. Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A. & Borges,


ACCELERATED ARTICLE PREVIEW

C. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *Journal of Applied Crystallography* 54, 343-355 (2021).
64. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSOL– a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *Journal of Applied Crystallography* 28, 768-773 (1995).
65. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. 42, 342-346 (2009).
66. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophysj* 76, 2879-2886 (1999).
67. Volkov, V. V. & Svergun, D. I. Uniqueness of ab initio shape determination in small-angle scattering. 36, 860-864 (2003).
68. Kozin, M. B. & Svergun, D. I. Automated matching of high- and low-resolution structural models. 34, 33-41 (2001).
69. Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. New developments in the ATSAS program package for small-angle scattering data analysis. 45, 342-350 (2012).
70. Milov, A., Salikohov, K. & Shirov, M. Application of Endor in Electron-Spin Echo for Paramagnetic Center Space Distribution in Solids. *Fizika Tverdogo Tela* 23, 975-982 (1981).
71. Pannier, M., Veit, S., Godt, A., Jeschke, G. & Spiess, H. W. Dead-time free measurement of dipole-dipole interactions between electron spins. *Journal of Magnetic Resonance (San Diego, Calif.: 1997)* 142, 331-340 (2000).
72. Larsen, R. G. & Singel, D. J. Double electron–electron resonance spin–echo modulation: Spectroscopic measurement of electron spin pair separations in orientationally disordered solids. *The Journal of chemical physics* 98, 5134-5146 (1993).
73. Worswick, S. G., Spencer, J. A., Jeschke, G. & Kuprov, I. Deep neural network processing of DEER data. *Science advances* 4, eaat5218 (2018).
74. Fábregas Ibáñez, L., Jeschke, G. & Stoll, S. DeerLab: a comprehensive software package for analyzing dipolar electron paramagnetic resonance spectroscopy data. *Magnetic Resonance* 1, 209-224 (2020).
75. Jeschke, G., Chechik, V., Ionita, P. & Godt, A. DeerAnalysis2006—a comprehensive software package for analyzing pulsed ELDOR data. *Applied Magnetic Resonance* 30, 473-498 (2006).
76. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. & Steinegger, M. ColabFold: making protein folding accessible to all. *Nature Methods* 1-4 (2022).
77. Cha, S. S., An, Y. J., Lee, C. R., Lee, H. S., Kim, Y. G., Kim, S. J., Kwon, K. K., De Donatis, G. M., Lee, J. H., Maurizi, M. R. & Kang, S. G. Crystal structure of Lon protease: molecular architecture of gated entry to a sequestered degradation chamber. *EMBO J* 29, 3520-3530 (2010).
78. Zorzini, V., Mernik, A., Lah, J., Sterckx, Y. G. J., De Jonge, N., Garcia-Pino, A., De Greve, H., Versées, W. & Loris, R. Substrate Recognition and Activity Regulation of the Escherichia coli mRNA Endonuclease MazF. *Journal of Biological Chemistry* 291, 10950-10960 (2016).
79. Simanshu, D. K., Yamaguchi, Y., Park, J.-H., Inouye, M. & Patel, D. J. Structural basis of mRNA recognition and cleavage by toxin MazF and its regulation by antitoxin MazE in Bacillus subtilis. *Molecular cell* 52, 447-458 (2013).
80. Hagelueken, G., Ward, R., Naismith, J. H. & Schiemann, O. MtsslWizard: In Silico Spin-Labeling and Generation of Distance Distributions in PyMOL. *Applied Magnetic Resonance* 42, 377-391 (2012).

ACCELERATED ARTICLE PREVIEW

81. Campagne, S., Marsh, M. E., Capitani, G., Vorholt, J. A. & Allain, F. H. T. Structural basis for− 10 promoter element melting by environmentally induced sigma factors. *Nature structural & molecular biology* 21, 269-276 (2014).
82. Lane, W. J. & Darst, S. A. The structural basis for promoter− 35 element recognition by the group IV σ factors. *PLoS biology* 4, e269 (2006).
83. Li, L., Fang, C., Zhuang, N., Wang, T. & Zhang, Y. Structural basis for transcription initiation by bacterial ECF σ factors. *Nat Commun* 10, 1153 (2019).

**Acknowledgements**

The synchrotron MX data were collected at beamline P13, operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). We would like to thank Gleb Bourenkov and Isabel Bento for the assistance in using the beamline. We thank Virginius Siksnys for helpful discussions. We thank Sally Shirran and Silvia Synowsky for the mass spectrometry analysis. We would like to thank Norbert Brenner for technical assistance. We thank Marcin Drag and Justyna Grzymska for discussions and an initial peptide screen. M.G. and J.L.S.B. are funded by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy– EXC2151–390873048. M.F.W. acknowledges a European Research Council Advanced Grant (grant number 101018608) and the China Scholarship Council (REF: 202008420207 to H.C.). G.H. is grateful for funding by the Deutsche Forschungsgemeinschaft (grant number HA6805/6-1).

**Author contributions**

C.R. and G.H. conceived and supervised the study and performed initial protein expression and crystallization experiments on CalpL. R.S. and W.B. cloned the initial CalpL construct. N.S., C.R., M.F.W. and G.H. designed experiments. N.S. optimized the purification of CalpL and CalpT, crystallized CalpL, CalpL/$cA_4$ and CalpL/$T_{10}$ and established and performed the cleavage assays, SEC-MALS and DLS experiments. N.S. and M.F.P. cloned all mutants. G.H. and N.S. solved and refined the CalpL, CalpL/$cA_4$ and CalpL/$T_{10}$ structures. J.M. and M.G. designed and performed the SPR experiments. H.C. and M.F.W. planned and performed the ribonuclease assay. H.C., N.S. and M.F.W. cloned and purified CalpS and performed binding and coexpression experiments involving CalpS. S.D.V. performed SAXS experiments. S.D.V. and D.S. performed the SAXS data analysis and interpretation and wrote the corresponding sections. B.E.B. and K.A. performed the pulsed EPR experiments, analyzed the data, prepared figures and wrote the corresponding sections. K.B. and J.L.S.B. designed and performed the RNase-Seq assay. C.R., M.F.W., H.C., N.S., and G.H. analyzed the data and wrote the paper. All authors discussed the results and commented on the manuscript at all stages.

ACCELERATED ARTICLE PREVIEW

**Competing interest statement**

The authors declare no competing interests.

**Additional information statement**

A previous version of this work has been published as preprint: https://doi.org/10.1101/2021.12.06.471393. Requests for materials should be addressed to G.H. (hagelueken@uni-bonn.de).

**Data availability statement**

The crystal structures have been deposited in the PDB with the accession codes 7QDA [https://doi.org/10.2210/pdb7QDA/pdb], 8B0R [https://doi.org/10.2210/pdb8B0R/pdb], 8B0U [https://doi.org/10.2210/pdb8B0U/pdb]. The SAXS data and models have been deposited in the SASBDB with the accession codes: SASDQM4 [https://www.sasbdb.org/data/SASDQM4], SASDQN4 [https://www.sasbdb.org/data/SASDQN4], SASDQP4 [https://www.sasbdb.org/data/SASDQP4], SASDQQ4 [https://www.sasbdb.org/data/SASDQQ4]. The following PDB entries have been used in this study: 2H27 [https://doi.org/10.2210/pdb2H27/pdb], 3K1J [https://doi.org/10.2210/pdb3K1J/pdb], 4ME7 [https://doi.org/10.2210/pdb4ME7/pdb], 4IZJ [https://doi.org/10.2210/pdb4IZJ/pdb], 4LUP [https://10.2210/pdb4LUP/pdb], 5ZX2 [https://doi.org/10.2210/pdb5ZX2/pdb], 5CR2 [https://doi.org/10.2210/pdb5CR2/pdb], 6VM6 [https://doi.org/10.2210/pdb6VM6/pdb], 6SCE [https://doi.org/10.2210/pdb6SCE/pdb], 7RWK [https://doi.org/10.2210/pdb7RWK/pdb]

**Code availability statement**

No custom code was used in this work.

ACCELERATED ARTICLE PREVIEW

## Extended Data Figure legends

**Extended Data Fig. 1 | Purification and Structure of CalpL. a,** Gelfiltration chromatography (Superdex 200 16/60) of CalpL. Inset: SDS-PAGE analysis of the fractions indicated by the black bar in the chromatogram. The experiment was performed multiple times (n > 3 biological replicates). **b,** TM-prediction by the TMHMM 2.0 server[28] vs experimental structure. **c,** Representative electron density of the SeMet CalpL crystal structure. The structural model is drawn in ball-and-stick representation. Selected residues are labeled. The black mesh is a $2mF_o$-$DF_c$ electron density map contoured at 1.0 σ. **d,** Topology diagram of CalpL. For gel source data, see Supplementary Figure 1.

**Extended Data Fig. 2 | CalpL in comparison to structurally related proteins**. **a,** CalpL is drawn as a cartoon model color-coded as in Fig. 1. The Lon protease from *T. onnorineus* (PDB-ID: 3K1J, DALI Z-score: 12.8[77]) is shown as a white cartoon model. **b,** Table listing proteins with similar domain structures. **c,** Surface electrostatics of the Lon protease active site region. The catalytic dyad is marked. The grey line marks the likely substrate binding site. **d,** Superposition of CalpL active site with the acyl-enzyme intermediate of yellowfin asciitis virus protease. CalpL is in sticks representation and color-coded as in Fig. 1. Chain D of structure 4IZJ[27] (residues 630-640) was superimposed on the corresponding residues of CalpL (150-160) leading to an r. m. s. d. of 0.314 Å. Of 4IZJ, only the acyl-enzyme intermediate is shown in sticks mode. Selected residues and the positions of the P1-P3 sites are indicated. **e,** Superposition of CalpL (color scheme as in Fig. 1) with the Cap4 protein (white, PDB-ID: 6VM6[23]). **f,** Superposition of the CalpL SAVED domain (color scheme as in Fig. 1) with the Cap5 protein (white, PDB-ID: 7RWK[29]). **g,** Superposition of the CalpL SAVED domain (color scheme as in Fig. 1) with the CARF domains of the Can1 protein (white, PDB-ID: 6SCE[17]).

**Extended Data Fig. 3 | The CalpL/$cA_4$ complex a,** Close-up of $cA_4$ (green) bound to the SAVED domain of CalpL. The blue mesh is a $2mF_o$-$DF_c$ electron density map contoured at 1.0 σ. **b,** Superposition of CalpL apo (white) onto the cA4 complex structure (color coded as in Fig. 1). **c,** Structural alignment of the SAVED domains of CALP/$cA_4$ and Cap4/$cA_3$ (white).

**Extended Data Fig. 4 | CalpT is a MazF homolog and the target of the CalpL protease. a, b,** A superposition of the predicted CalpT structure (compare Fig. 2B) with one monomer of the MazF/ssRNA complex (purple/orange) (PDB-IDs: 5CR2[78]). The AlphaFold2[33] prediction confidence is mapped onto the CalpT structure (pLDDT[49], predicted local distance difference test). **b,** A superposition of the predicted CalpT structure (compare Fig. 2b) with one monomer of the MazE/F complex (PDB-IDs: 4ME7[79]). The AlphaFold2[33] prediction confidence is mapped onto the CalpT structure (pLDDT[49], predicted local distance difference test). **c,** Gel filtration chromatography (Superdex 75 16/60) of CalpT The experiment was performed multiple times (n > 3 biological replicates). According to the MALS data in Fig. 3, isolated CalpT behaves as a monomer. **d,** Peptide fingerprints of cleavage bands. The indicated gel-bands were cut from the gel and submitted for identification at the Mass spectrometry and proteomics facility at the University of St Andrews (Fife, UK, https://mass-spec.wp.st-andrews.ac.uk). Red letters indicate peptides that were identified in the respective sample. The experiment was performed once. **e,** Mutational analysis of potential CalpL cleavage sites in CalpT. The positions of the mutants are indicated as magenta spheres on the right. (pLDDT[49], predicted local distance difference test) The experiment was performed twice (n = 2 technical replicates). **e)** SDS-PAGE analysis of the fractions indicated by the black bar in a) The experiment was performed multiple times (n > 3 technical replicates). For gel source data, see Supplementary Figure 1.

ACCELERATED ARTICLE PREVIEW

**Extended Data Fig. 5 | Characterization of the CalpL/T complex. a,** Single cycle kinetics SPR data of the CalpL/T interaction. The interaction is very strong but cannot be satisfyingly fitted with a 1:1 binding model. The experiment was performed twice (n = 2 technical replicates). **b,** As **a**), but an artificial construct of an unspecific VHH fused to CalpT was used as analyte in this experiment. The interaction is very similar to the CalpL/T interaction. The experiment was performed twice (n = 2 technical replicates). **c,** Schematics of two artificial constructs containing the CalpL cleavage site. **d,** CalpL cleaves an artificial construct of an unspecific VHH fused to $CalpT_{10}$ but not a construct of two VHHs fused by the CalpL cleavage site. The experiment was performed once. **e,** SEC-MALS traces (solid lines: $UV_{280}$, dashed lines: $MW_{MALS}$) of proteolysis reactions with different combinations of CalpL S152A, CalpT, and cOA. The schematic indicates the molecular species behind the individual peaks. The experiments were performed twice with slightly different buffer conditions (n=2 technical replicates). **f,** Binding of CalpL wt to the indicated CalpT mutants in the absence of $cA_4$. The schematic indicates the position of the mutant in the CalpL/T complex. The experiments were performed once. **g,** Representative electron density of the $CalpL_{/T10}$ crystal structure. Selected residues are labeled. The black mesh is a $2mF_o$-$DF_c$ electron density map contoured at 1.0 σ. **h,** SEC-SAXS experiment of the CalpL/$T_{10}$ complex. The experiment was performed once. Thirty sample intensity frames and sixty buffer intensity frames were collected and averaged. For each data set and angular point the errors were computed following the Poisson statistics. The data points represent the average intensity difference (sample-buffer) and the error bars represent the standard deviation. For gel source data, see Supplementary Figure 1.

**Extended Data Fig. 6 | $cA_4$ induced oligomerization of CalpL studied by DLS and SAXS. a,** Dynamic light scattering experiments (six timeseries, each series marked by a dashed circle, single data points are shown) at different protein concentrations and in the absence (t=0: light grey to t= 60 min: dark grey) and presence (t=0: cyan to t= 60 min: violet) of $cA_4$ reveal a $cA_4$-dependent oligomerization of CalpL. The experiment was performed twice (n=2 technical replicates) **b,** SAXS experiments at different concentrations. The experiments were performed once. For each experiment, thirty sample intensity frames and sixty buffer intensity frames were collected and averaged. For each data set and angular point the errors were computed following the Poisson statistics. The data points represent the average intensity difference (sample-buffer) and the error bars represent the standard deviation. **c,** *Ab initio/rigid-body* model of a CalpL dimer created with DAMMIF and SASREFMX by a global fit of a monomer-dimer mixture to the different concentrations (red lines). The crystal structure of the CalpL monomer is shown on the same scale.

**Extended Data Fig. 7 | Probing the RNase activity of the activated toxin and checking for cA4 induced dimerization of CalpT with pulsed EPR. a**, Fluorescence image of the denaturing PAGE to determine ribonuclease activity of the reactions in **b**) against six fluorescently labelled RNA substrates (listed in **c**)). No cleavage was observed after 30 min incubation with RNAs at 60 °C. The experiment was performed three times (n=3 biological replicates) **b,** SDS-PAGE analysis of $cA_4$-induced cleavage of CalpT (33 kDa) by CalpL. Cleavage is complete after 60 min at 60 °C. The experiment was performed three times (n=3 biological replicates) **c,** Sequences of the RNA substrates **d,** left: MazF was incubated with a single stranded RNA library containing 10 random bases. Illumina sequencing was used to check for sequences that were cleaved by MazF. Compared to a control reaction without MazF, sequences containing the known MazF target site (ACA) were depleted. right: same experiment but with CalpL/T ± $cA_4$ instead of MazF. No off-diagonal sequences and hence no ssRNase activity were observed. The experiment was performed two times (n=2 biological replicates). **e,** Oligonucleotides for the experiments in **d**) **f,** AlphaFold2 dimer models of $CalpT_{23}$. **g,** Best model (pLDDT[49], predicted local distance difference test) including MTSSL spin label[80]. **h,** X-

ACCELERATED ARTICLE PREVIEW

band cw-EPR spectrum of of CalpL/T E119R1. The amount of free label (sharp spikes) is ~10%. The labelling efficiency determined as ~100%. **i,** PELDOR time traces of CalpL/T E119R1 in the presence (red) and absence (black) of $cA_4$. **j,** Consensus distributions and corresponding uncertainty bands. Colored bars indicate reliability ranges (green: shape reliable; yellow: mean and width reliable; orange: mean reliable; red: no quantification possible). Predicted distance calculated with mtsslWizard[80]. The EPR experiment was performed twice (n=2 technical replicates). For gel source data, see Supplementary Figure 1.

**Extended Data Fig. 8 | AlphaFold2 predictions of CalpS. a**, Prediction of CalpS alone. The protein is shown as cartoon and colored according to the prediction confidence (pLDDT[49], predicted local distance difference test) **b,** Prediction of the CalpT/S complex. **c,** Superposition of CalpS with 4LUP[81] and 2H27[82] identify the DNA binding regions of CalpS. **d,** Model of CalpS in the context of a RNAP/ECF σ-factor/promotor complex (PDB: 5ZX2[83], grey, yellow, green) from *M. tuberculosis*. Note that the linker region between the σ2 and σ4 subunits of CalpS has been cut to allow the superposition of the $\sigma_2$ and $\sigma_4$ domains onto those of 5ZX2. The linker is long enough to bind to the RNAP in a similar way as the σ-factor in the 5ZX2 structure (yellow).

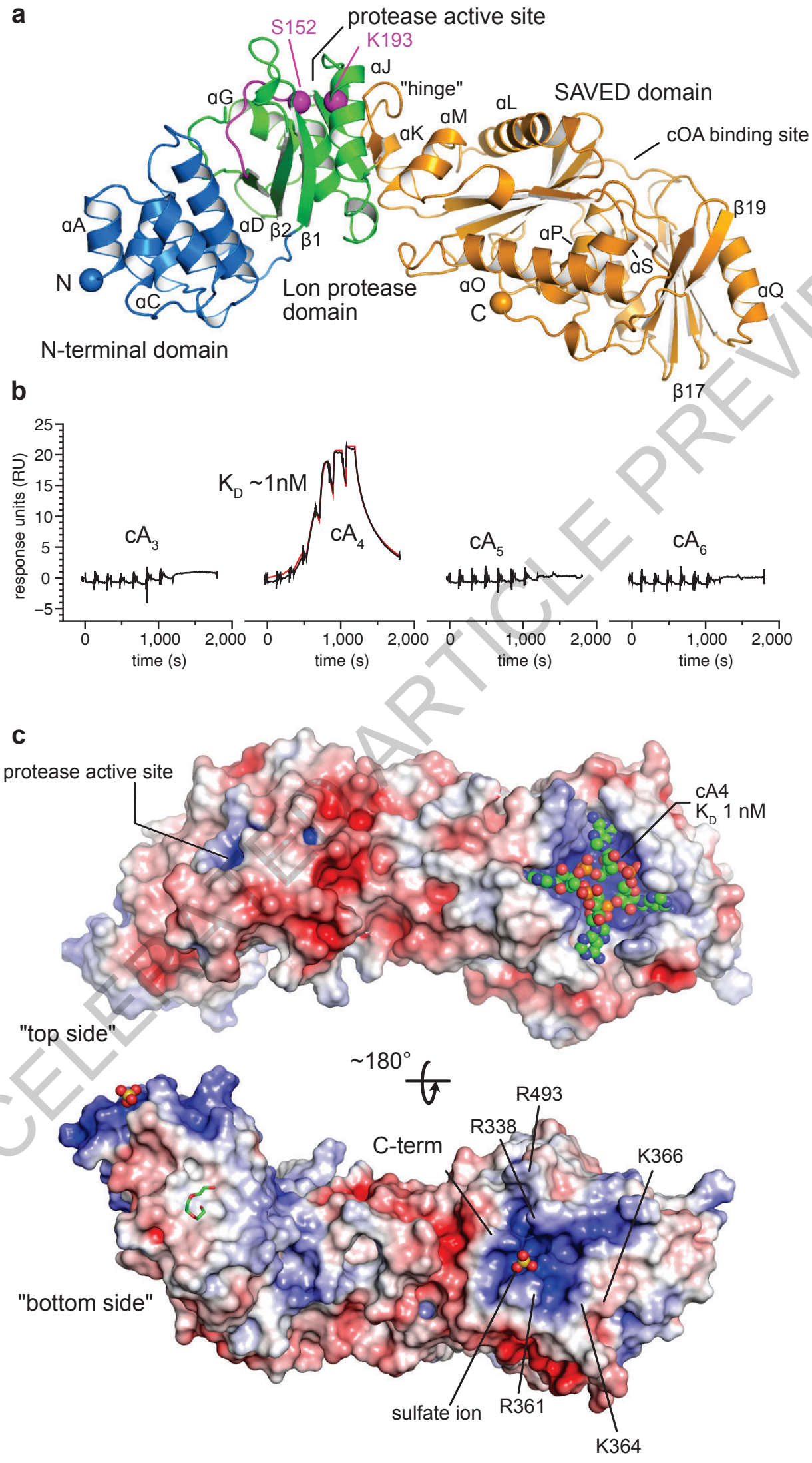
**Extended Data Table 1 | Data collection and refinement statistics for crystallographic structures.**

*Values in parenthesis are for the highest-resolution shell.
One crystal was used for each data collection.

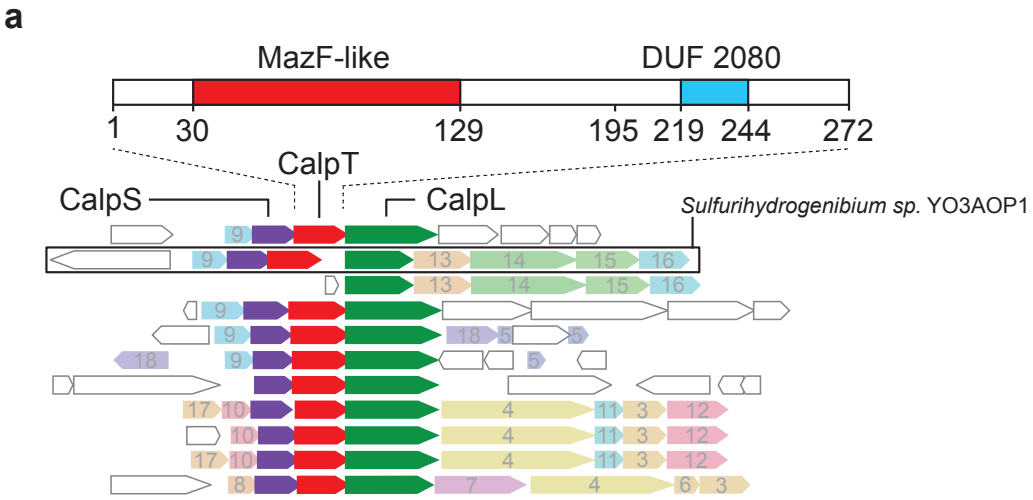
**Extended Data Table 2 | SAXS data collection and parameters.**

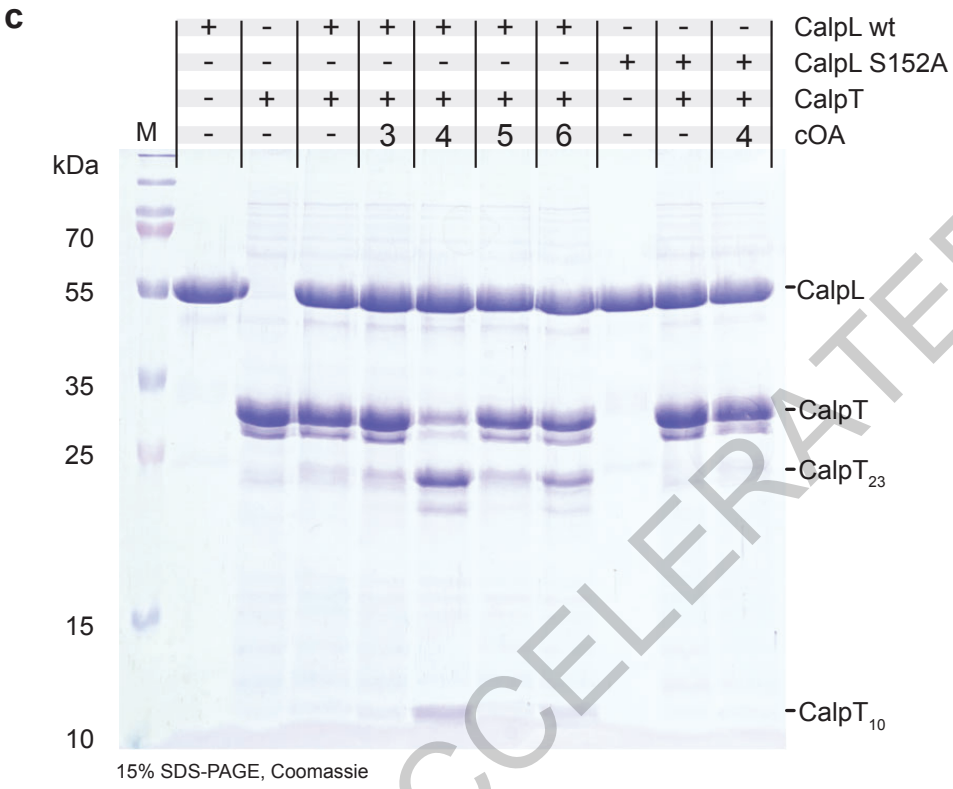
*Rg (radius of gyration) from Guinier approximation, †Rg from real-space pair distance distribution function, ‡Largest intramolecular distance, Dmax, §Bayesian molecular mass (Mr) estimate and credibility interval (>90% probability), ‖Porod volume from regularized curve, ¶Mr from VP, #,☆Volume and Mr of the *ab initio* models, **Normalized Spatial Discrepancy of the *ab initio* reconstructions, ††The $\chi^2$ value is given for the most representative DAMMIF *ab initio* reconstruction.

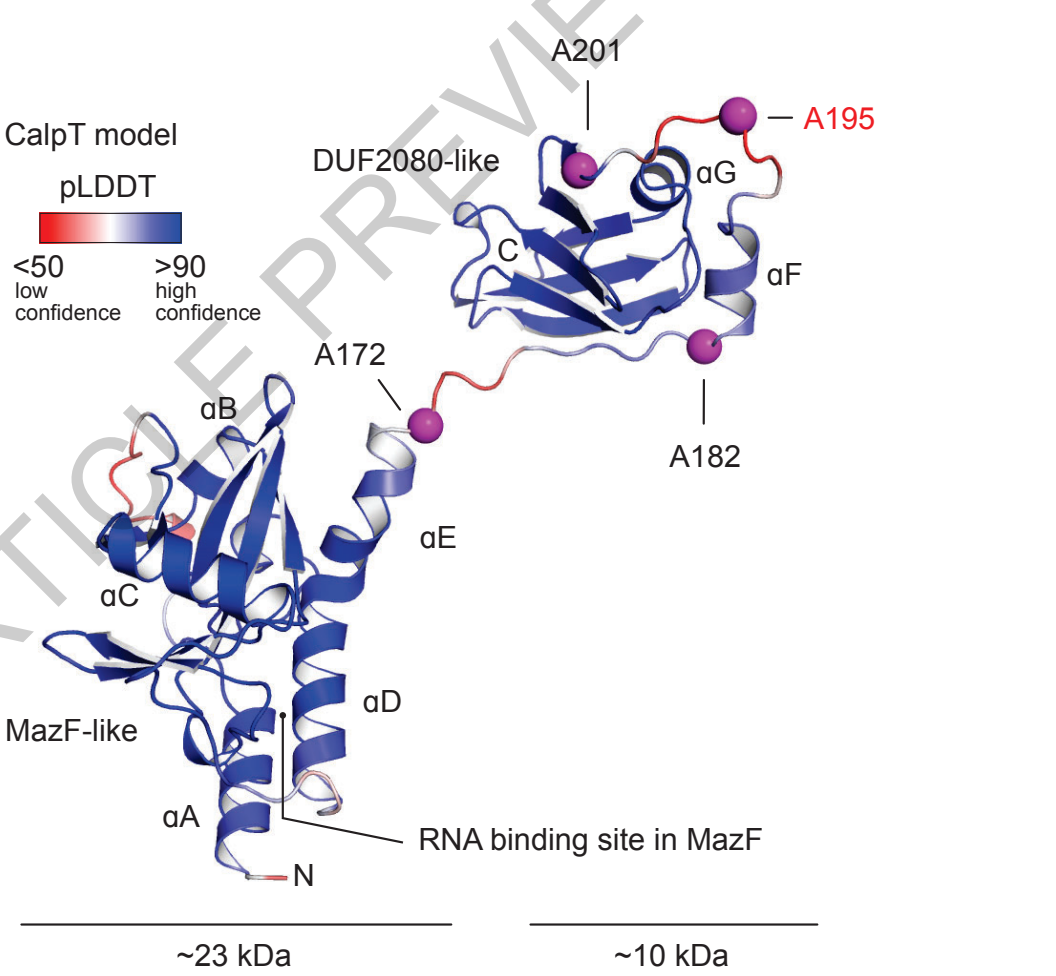
ACCELERATED ARTICLE PREVIEW

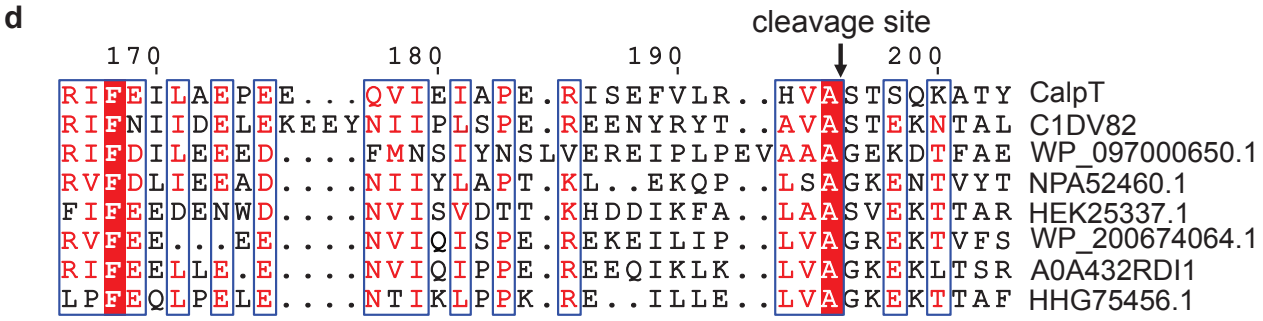



ACCELERATED ARTICLE PREVIEW

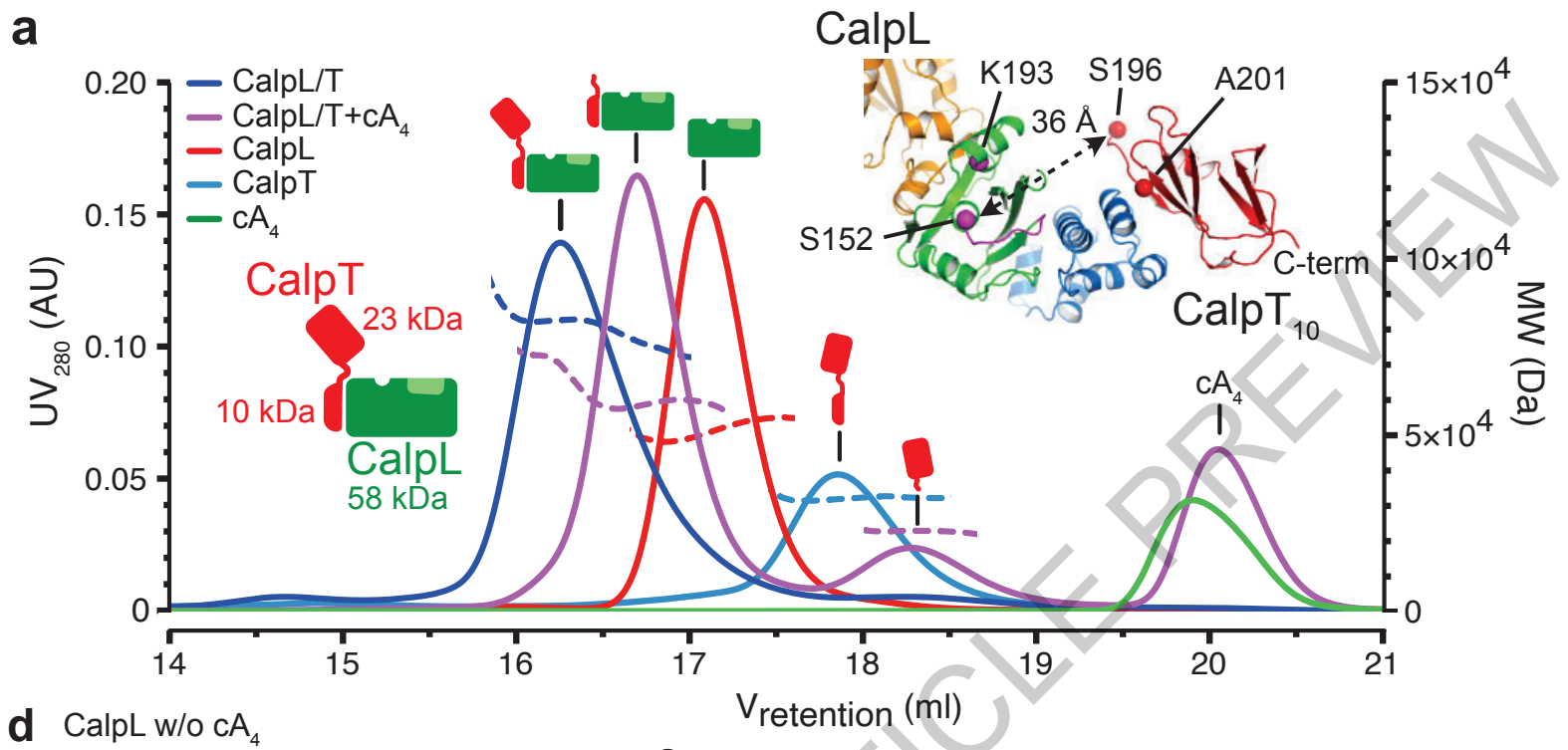


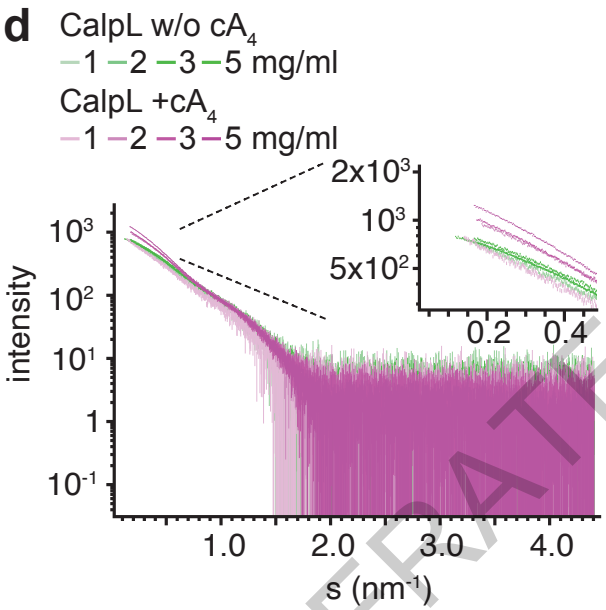



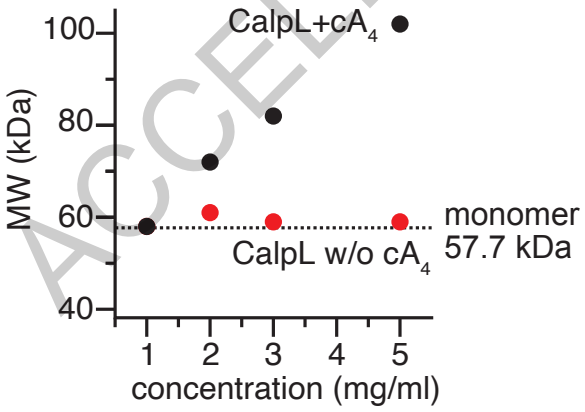



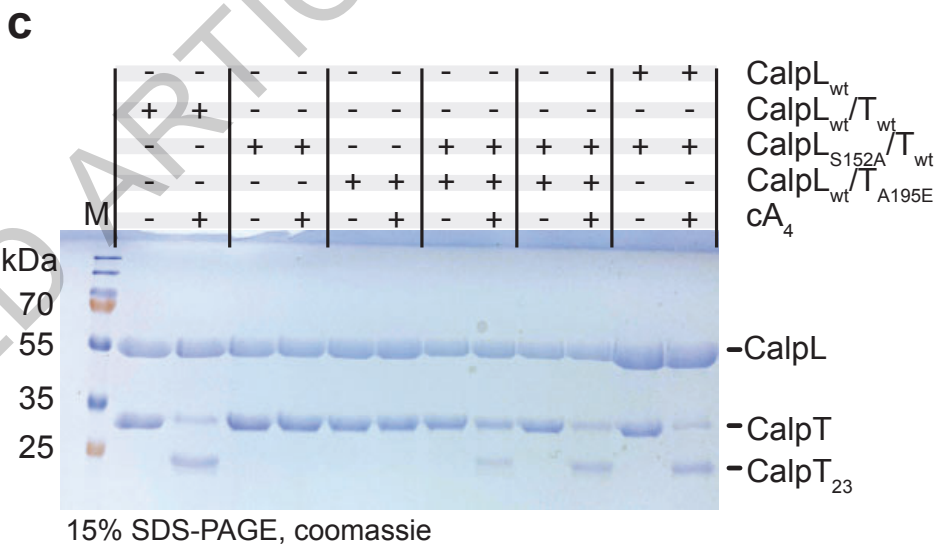



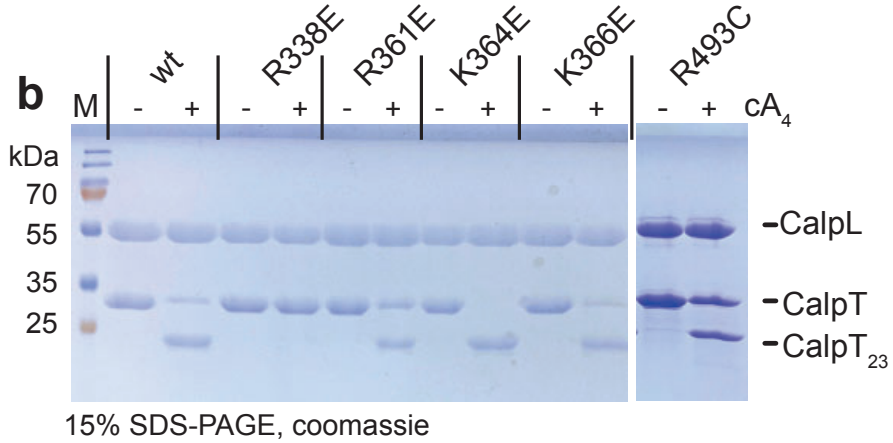

ACCELERATED ARTICLE PREVIEW



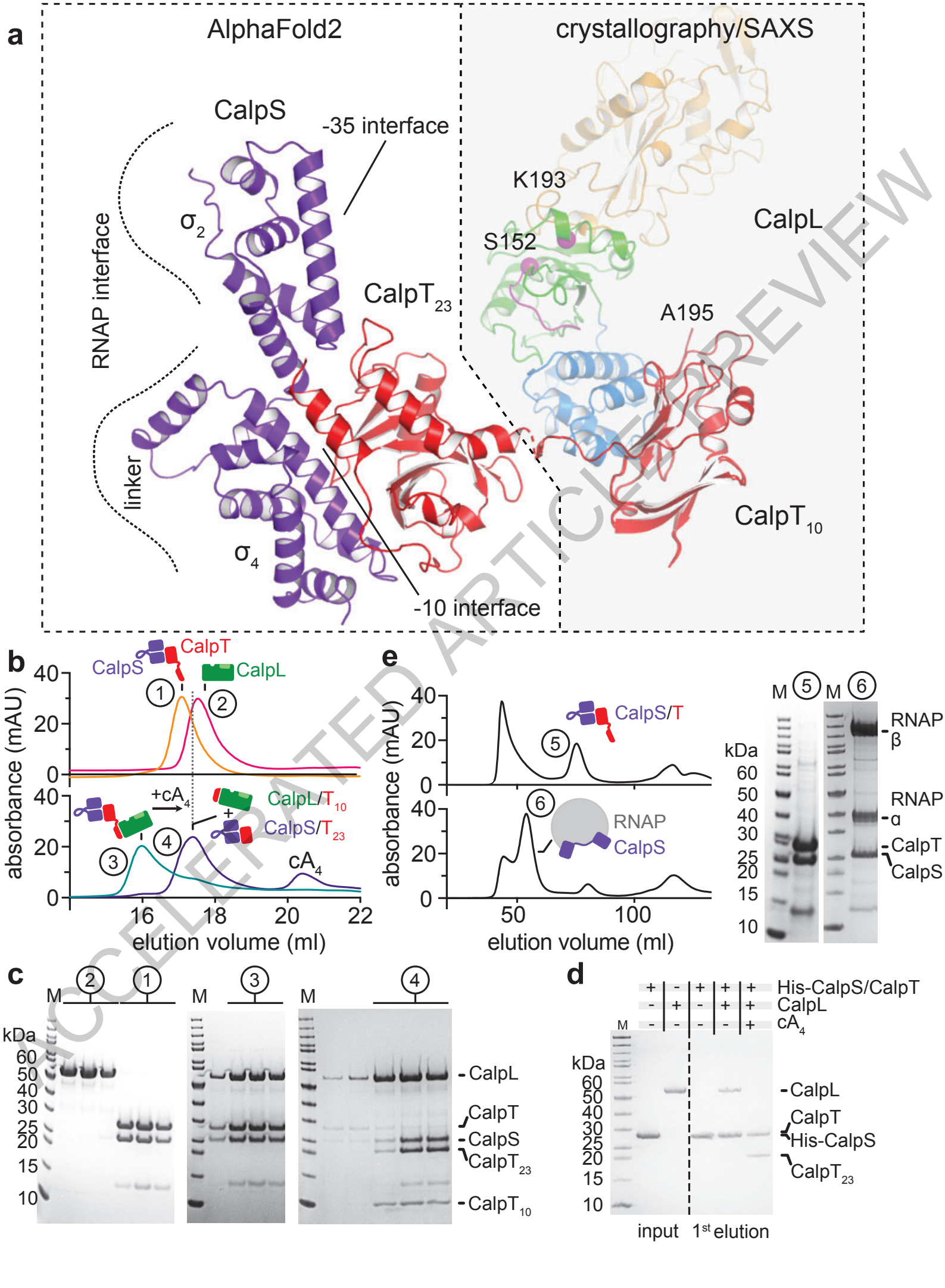



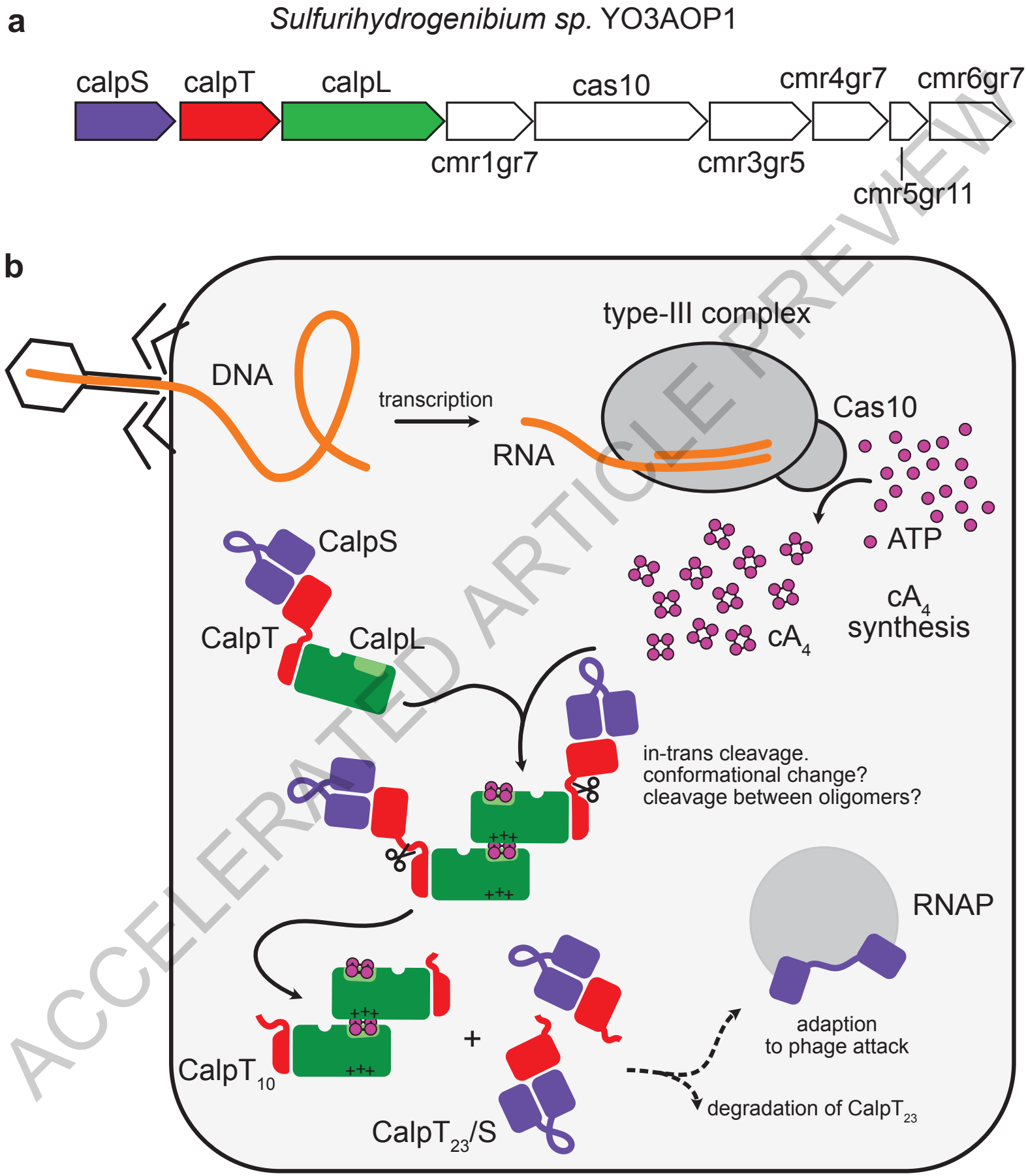



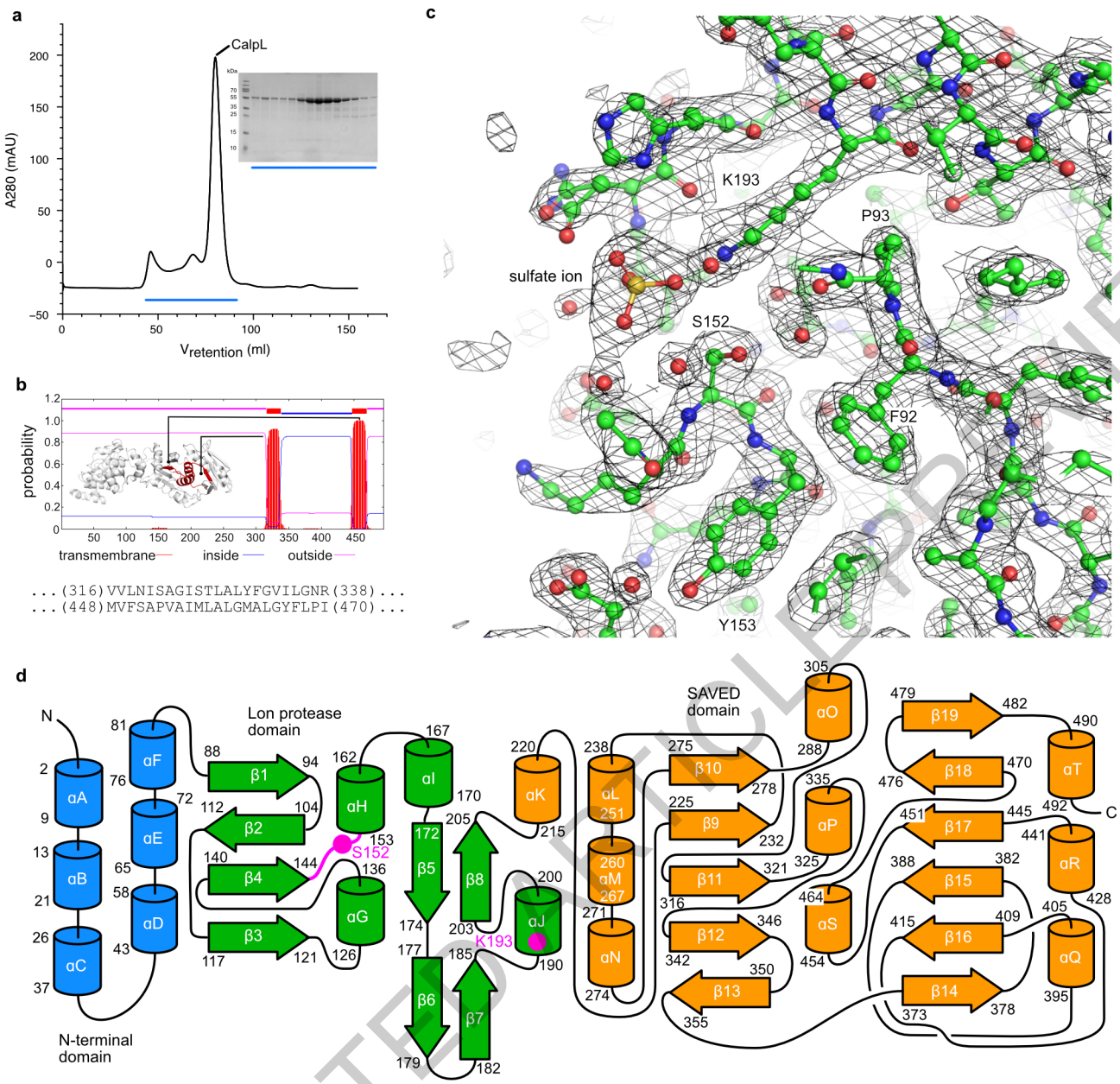

ACCELERATED ARTICLE PREVIEW




ACCELERATED ARTICLE PREVIEW

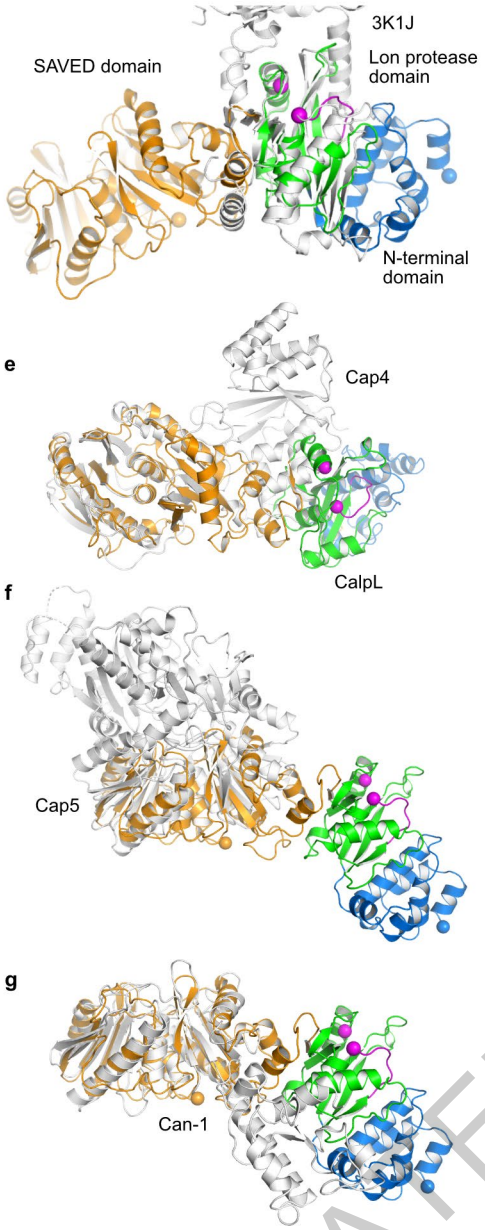



ACCELERATED ARTICLE PREVIEW



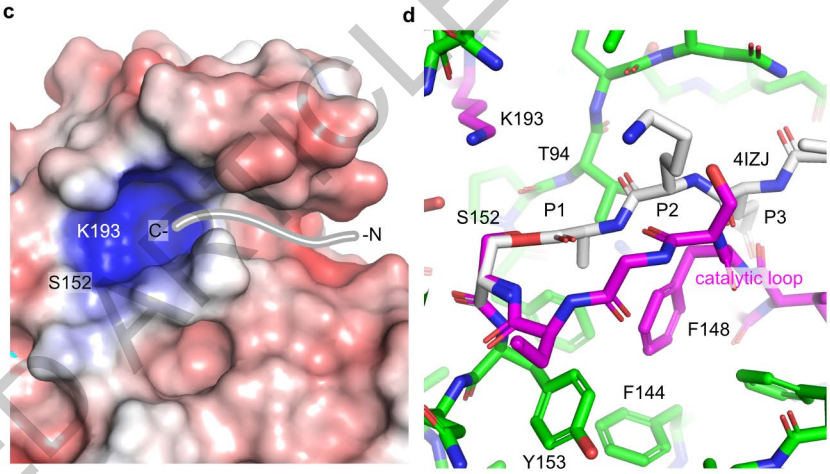

**Extended Data Fig. 1**

ACCELERATED ARTICLE PREVIEW



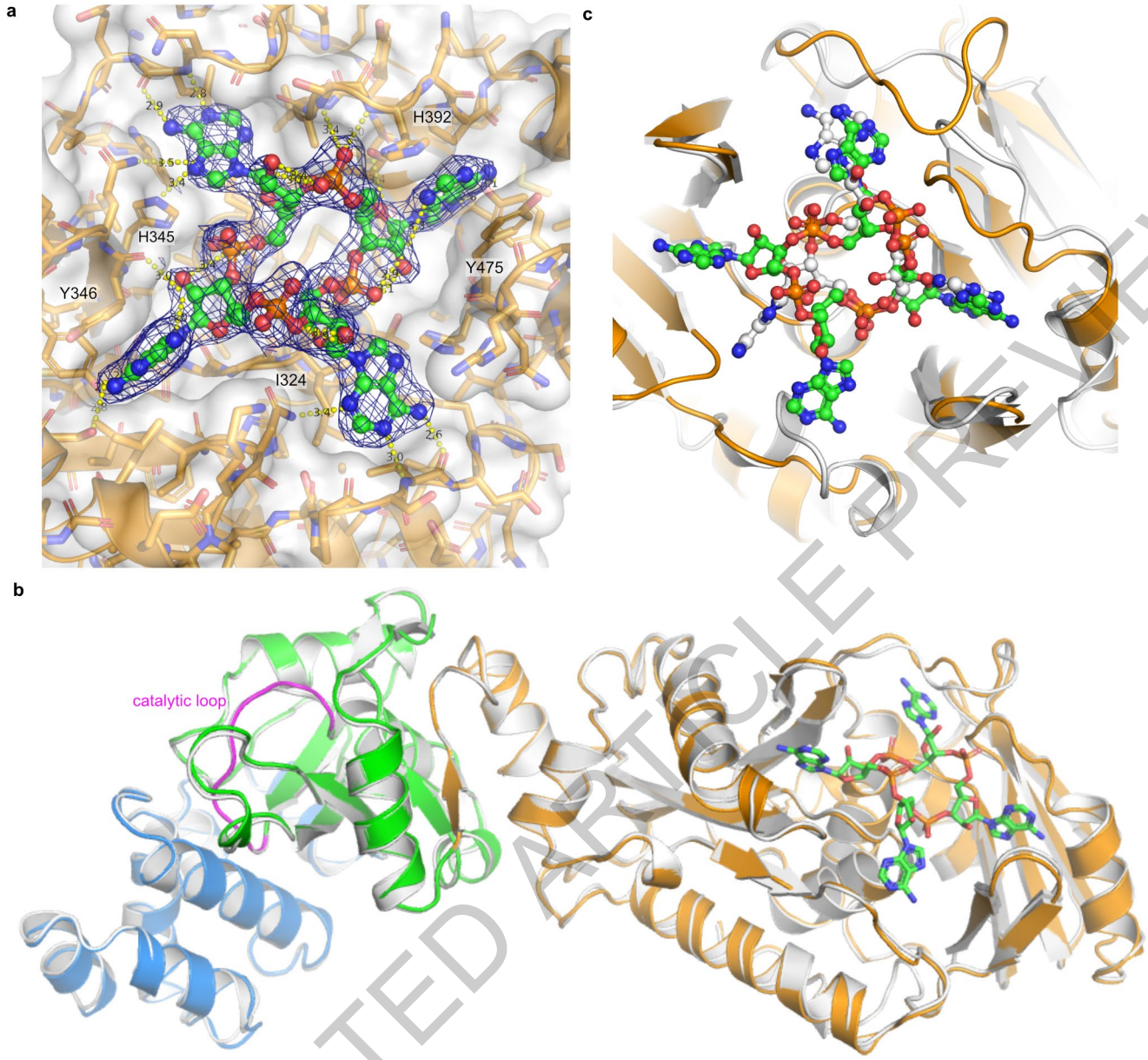

b

| Name | Organism | Domain | PDB | DALI[35] Z-score | Similar to CalpL domain |
|---|---|---|---|---|---|
| ATP-dependent Lon protease | *Thermococcus onnurineus* | Lon | 3KIJ[75] | 12.8 | Lon |
| VP-4 protease | *Tellina virus* | Lon | 3P06[76] | 11.0 | Lon |
| RadA helicase | *Streptococcus pneumoniae* | Lon | 5LKM[77] | 11.7 | Lon |
| ATP-independent Lon protease | Yellowfin ascites virus | Lon | 4IZJ[28] | 9.8 | Lon |
| Cap4 | *Acinetobacter baumannii* | SAVED | 6VM6[23] | 17.1 | SAVED |
| Cap5 | *Asticcacaulis sp.* | SAVED | 7RWK[30] | 9.9 | SAVED |
| Can1 | *Thermus thermophilus* | CARF | 6SCE[17] | 8.0 | SAVED |
| Katanin | *Mus musculus* | N-terminal protein-protein interaction domain | 5OW5[31] | 6.2 | N-terminal domain |



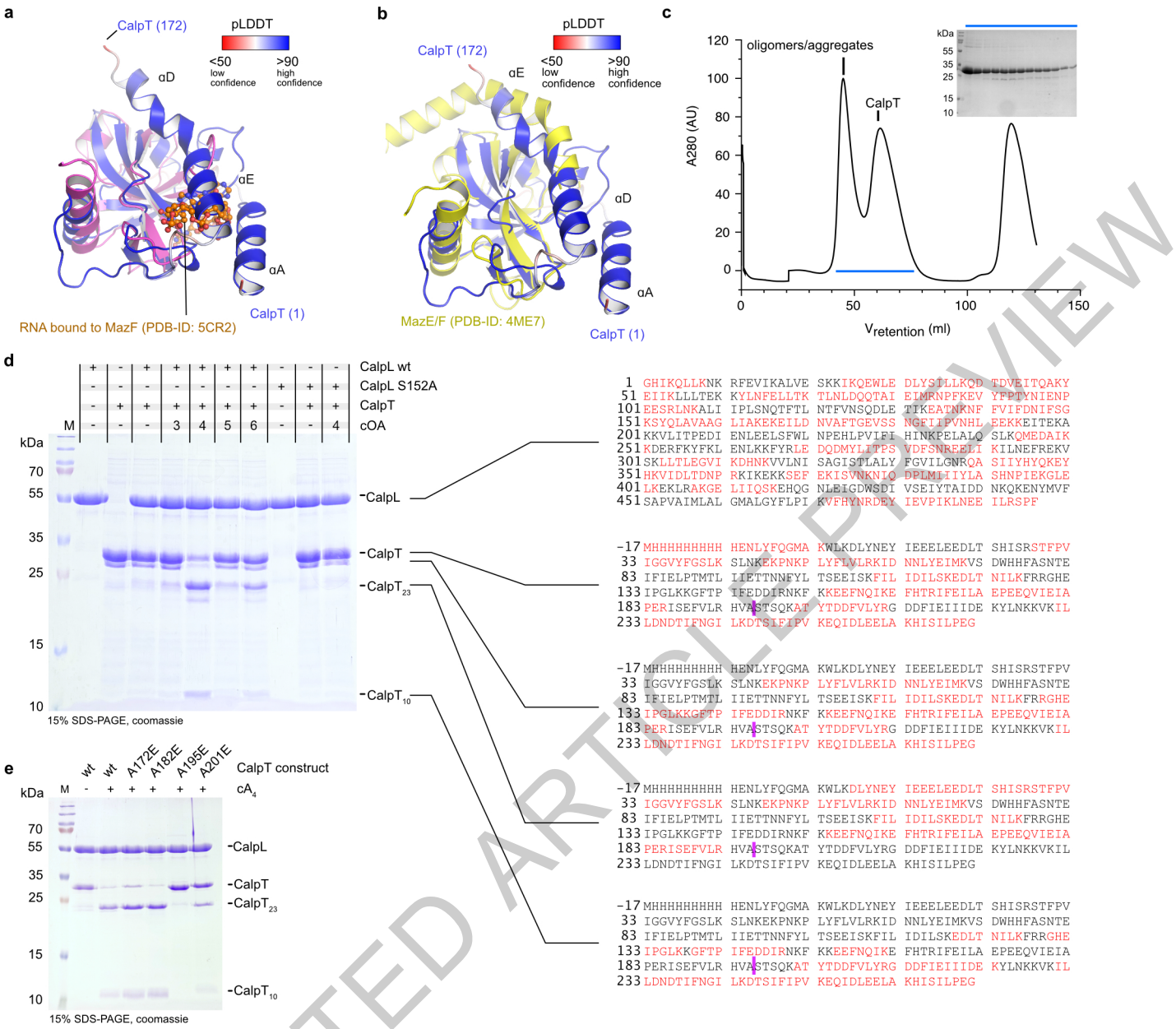

**Extended Data Fig. 2**

ACCELERATED ARTICLE PREVIEW



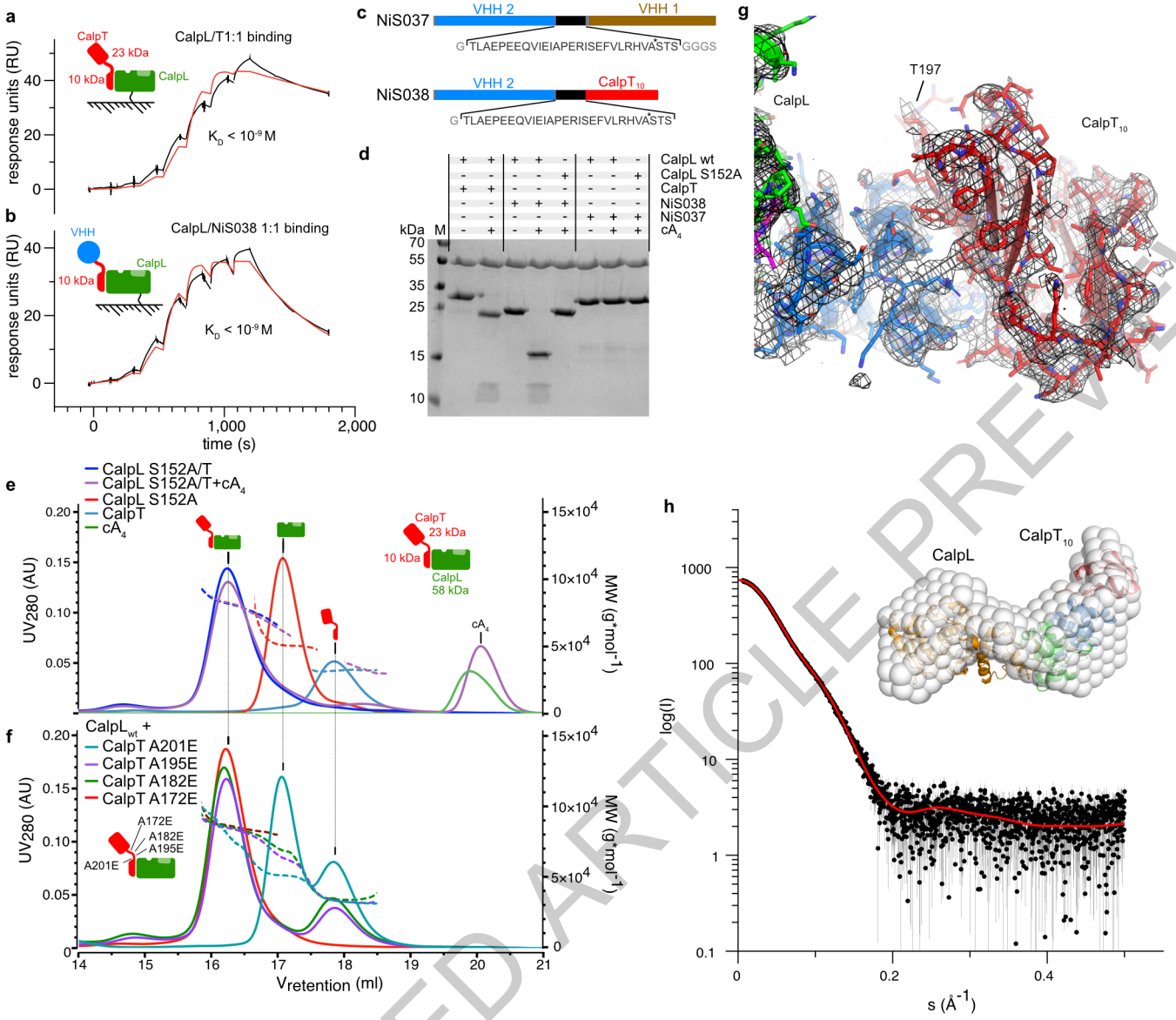

**Extended Data Fig. 3**

ACCELERATED ARTICLE PREVIEW



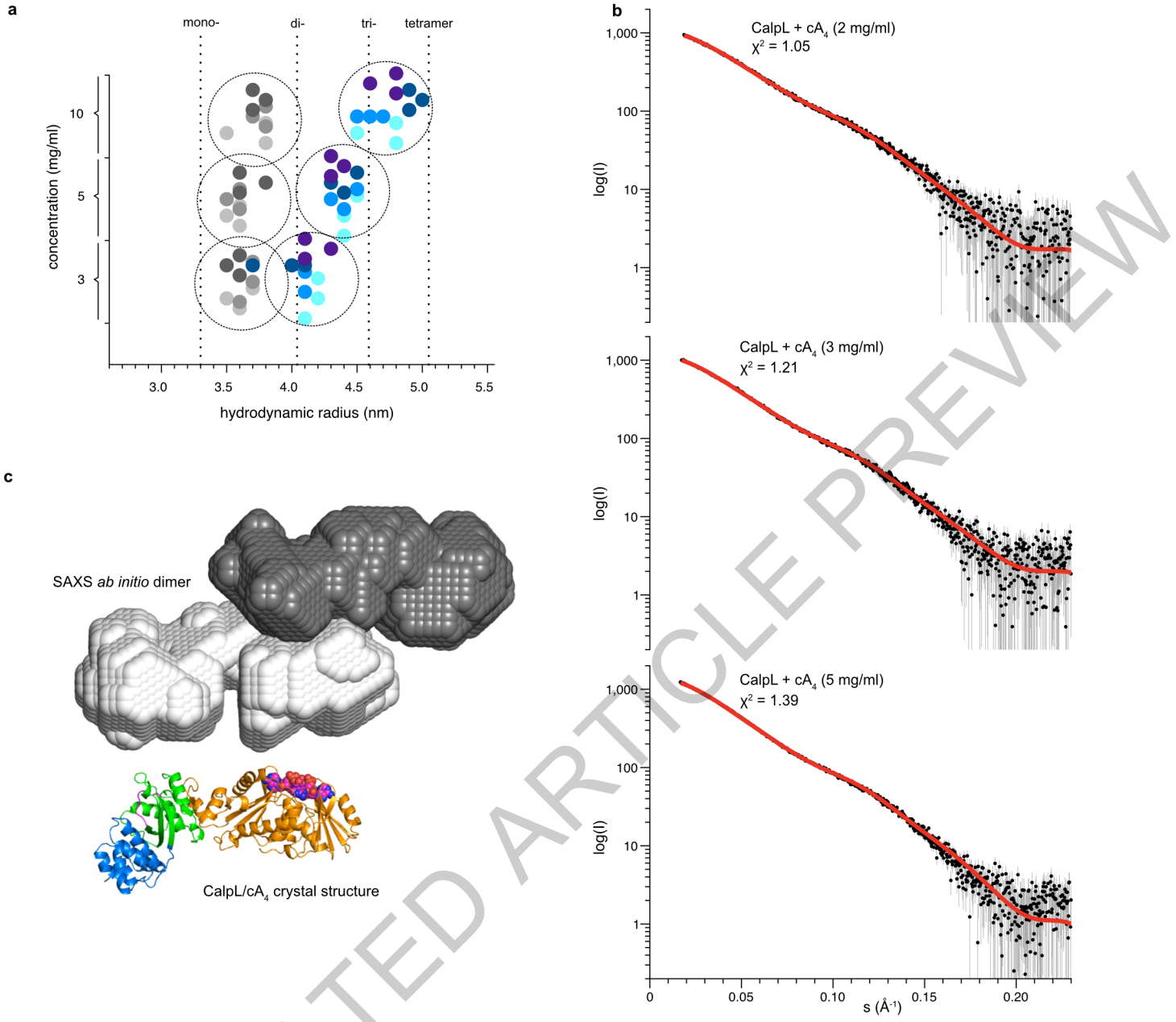

**Extended Data Fig. 4**

ACCELERATED ARTICLE PREVIEW



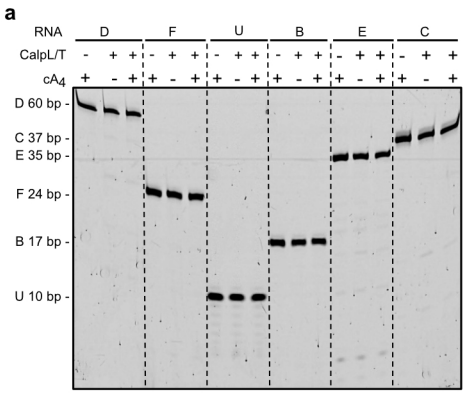

**Extended Data Fig. 5**

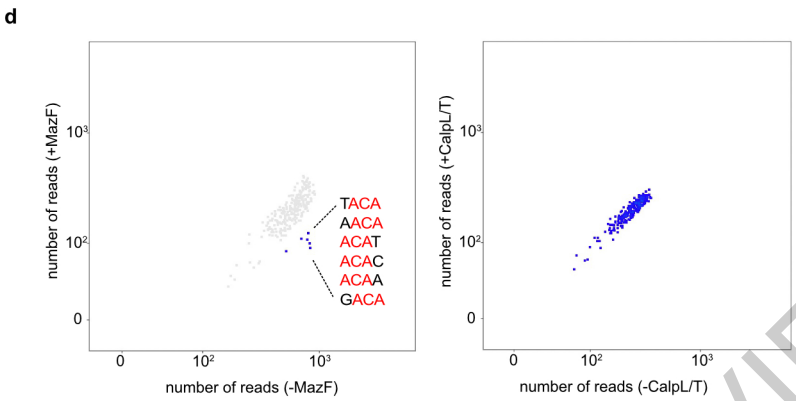
ACCELERATED ARTICLE PREVIEW

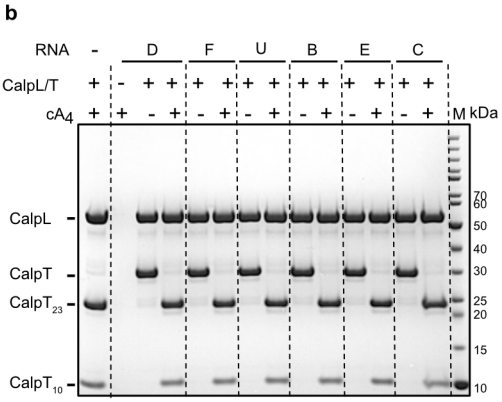



**Extended Data Fig. 6**

ACCELERATED ARTICLE PREVIEW



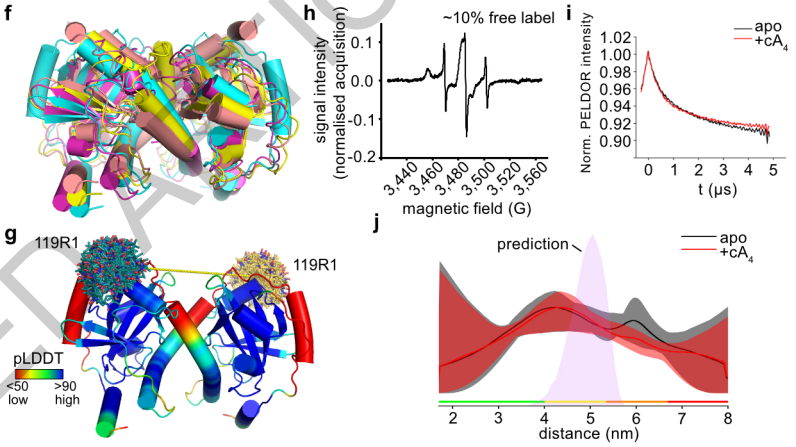




e

| Name | Sequence |
|---|---|
| RNase-Seq_fwd_S0 | acACTCTTTCCCTACACGACGCTCTTCCGATCTATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S1 | acACTCTTTCCCTACACGACGCTCTTCCGATCT C ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S2 | acACTCTTTCCCTACACGACGCTCTTCCGATCT GC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S3 | acACTCTTTCCCTACACGACGCTCTTCCGATCT AGC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S4 | acACTCTTTCCCTACACGACGCTCTTCCGATCT CAAC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S6 | acACTCTTTCCCTACACGACGCTCTTCCGATCT TGCACC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S7 | acACTCTTTCCCTACACGACGCTCTTCCGATCT ACGCAAC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_fwd_S8 | acACTCTTTCCCTACACGACGCTCTTCCGATCT GAAGACCC ATCTTGTGGAAAGGACGAAACAC |
| RNase-Seq_rev | tgACTGGAGTTCAGACGTGTGCTCTTCCGATCTTCAAGTTGATAACGGACTAGCCT |
| Indexing_fwd | AATGATACGGCGACCACCGAGATCTACACNNNNNNNNNNACACTCTTTCCCTACACGACGCT |
| Indexing _rev | CAAGCAGAAGACGGCATACGAGATNNNNNNNNNNGTGACTGGAGTTCAGACGTGTGCT |
| random-RNA N10 | GGGAUCUUGUGGAAAGGACGAAACACA NNNNNNNNNN AAGGCUAGUCCGUUAUCAACUUGA |

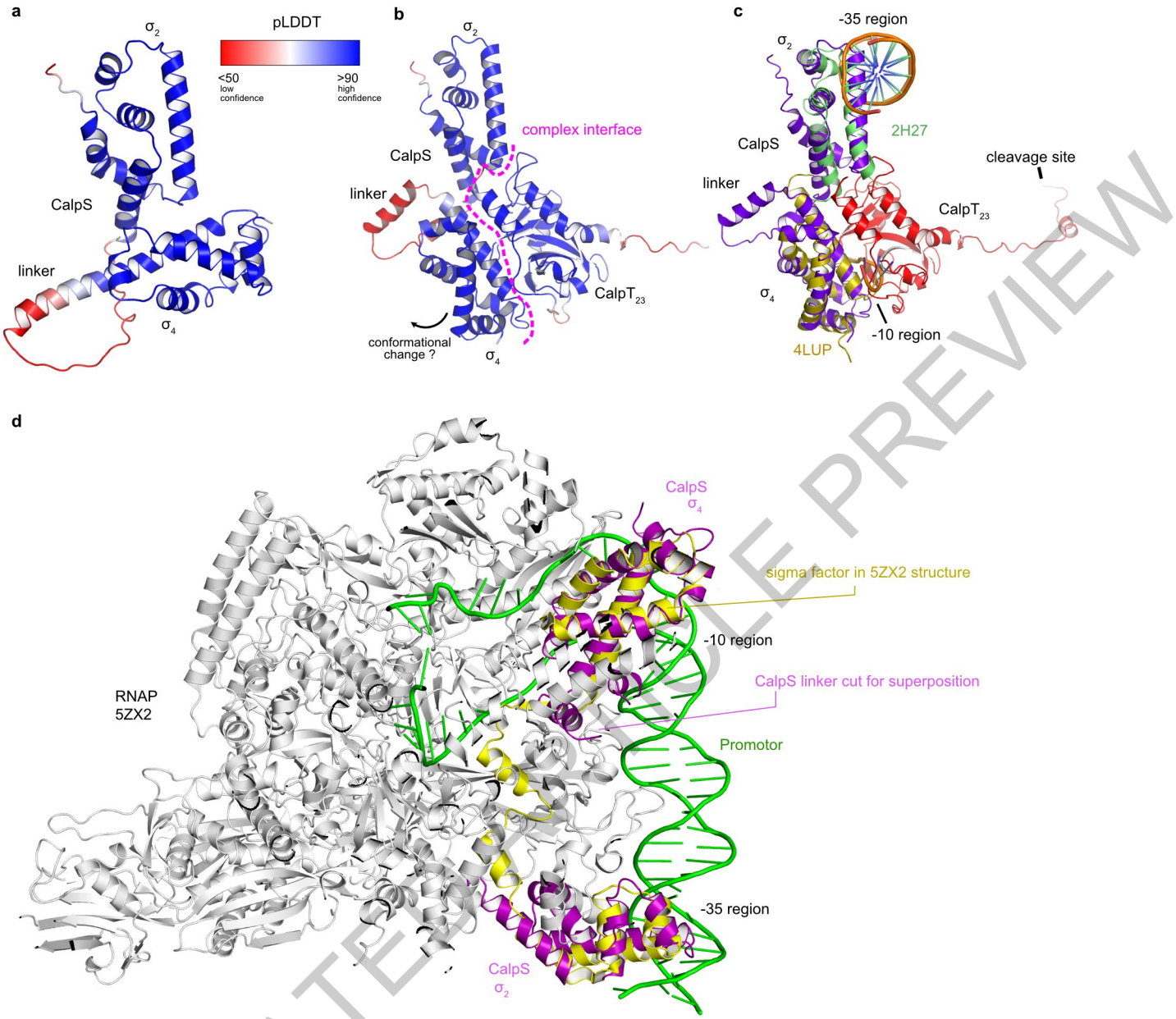
c

| Name | Sequence | Bases |
|---|---|---|
| RNA D | 5'-FAM-AUUGAAAGACCAUACCCAACUUCUAACAAC GUCGUUCUUAACAACGGAUUAAUCCCAAAA | 60 |
| RNA F | 5'-FAM-CUUU/CAAU/UCUAUAGUAGAUUAGC | 24 |
| RNA U | UUUUUUUUUU-FAM-3' | 10 |
| RNA B | 5'-FAM-UGAUAAUCUCUUAUAGA | 17 |
| RNA E | 5'-FAM-AUGUAGAUGUAUUCCAGUAUAAUAAGGAU UAAGAC | 35 |
| RNA C | 5'-FAM-UGUCGUCAGACCCAAAACCCCGAGAGGGGA CGGAAAC | 37 |




**Extended Data Fig. 7**

ACCELERATED ARTICLE PREVIEW




**Extended Data Fig. 8**

ACCELERATED ARTICLE PREVIEW

| | CalpL | CalpL/cA4 | CalpL/T10 |
|---|---|---|---|
| Data collection | | | |
| Space group | P 3(2)21 | P 3(2)21 | P 2(1) |
| Cell dimensions | | | |
| *a*, *b*, *c* (Å) | 68.9, 68.9, 255.4 | 69.5, 69.5, 256.3 | 89.8, 96.6, 131.7 |
| a, b, g(°) | 90.0, 90.0, 120.0 | 90.0, 90.0, 120.0 | 90.0,98.2, 90.0 |
| Resolution (Å) | 54.06(2.18 - 2.1)* | 49.2(2.28-2.2) | 49.31(3.41- 3.29) |
| *R*merge | 0.054(0.3444) | 0.03466(0.2469) | 0.1467(1.559) |
| *I* / s*I* | 23.20(2.36) | 8.37(2.33) | 9.92(1.29) |
| Completeness (%) | 99.02 (90.28) | 99.98 (100.00) | 99.12 (99.34) |
| Redundancy | 2.0 (2.0) | 2.0 (2.0) | 6.8 (6.9) |
| | | | |
| Refinement | | | |
| Resolution (Å) | 54.06(2.18-2.1)* | 49.2(2.28-2.2) | 49.31(3.41-3.29) |
| No. reflections | 41842(3732) | 37614(3685) | 33758(3319) |
| *R*work / *R*free | 0.1932/0.2255 | 0.2450/0.2747 | 0.1968/0.2313 |
| No. atoms | 4487 | 4246 | 9441 |
| Protein | 4175 | 4096 | 9392 |
| Ligand/ion | 69 | 113 | 47 |
| Water | 243 | 37 | 2 |
| *B*-factors | | | |
| Protein | 43.7 | 41.7 | 116.6 |
| Ligand | - | 48.1 | - |
| Water | 46.6 | 41.8 | 69.3 |
| R.m.s. deviations | | | |
| Bond lengths (Å) | 0.003 | 0.002 | 0.002 |
| Bond angles (°) | 0.567 | 0.535 | 0.533 |
| PDB accession code | 7QDA | 8B0R | 8B0U |

## Extended Data Table 1

ACCELERATED ARTICLE PREVIEW

| Data collection parameters (I, II refer to the experimental sessions) | |
|---|---|
| Radiation source | Petra III (DESY) |
| Current | 120 mA (I) or 100 mA (II) |
| Beamline | P12 |
| Detector | Pilatus 6M |
| Wavelength (X-ray Energy) | 0.124 nm (10 keV) (I) or 0.155 nm (8 keV) (II) |
| Sample-detector distance | 3.1 m |
| s-range | 0.024-7.36 $nm^{-1}$ (I) or 0.022-5.83 $nm^{-1}$ |
| Capillary | 0.9 mm, quartz |
| Exposure time | 1 s/frame (SEC-SAXS) (I,II) or 30 x 0.095 s (Batch) (II) |
| Exposure temperature | 293.2 K |
| Attenuation | 30 % Transmission (I) or 50% Transmission (II) |

| Parameter (unit) | (I) CalpL | (I) CalpL/T10 | (II) CalpL+cA4 |
|---|---|---|---|
| Mode | SEC-SAXS | SEC-SAXS | SEC-SAXS |
| Rg Guinier (nm)* | 3.11±0.01 | 3.41±0.01 | 3.21±0.01 |
| Rg P(r) (nm)† | 3.18±0.01 | 3.50±0.01 | 3.30±0.01 |
| Dmax (nm) ‡ | 10.3±0.5 | 11.7±0.5 | 10.3±0.5 |
| Mr Bayes (kDa)§ | 53.2 | 62.4 | 58.2 |
| | (51.4-57.5) | (57.5-64.6) | (50.3-60.2) |
| VP ($nm^3$)‖ | 79.8 | 93.1 | 82.0 |
| Mr VP (kDa) ¶ | 49.9 | 58.2 | 51.2 |
| VDAM ($nm^3$)# | 99.5 | 119 | 102 |
| MrDAM (kDa)☆ | 49.7 | 59.3 | 51.2 |
| NSD** | 0.845±0.026 | 0.953±0.157 | 0.829±0.016 |
| Fit χ2 *ab initio*†† | 1.08 | 0.94 | 1.25 |
| Concentration, volume injected (mg/mL, µL) | 8.8, 40 | 7.6, 40 | 10, 35 |
| Mr from sequence (kDa) | 57.7 | 66.5 | 57.7 |

## Extended Data Table 2

ACCELERATED ARTICLE PREVIEW

nature portfolio

Corresponding author(s): Christophe Rouillon, Gregor Hagelueken

Last updated by author(s): Nov 10, 2022

# Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

## Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

| n/a | Confirmed |
|---|---|
| ☐ | ☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement |
| ☐ | ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| ☒ | ☐ The statistical test(s) used AND whether they are one- or two-sided<br>*Only common tests should be described solely by name; describe more complex techniques in the Methods section.* |
| ☒ | ☐ A description of all covariates tested |
| ☒ | ☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| ☐ | ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| ☒ | ☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted<br>*Give P values as exact values whenever suitable.* |
| ☒ | ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| ☒ | ☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| ☒ | ☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated |

*Our web collection on statistics for biologists contains articles on many of the points above.*

## Software and code

Policy information about availability of computer code

| | |
|---|---|
| Data collection | Crystallography: XDS (v. 2020), SAXS: SASFLOW (v. 3.0). |
| Data analysis | phenix (1.19.1-4122-000), Coot (0.9), alphafold2(https://github.com/deepmind/alphafold; v 2.0.0), Colabfold (v.1.0-1.4), PyMOL (v 2.0), ASTRA 8 (Wyatt), ATSAS package (v. 3.1.3, contains all SAXS analysis software that was used in this study), Molprobity (http://molprobity.biochem.duke.edu/ v. 4.5.1), webflags server (http://www.webflags.se/ v. 2020), Biacore Insight Evaluation (v. 3.0.12.15655), DeerAnalysis 2022 including ComparativeDeerAnalyzer ( v. 2.0), mtsslWizard (Webserver version, http://www.mtsslsuite.isb.ukbonn.de/) |

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

## Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The crystal structures have been deposited in the PDB with the accession codes 7QDA [https://doi.org/10.2210/pdb7QDA/pdb], 8B0R [https://doi.org/10.2210/pdb8B0R/pdb], 8B0U [https://doi.org/10.2210/pdb8B0U/pdb]. The SAXS data and models have been deposited in the SASBDB with the accession codes: SASDQM4 [https://www.sasbdb.org/data/SASDQM4], SASDQN4 [https://www.sasbdb.org/data/SASDQN4], SASDQP4 [https://www.sasbdb.org/data/SASDQP4], SASDQQ4 [https://www.sasbdb.org/data/SASDQQ4]. The following PDB entries have been used in this study: 2H27 [https://doi.org/10.2210/pdb2H27/pdb], 3K1J [https://

doi.org/10.2210/pdb3K1J/pdb], 4ME7 [https://doi.org/10.2210/pdb4ME7/pdb], 4IZJ [https://doi.org/10.2210/pdb4IZJ/pdb], 4LUP [https://10.2210/pdb4LUP/pdb], 5ZX2 [https://doi.org/10.2210/pdb5ZX2/pdb], 5CR2 [https://doi.org/10.2210/pdb5CR2/pdb], 6VM6 [https://doi.org/10.2210/pdb6VM6/pdb], 6SCE [https://doi.org/10.2210/pdb6SCE/pdb], 7RWK [https://doi.org/10.2210/pdb7RWK/pdb]

# Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

# Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| | |
|---|---|
| Sample size | No sample size calculation was performed, because not needed for this biochemical data. |
| Data exclusions | No data were excluded from the analyses. |
| Replication | The key findings such as the protease activation by cA4 were repeated multiple times in different laboratories (Gregor Hagelueken and Malcolm F. White). The formation of the protein complexes was also independently shown in both laboratories. All attempts at replication were successful. The number and type of replications (biological or technical) are given for each experiment in the particular figure legend. No replication attempts were unsuccessful. |
| Randomization | Randomization was not relevant to this study. |
| Blinding | Blinding was not relevant to this study. |

# Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials & experimental systems

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ Antibodies |
| ☒ | ☐ Eukaryotic cell lines |
| ☒ | ☐ Palaeontology and archaeology |
| ☒ | ☐ Animals and other organisms |
| ☒ | ☐ Human research participants |
| ☒ | ☐ Clinical data |
| ☒ | ☐ Dual use research of concern |

## Methods

| n/a | Involved in the study |
|---|---|
| ☒ | ☐ ChIP-seq |
| ☒ | ☐ Flow cytometry |
| ☒ | ☐ MRI-based neuroimaging |