



<https://doi.org/10.1038/s42003-023-04654-z>

OPEN

## In vivo characterization of the bacterial intramembrane-cleaving protease RseP using the heme binding tag-based assay iCLiPSpy

Thomas Kupke <sup>1</sup>✉, Rabea M. Götz <sup>2</sup>, Florian M. Richter <sup>2</sup>, Rainer Beck <sup>1</sup>, Fabio Lolicato <sup>1,3</sup>, Walter Nickel <sup>1</sup>, Carsten Hopf <sup>2</sup> & Britta Brügger <sup>1</sup>✉

Regulated intramembrane proteolysis (RIP) describes the protease-dependent cleavage of transmembrane proteins within the hydrophobic core of cellular membranes. Intramembrane-cleaving proteases (I-CLiPs) that catalyze these reactions are found in all kingdoms of life and are involved in a wide range of cellular processes, including signaling and protein homeostasis. I-CLiPs are multispreading membrane proteins and represent challenging targets in structural and enzyme biology. Here we introduce iCLiPSpy, a simple assay to study I-CLiPs in vivo. To allow easy detection of enzyme activity, we developed a heme-binding reporter based on TNF $\alpha$  that changes color after I-CLiP-mediated proteolysis. Co-expression of the protease and reporter in *Escherichia coli* (*E. coli*) results in white or green colonies, depending on the activity of the protease. As a proof of concept, we use this assay to study the bacterial intramembrane-cleaving zinc metalloprotease RseP in vivo. iCLiPSpy expands the methodological repertoire for identifying residues important for substrate binding or activity of I-CLiPs and can in principle be adapted to a screening assay for the identification of inhibitors or activators of I-CLiPs, which is of great interest for proteases being explored as biomedical targets.

<sup>1</sup>Heidelberg University Biochemistry Center (BZH), Im Neuenheimer Feld 328, 69120 Heidelberg, Germany. <sup>2</sup>Center for Mass Spectrometry and Optical Spectroscopy (CeMOS), Mannheim University of Applied Sciences, Paul-Wittsack-Str. 10, 68163 Mannheim, Germany. <sup>3</sup>Department of Physics, University of Helsinki, Helsinki, Finland. ✉email: [thomas.kupke@bzh.uni-heidelberg.de](mailto:thomas.kupke@bzh.uni-heidelberg.de); [britta.bruegger@bzh.uni-heidelberg.de](mailto:britta.bruegger@bzh.uni-heidelberg.de)

I-CLiPs cleave transmembrane proteins within the lipid bilayer (reviewed in<sup>1</sup>). This type of proteolytic activity is conserved throughout all domains of life, and contributes to a wide range of cellular processes, including protein maturation and degradation, and cellular signaling. I-CLiPs can be divided into different groups depending on the type of catalysis. Although these proteases generally have low substrate specificity, some exclusively cleave type I transmembrane proteins, such as mammalian gamma-secretase<sup>2</sup>, which is associated with Alzheimer's disease, while other proteases target type II transmembrane proteins. Examples of the latter are members of the SPPL2 (signal peptide peptidase-like 2) family, whose best-known substrate is tumor necrosis factor alpha (TNF $\alpha$ ), a key regulator of the inflammatory response<sup>3</sup>. I-CLiPs are involved in highly coordinated RIP signaling events. In general, signal transduction by RIP begins with cleavage of the extracellular domain of transmembrane proteins in a process known as ectodomain shedding, and then the remaining membrane-bound fragment is processed by I-CLiPs<sup>4–6</sup>. A well-characterized I-CLiP in *E. coli* is the inner membrane zinc metalloprotease RseP<sup>7–9</sup> (formerly called YaeL or EcfE). RseP activates the cell envelope stress response in *E. coli* by cleaving the anti- $\sigma^E$  type II transmembrane protein RseA. This signaling cascade is initiated by the membrane-bound serine protease DegS<sup>10–12</sup>, which acts as a sheddase and removes the periplasmic domain of RseA. The remaining N-terminal fragment 1–148 of RseA is then cleaved by RseP between residues 108 and 109<sup>13</sup>.

RseP is a transmembrane protein with four transmembrane helices and two tandemly arranged periplasmic PDZ domains that serve as size exclusion filters for the substrate<sup>8,9,14–17</sup>. The Zn<sup>2+</sup> ion required for activity is bound by the side chains of His22, His26 and Asp402<sup>7,8,18</sup>. An important structural element involved in substrate recognition and cleavage is the membrane-embedded  $\beta$ -hairpin-like structure formed by residues Ile61 to Glu75 [membrane-reentrant  $\beta$ -loop,  $\beta$ -MRE]<sup>19</sup>. In addition, the GFG motif of RseP and a membrane-bound amphiphilic helix (residues Pro323 to Thr350) contribute to substrate discrimination and binding<sup>20,21</sup>. The recently deposited crystal structure of RseP with the inhibitor batimastat provides mechanistic insights into intramembrane proteolysis and suggests a gating mechanism to regulate substrate entry<sup>22</sup>.

RseP also cleaves signal peptides generated by leader peptidase-mediated processing of presecretory proteins<sup>9,23,24</sup>. This low substrate specificity prompted us to investigate whether reporter substrates could be used to study RseP in vitro and in vivo. Ideally, such a reporter would change its spectroscopic characteristics upon proteolytic processing and also contain a tag that would allow affinity purification and detection.

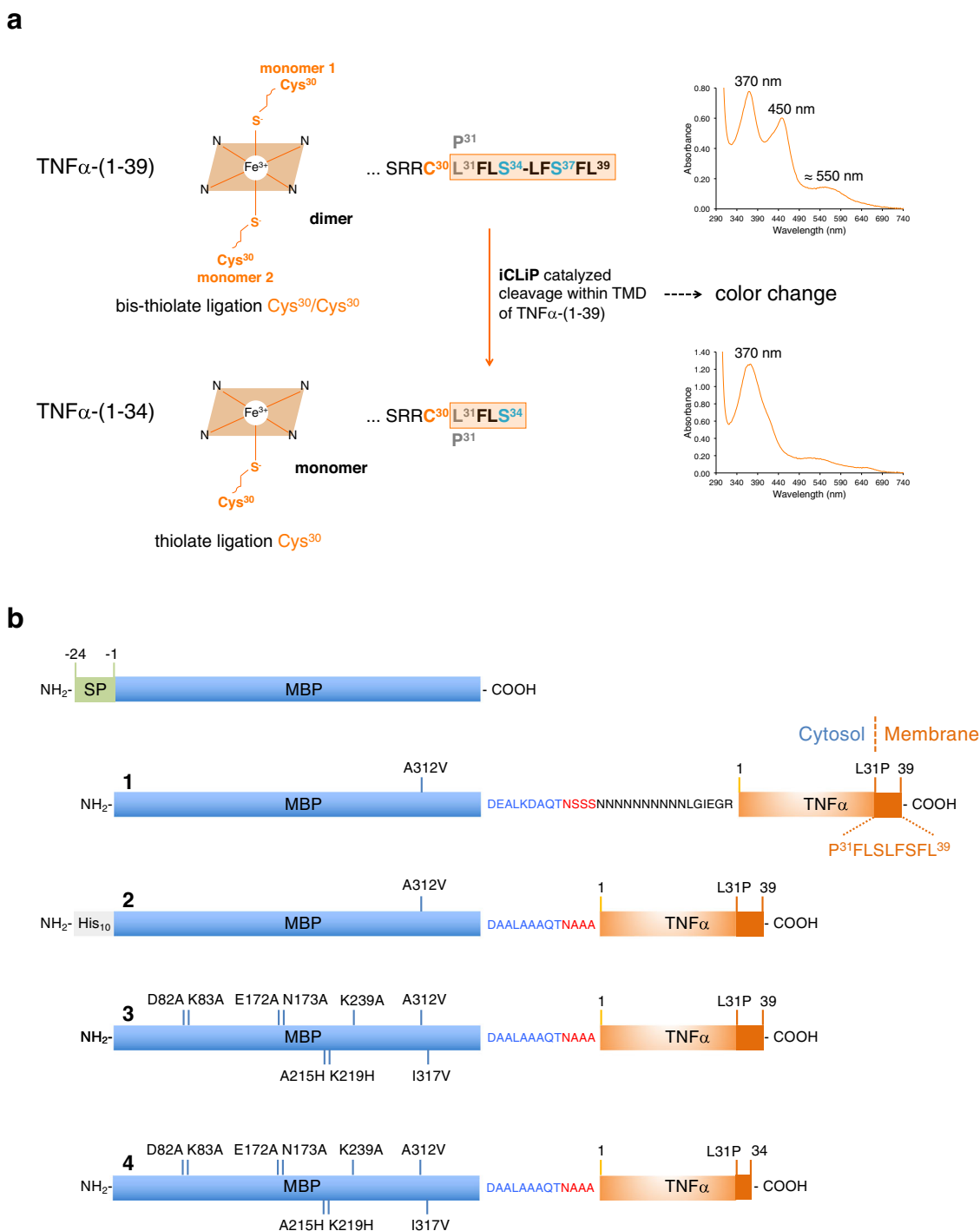
In a recent study, we used an *E. coli* expression model to identify TNF $\alpha$ , its proteolytic processing products and other substrates of SPPL2a/b proteases as heme (= heme B) binding proteins<sup>25</sup>. In vivo, TNF $\alpha$  is cleaved by the sheddase TACE, releasing an N-terminal fragment (NTF) and a soluble cytokine<sup>26</sup>. The NTF is further processed by the intramembrane-cleaving proteases SPPL2a/b, resulting in two intracellular domains (ICDs) that differ in length [TNF $\alpha$ -(1-39) and TNF $\alpha$ -(1-34)]<sup>27</sup>. We have described that heme binding depends on both a cytosolic cysteine residue proximal to the membrane (Cys30) and oligomerization of the transmembrane domain (TMD). Proteolytic processing of the N-terminal fragment of TNF $\alpha$  alters the nature of ferric heme binding, switching from stable, hexa-coordinated ligation by two Cys30 side chains in dimeric TNF $\alpha$ -(1-39) (so-called bis-thiolate ligation) to low affinity, penta-coordinated heme bound by only one Cys30 side chain in the shorter, monomeric ICD TNF $\alpha$ -(1-34)<sup>25</sup> (Fig. 1a). Deletion of five of the nine TMD residues still

present in TNF $\alpha$ -(1-39) (residues Leu35 to Leu39) removes the dimerization motif S<sup>34</sup>XXS<sup>37</sup> and thus prevents heme promoted dimerization of TNF $\alpha$ -(1-34) (Fig. 1a). Bis-thiolate ligated ferric heme shows a split Soret band with absorbance maxima at about 370 nm and 450 nm and an additional absorbance maximum at about 550 nm, so that TNF $\alpha$ -(1-39) with bound heme is a green colored peptide. In contrast, TNF $\alpha$ -(1-34) has only one Soret band at about 370 nm and appears uncolored due to its low heme binding affinity<sup>25</sup> (Fig. 1a).

In the present study, we exploit the different spectroscopic properties of substrate TNF $\alpha$ -(1-39) and product TNF $\alpha$ -(1-34) to investigate intramembrane processing catalyzed by RseP in vivo. We show that *E. coli* cells co-producing an MBP-TNF $\alpha$ -(1-39) fusion protein (MBP, maltose binding protein) optimized for iClIPSpy and an active RseP protease are uncolored/white, while cells expressing the inactive RseP enzyme together with the reporter are green. To confirm the processing of TNF $\alpha$ -(1-39) catalyzed by RseP, we used mass spectrometry. By determining the intact mass and additionally the sequence of tryptic fragments of the processed substrate proteins, we show that RseP trims tripeptides and cleaves L/F-S peptide bonds within the substrate used. To test whether this assay for I-CLiPs allows the identification of residues within RseP that are important for activity and/or substrate binding, we generated mutant *rseP* genes by error-prone PCR<sup>28</sup>. Using this approach, we identified Tyr69 as a critical residue for RseP activity and/or substrate binding. Molecular dynamics (MD) simulations of the AlphaFold<sup>29,30</sup> structure of RseP (with bound Zn<sup>2+</sup> ion) in a lipid bilayer with a lipid composition mimicking *E. coli* membranes show that residue Tyr69 is located in the membrane/cytosol environment near the active site and may play a role in positioning the substrate peptide through hydrophobic interactions. These results are in agreement with the recently published crystal structure of RseP and the model proposed for substrate binding, in which the so-called edge strand of RseP (residues Gly67 to Val70 including Tyr69 and part of the  $\beta$ -MRE) forms an antiparallel  $\beta$ -sheet with the substrate TMD<sup>22</sup>. Our data show that iClIPSpy is a simple and sensitive assay to identify and study I-CLiP activities in vivo.

## Results

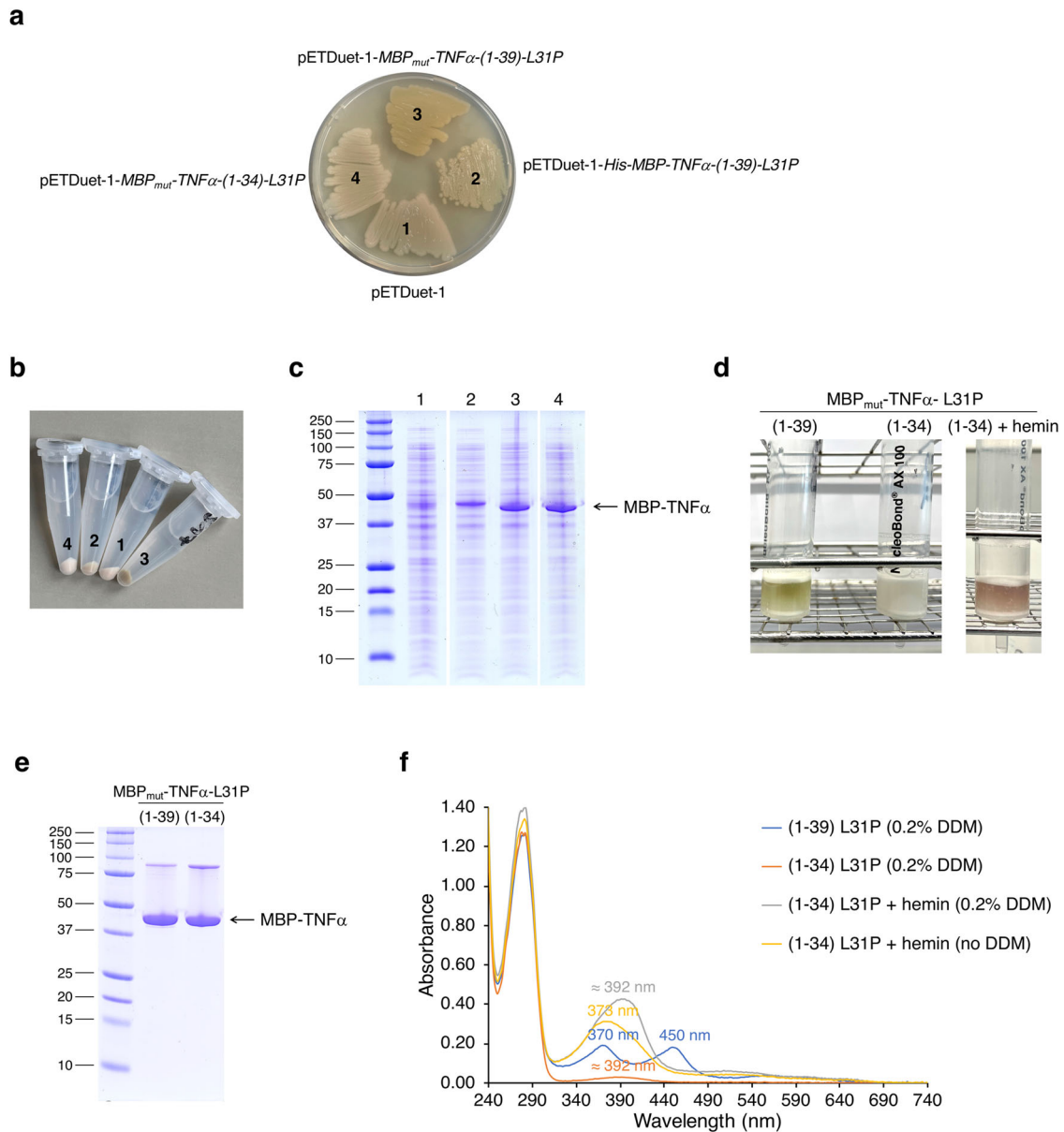
**Substrate and product of the reaction catalyzed by RseP.** An effective production strategy for the heme-binding peptide TNF $\alpha$ -(1-39) as potential substrate of RseP and other I-CLiPs, is its expression as MBP fusion protein in *E. coli*<sup>25</sup>. To ensure the correct topology [considering that full-length TNF $\alpha$  is a type II transmembrane protein: N-terminus *in* (cytoplasm) and C-terminus *out* (periplasm/lumen of the ER)], MBP was fused to the N-terminus of TNF $\alpha$ -(1-39) without the signal peptide (residues -24 to -1, Fig. 1b). TNF $\alpha$ -(1-39) contains only the N-terminal nine amino acid residues L<sup>31</sup>FLSLFSFL<sup>39</sup> of the TNF $\alpha$  TMD and is therefore found in both the membrane and cytosolic fractions<sup>25,27</sup>. Instead of the MBP construct described in our recent study<sup>25</sup>, we used a mutant MBP protein (MBP<sub>mut</sub>) containing additional substitutions, including I317V, which further increases the affinity for amylose<sup>31</sup> (Fig. 1b). To increase the amount of heme bound by TNF $\alpha$ , we used the variant peptide TNF $\alpha$ -(1-39)-L31P, which shows a stronger interaction with heme compared to the wildtype peptide<sup>25</sup>. In addition, MBP<sub>mut</sub> and TNF $\alpha$ -(1-39)-L31P are linked in MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P by a shortened linker peptide (Fig. 1b), which should increase the proteolytic stability of the fusion protein. We expressed the putative product of RseP-mediated intramembrane proteolysis, TNF $\alpha$ -(1-34), in the same manner [MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P]. The MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P, MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P and His-MBP-TNF $\alpha$ -(1-39)-L31P genes (Fig. 1b) were cloned into the



**Fig. 1 The principle of the heme binding tag-based iCLIPSpY assay.** **a** Bis-thiolate ligation of heme depends on dimerization of TNF $\alpha$ -(1-39). Cleavage within the TMD part L31-L39 of TNF $\alpha$ -(1-39) prevents dimerization by removing or disrupting the S<sup>34</sup>XXS<sup>37</sup> motif and results in mono-thiolate ligation of heme by Cys<sup>30</sup>. Bis-thiolate and mono-thiolate-coordinated heme molecules differ in their absorbance spectra, so that TNF $\alpha$ -(1-39) (substrate for intramembrane-cleaving protease) and TNF $\alpha$ -(1-34) (product of intramembrane proteolysis) with bound heme show different colors. Because of increased heme binding affinity, we use TNF $\alpha$ -(1-39)-L31P instead of TNF $\alpha$ -(1-39) wt as heme binding tag. **b** MBP-TNF $\alpha$  constructs used in our studies [lacking the signal peptide (SP) of *E. coli* MBP shown at the top of the figure]. **1**, MBP-TNF $\alpha$ -(1-39)-L31P with a poly-Asn linker was used in our initial study to characterize heme binding of TNF $\alpha$ <sup>25</sup>. **2-4** MBP-TNF $\alpha$  fusions used in the present study all contain the peptide DAALAAQTNAAA, which links MBP to TNF $\alpha$ -(1-39). The topology of MBP-TNF $\alpha$ -(1-39) fusion proteins is exemplary shown for construct 1. For co-production with I-CLIP RseP, construct **3** [named MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P] was used, which contains additional eight substitutions that further increase affinity for amylose, reduce the surface entropy and increase the crystal-packing interactions between adjacent MBP molecules<sup>31</sup>.

pETDuet-1 vector (Supplementary Table 1 and Supplementary Note 1) and the recombinant plasmids were transformed into *E. coli* T7 Express cells. For in vivo experiments, expression of MBP fusions was performed without IPTG induction on agar

plates (Fig. 2a). A general disadvantage of the T7 expression system is the already existing low protein expression without IPTG induction, also called leaky expression. However, leaky expression can be an advantage, especially in the expression of



**Fig. 2** Characterization of substrate and product of the iClIPspy assay. **a** *E. coli* T7 Express cells harboring the plasmids **1**, pETDuet-1 (control), **2**, pETDuet-1-*His-MBP-TNF $\alpha$ -(1-39)-L31P* (expressing a substrate for I-CLiPs), **3**, pETDuet-1-*MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* (expressing an alternative substrate for I-CLiPs) or **4**, pETDuet-1-*MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P* (expressing the putative product of intramembrane proteolysis), were incubated overnight on an agar plate at 37 °C followed by another day at room temperature. **b** *E. coli* T7 Express cells harboring the plasmids **1**, pETDuet-1, **2**, pETDuet-1-*His-MBP-TNF $\alpha$ -(1-39)-L31P*, **3**, pETDuet-1-*MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* or **4**, pETDuet-1-*MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P*, were grown on an agar plate for six days at room temperature, scraped from the plate, resuspended in TN buffer, and harvested by centrifugation and **c** then further analyzed by SDS-PAGE (after solubilization by boiling in Laemmli sample buffer). **d** MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P were purified by amylose affinity chromatography from the membrane fraction using DDM as solubilizing detergent (left and middle columns) from IPTG induced *E. coli* T7 Express cells harboring the above-listed plasmids. Purification of MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P was repeated after addition of 25  $\mu$ M hemin (right column) to membrane proteins solubilized with DDM. Purified MBP fusion proteins were then analyzed by **e** SDS PAGE and **f** UV/Vis spectroscopy [plus 0.2% DDM: blue trace, MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P; red trace, MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P; grey trace, MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P plus hemin; minus DDM: yellow trace, MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P plus hemin]. When DDM is removed from hemin reconstituted MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P by washing the amylose column with buffer containing no DDM, the typical absorbance maximum of about 373 nm for mono-thiolate ligated heme is observed<sup>25,56</sup>, showing that presence of DDM leads to a bathochromic shift of the absorbance maximum from 373 nm to about 392 nm for mono-thiolate ligated heme. Free hemin in the elution buffer used with 0.2% DDM has an absorbance maximum at 405 nm. In vivo detection of MBP-TNF $\alpha$  fusion proteins, as shown in **a**, was repeated once (and in a different way in Fig. 3 a and Supplementary Fig. 2a), and one representative result is shown. Detection of MBP-TNF $\alpha$  fusion proteins in cells by SDS-PAGE (**b** and **c**) was done once, and the UV/Vis spectra of purified MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P fusion proteins (**f**), were checked in a second experiment.

membrane proteins, as the low-level expression prevents the formation of inclusion bodies<sup>32,33</sup> and also favors heme-binding of TNF $\alpha$ -(1-39)-L31P, as the heme concentration in cells is extremely low<sup>34</sup>.

After several days at room temperature cells expressing MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P were ochre green, while cells expressing His-MBP-TNF $\alpha$ -(1-39)-L31P were only faintly colored and cells expressing MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P or the empty vector control

pETDuet-1 were white (Fig. 2a, b). SDS-PAGE analysis of whole cells scraped from agar plates confirmed the production of the different MBP-TNF $\alpha$  fusion proteins and showed that reduced expression of *His-MBP-TNF $\alpha$ -(1-39)-L31P* is responsible for the less intense color of these cells (Fig. 2c). Therefore, we selected the *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P*, and *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P* constructs for further experiments. The membrane-associated MBP fusion proteins were then solubilized with the non-ionic detergent n-dodecyl- $\beta$ -D-maltopyranoside (DDM) and purified by amylose affinity chromatography in the presence of DDM. UV/Vis spectroscopy of the purified fusion proteins confirmed the ligation of heme by bis-thiolate in *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* and by thiolate in *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P* (Fig. 2d-f). *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P* with thiolate ligated heme shows an absorbance maximum at about 392 nm, as the addition of DDM results in a bathochromic shift of the absorbance maximum by  $\sim$ 20 nm (Fig. 2f). Cells expressing *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P* are white because the heme content is very low. Heme reconstitution of *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P*, however, resulted in a red-brown color of the fusion protein (Fig. 2d, f). Taken together, our results show that recombinantly expressed reporter substrate and mimicked product of I-CLiP-mediated proteolysis in *E. coli* differ in their spectroscopic properties, resulting in visible changes in the coloring of bacterial colonies. Next, we tested whether our reporter *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* also serves as a substrate for bacterial I-CLiPs.

**Visualization of the in vivo activity of RseP.** A well-characterized bacterial I-CLiP is *E. coli* RseP, which we used to develop the I-CLiP assay iCLiPSpy. A previously published in vivo assay for monitoring RseP activity<sup>9,23</sup> uses *E. coli rseP rseA* double-disrupted cells carrying plasmids expressing RseP-Myc<sup>8</sup> and an HA-MBP-tagged, truncated version of RseA, RseA140, as a substrate for RseP. RseA140 cleavage by RseP was monitored by immunoblot analysis<sup>9,23</sup> and this assay has been successfully used to study the structure, function and substrate binding of RseP by site-directed mutagenesis<sup>19–21,35</sup>. Since the in vivo activity of RseP is not directly visible when looking at colonies of *E. coli* cells over-expressing RseP-Myc and HA-MBP-RseA140, we sought to develop an assay that allows simple detection of I-CLiP activity based on the color of the colonies.

To test whether *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* can be cleaved by RseP, RseP-Myc<sup>8</sup> wild type (wt, Supplementary Fig. 1a) or the catalytically inactive variant RseP-Myc H22F were expressed together with our I-CLiP reporter. Since RseP is essential for bacterial growth<sup>7</sup>, we performed iCLiPSpy in presence of the chromosomal copy of *rseP*. In principle, the RseP-dependence of growth can be circumvented by down-regulating the porins OmpA and OmpC<sup>36</sup>. Our data, however, show that the amount of chromosomally encoded RseP is not sufficient to degrade overproduced *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* to any appreciable extent (Fig. 3a), so that additional genetic modification was not required.

Co-expression of the I-CLiP reporter with RseP-Myc wt resulted in white cells (Fig. 3a), indicating that the overexpressed RseP had cleaved *TNF $\alpha$ -(1-39)-L31P*. The H22F substitution in RseP, which leads to a loss of zinc binding, inhibited RseP activity, so that co-expression of *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* with *rseP-Myc H22F* resulted in green *E. coli* colonies (Fig. 3a). SDS-PAGE analysis of whole cells scraped from the agar plate confirmed comparable expression levels of the *MBP<sub>mut</sub>-TNF $\alpha$*  fusions and showed that the RseP cleavage product of *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* has a similar apparent molecular weight as *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P*, indicating that only a few amino acid residues were cleaved off by RseP (Fig. 3b). Expression of

*rseP-Myc wt* and *rseP-Myc H22F* genes was confirmed by western blotting (Fig. 3c). Higher expression levels of *rseP-Myc H22F* compared to *rseP-Myc wt* were less pronounced in case of co-expression with *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* (compare the first two lanes with lanes five and six of the immunoblot shown in Fig. 3c). As expected and despite this higher expression, proteolytic cleavage of *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* by RseP-Myc H22F was not observed. In vivo processing of *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* was also confirmed for untagged RseP (Supplementary Figs. 1a-c and 2a, b). However, when a 10x HisTag (His<sub>10</sub>) was fused directly without a linker peptide to the carboxy terminus of RseP, the activity of RseP was decreased (Supplementary Fig. 1a-c), suggesting that the interaction of the periplasmic His<sub>10</sub> Tag with the PDZ domains (Supplementary Fig. 1b) could affect folding, stability or activity of RseP.

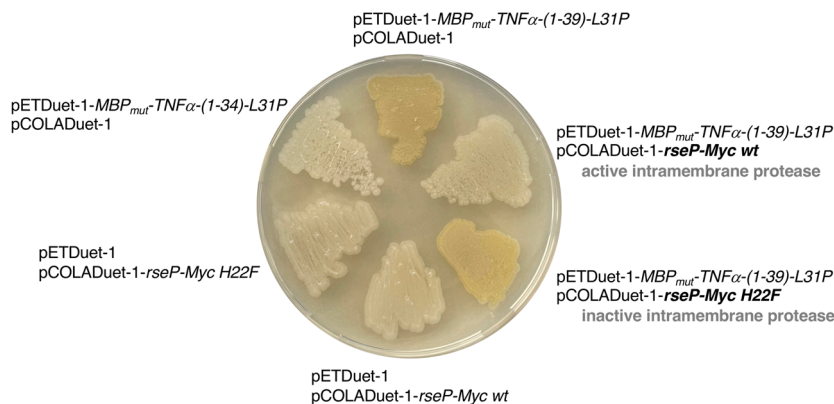
In general, our results suggest that green-colored colonies appear when sufficient *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* is present. White colonies, however, are observed when *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* has been cleaved or when the reporter is expressed at very low concentrations. Therefore, an expression rate of the I-CLiP reporter that is independent of the sequence of the co-expressed *rseP* gene (Fig. 3b and Supplementary Fig. 2b) is an essential requirement for this assay.

#### Processing changes heme binding properties of the substrate.

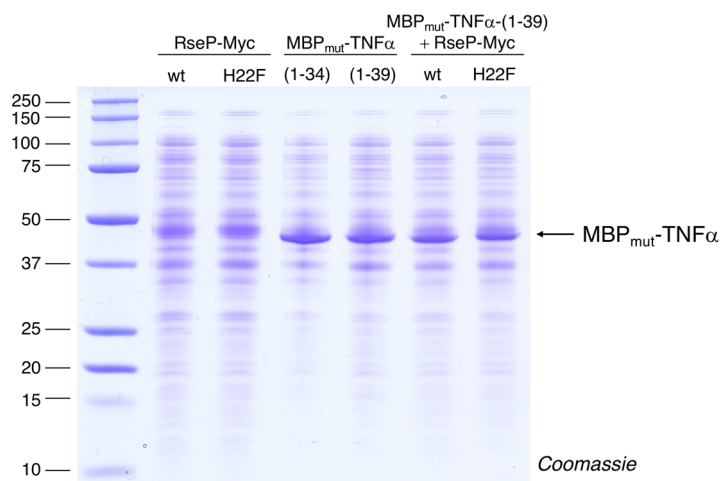
Our data suggest that intramembrane proteolysis catalyzed by RseP affects the dimerization motif S<sup>34</sup>LFS<sup>37</sup> of *TNF $\alpha$ -(1-39)* and consequently alters the heme binding mode of the *TNF $\alpha$  peptide*, consistent with the color change. To test whether *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* indeed serves as a substrate for RseP cleavage, membrane-associated proteins solubilized with DDM from IPTG-induced *E. coli* cells co-expressing *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* with either *rseP wt* or *rseP H22F* were enriched via amylose affinity chromatography (Fig. 4a), analyzed by SDS-PAGE (Fig. 4b) and the heme coordination mode was determined by UV/Vis spectroscopy (Fig. 4c). The ratio of active and correctly folded protease to substrate may differ after IPTG-induced overexpression compared to leaky expression in the in vivo assay. This would lead to different proportions of fully processed substrate. Indeed, we observed that both amylose affinity columns were green, but the color was more intense in the case of co-expression with *rseP H22F* (Fig. 4a). *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* proteins purified from *rseP wt* over-expressing cells showed a slightly lower apparent molecular weight than *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* proteins purified from *rseP H22F* over-expressing cells (Fig. 4b). The UV/Vis spectrum of *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* purified from *rseP H22F* over-expressing cells showed absorbance maxima at 371 nm and 451 nm, demonstrating bis-thiolate ligation and confirming that RseP H22F does not cleave *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P*. For *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* purified from *rseP wt* over-expressing cells, however, the absorbance at 451 nm was significantly reduced, the absorbance maximum at 371 nm was shifted to 373 nm and the absorbance at 392 nm was increased [Fig. 4c, see also Fig. 2f, which shows a UV/Vis absorbance spectrum of the putative RseP cleavage product *MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P*, which would be expected if proteolysis was complete and the dimerization motif disrupted]. As expected, unprocessed and RseP-processed *MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P* fusion proteins were also present in the cytosolic fraction (Supplementary Fig. 3a-e; compare also Kupke et al.<sup>25</sup>).

Solubilized membrane-associated *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* proteins purified from *E. coli* co-overexpressing either inactive *rseP H22F* or active *rseP wt* were then separated by gel filtration in presence of DDM into heme-binding (monomeric or dimeric) *MBP<sub>mut</sub>-TNF $\alpha$ -L31P* and heme-free monomeric *MBP<sub>mut</sub>-TNF $\alpha$ -*

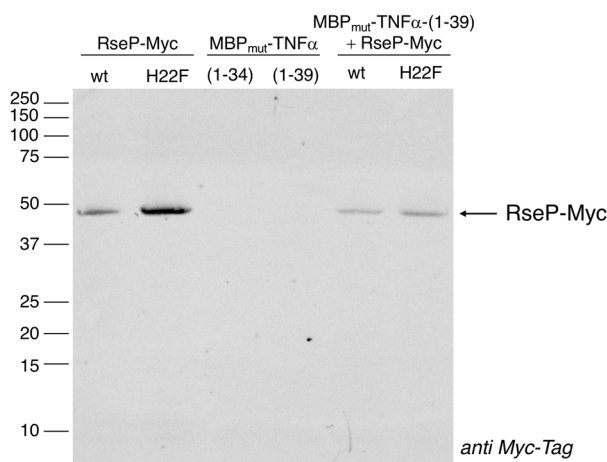
**a**



**b**



**c**



L31P fusion proteins (Fig. 4d). Heme-free monomers are always present as heme is not stoichiometrically bound by TNF $\alpha$ -ICDS<sup>25</sup>. The higher content of heme-free monomers of MBP<sub>mut</sub>-TNF $\alpha$ -L31P in the case of RseP wt compared to RseP H22F is probably related to a lower heme binding affinity of the processed fusion protein (compare UV/Vis spectra in Fig. 4c), because hydrophobic residues within the heme-binding motif SRRCPFLSFLSFL

of TNF $\alpha$  may also contribute to heme binding and because bis-thiolate ligation stabilizes heme binding<sup>25</sup>. The heme-binding MBP<sub>mut</sub>-TNF $\alpha$ -L31P fusion protein obtained in the case of RseP H22F eluted at 15.40 ml, whereas the protein observed for RseP wt eluted at 16.05 ml (corresponding to a lower apparent molecular weight) from the size exclusion column (Fig. 4d). Since MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P contains only a single cysteine

**Fig. 3 Visualization of in vivo intramembrane proteolysis catalyzed by RseP-Myc.** **a** *E. coli* T7 Express cells harboring the two compatible plasmids pETDuet-1 and pCOLADuet-1, which co-express genes encoding the heme-binding protein MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and the active RseP-Myc wt or inactive RseP-Myc H22F protease, were incubated for five days at room temperature on an agar plate. Control cells expressed only one gene encoding either RseP-Myc wt, RseP-Myc H22F, the substrate MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P or the putative proteolysis product MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P. Green *E. coli* colonies are only observed if the substrate present, MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P, is not C-terminally truncated by RseP. **b** and **c** Analysis of *E. coli* T7 Express cells expressing genes for RseP-Myc wt, RseP-Myc H22F, MBP<sub>mut</sub>-TNF $\alpha$ -(1-34)-L31P, MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P or coexpressing genes for MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and RseP-Myc wt/H22F, respectively, and grown on a selection plate for four days at room temperature by **b**, SDS-PAGE, and **c**, by immunoblotting using a monoclonal anti-Myc-Tag antibody. Comparative in vivo analysis of RseP activities as shown in **a** and **b** was repeated with untagged RseP proteases with the same result (compare Supplementary Fig. 2).

residue [Cys30 of TNF $\alpha$ -(1-39)-L31P,] bis-thiolate ligation of heme requires dimerization<sup>25</sup> (Fig. 1a). Therefore, the eluted proteins were again characterized by UV/Vis spectroscopy to distinguish between monomeric and dimeric heme-binding proteins (Fig. 4e). The UV/Vis spectra of the heme-binding proteins purified by gel filtration confirm that dimeric MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P is present when MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was co-expressed with *rseP H22F* but not with *rseP wt*. In the latter case, the absorbance maximum at about 392 nm (measured in the presence of DDM) shows that processed MBP<sub>mut</sub>-TNF $\alpha$ -L31P coordinates heme *via* a single cysteine residue and is monomeric. It appears that RseP-processed (monomeric) heme binding MBP<sub>mut</sub>-TNF $\alpha$ -L31P is not uniform, as the observed elution volume is as mentioned either 16.05 ml (when measuring the absorbance at 280 nm) or 15.85 ml (absorbance at 371 nm; Fig. 4d). We assume that at least two different processing products with different heme binding affinities (and thus different heme content) are present. At higher protein concentrations, a processing product that still contains part of the dimerization motif may dimerize to a small extent and therefore also shows an absorbance maximum at about 450 nm (and 373 nm; compare spectrum “+ RseP wt” of Fig. 4e with the corresponding spectrum of Fig. 4c).

In summary, gel filtration shows that cleavage by RseP results in MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins that differ in molecular weight and heme binding properties from the substrate MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P.

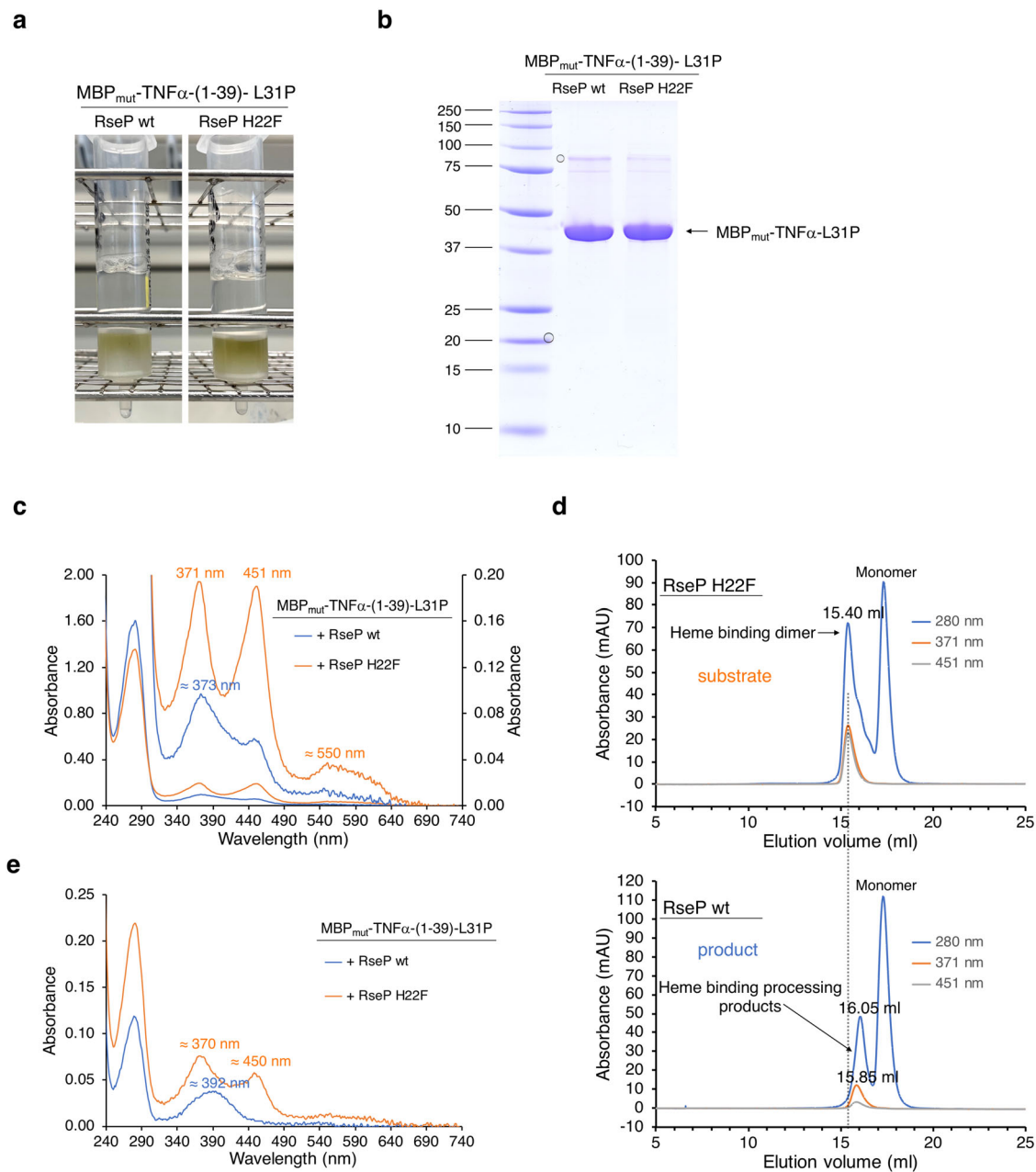
**RseP processing sites in the MBP<sub>mut</sub>-TNF $\alpha$  substrate.** The UV/Vis absorbance spectrum of the MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P fusion protein purified from cells co-overexpressing *rseP wt* shows that the heme-coordinating residue Cys30 is still present in the proteolytic cleavage product, but that the bis-thiolate ligation of heme has been largely lost. Consequently, cleavage occurred within TNF $\alpha$  TMD residues 31-39 of the reporter protein, disrupting the dimerization motif (compare Fig. 1). To determine the exact RseP processing site(s) within the C-terminal TNF $\alpha$  TMD residues P<sup>31</sup>FLSLFSFL<sup>39</sup> by mass spectrometry, MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was again co-expressed with either *rseP wt*, empty vector control pCOLADuet-1, *rseP-Myc wt* or *rseP-Myc H22F*. The MBP-TNF $\alpha$  fusion proteins produced were purified from the membrane fraction by amylose affinity chromatography and then analyzed by SDS-PAGE (Fig. 5a, b) and UV/Vis spectroscopy (Fig. 5c) to confirm the experimental data shown in Fig. 4 and to ensure RseP catalyzed processing within TNF $\alpha$ -(1-39)-L31P prior to mass spectrometric analysis. As previously observed, the MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins that still bind heme *via* bis-thiolate ligation have a slightly higher molecular weight than the fusion proteins that bind heme *via* mono-thiolate ligation (Fig. 5a, b), but the resolution of SDS-PAGE is too low to separate differently truncated proteins. Two complementary mass spectrometric analyses were then performed to determine the proteolytic cleavage sites: (i) analysis of the intact mass of the MBP-TNF $\alpha$  fusion proteins (Fig. 5d, e and Supplementary

Fig. 4a-d). and (ii) sequence analysis and quantification of their tryptic peptides (Supplementary Fig. 4e). For the first approach, MBP-TNF $\alpha$  fusion proteins obtained from the four different genetic backgrounds were separated by reversed-phase (RP)-liquid chromatography (LC) and the intact mass information was obtained by coupled electrospray ionization mass spectrometry (ESI-MS) and compared with the calculated average molecular weights of C-terminally truncated MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P fusion proteins (Fig. 5d, e). The detected major degradation product MBP<sub>mut</sub>-...CP<sup>31</sup> [lacking the terminal eight amino acid residues of TNF $\alpha$ -ICD-(1-39)] is not an RseP processing product, as it is also present in the absence of over-expressed active protease (Fig. 5e). Residues F<sup>32</sup>LSLFSFL<sup>39</sup> could be cleaved off by unspecific proteolytic processing (independent of RseP), as to our knowledge there is no *E. coli* K12 endopeptidase that specifically cleaves Pro-X peptide bonds. We assume that MBP<sub>mut</sub>-...CP<sup>31</sup> is detectable, despite the absence of TNF $\alpha$  TMD residues, because the membrane fraction has not been washed and therefore also contains soluble proteins (but also the hydrophobic heme cofactor bound to the CP motif might target this fusion protein to the membrane).

The full-length MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P fusion protein characterized by the C-terminal sequence ...C<sup>30</sup>PFLSLFSFL<sup>39</sup>, however, was identified both by its intact mass and by sequence analysis of tryptic peptides (see Supplementary Note 2), for “Empty vector control” and “RseP-Myc H22F” (inactive RseP protein) purifications (Fig. 5e, and Supplementary Fig. 4a, b, e). This full-length MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P protein is able to bind heme by bis-thiolate ligation, resulting in the characteristic absorbance maxima of 370 nm, 450 nm and 550 nm<sup>25,37,38</sup>.

However, when MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was co-expressed with *rseP wt* or *rseP-Myc wt*, MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P (= MBP<sub>mut</sub>-...CPFLSLFSFL<sup>39</sup>) was cleaved between residues F36 and S37 as well as between L33 and S34 of TNF $\alpha$ -(1-39)-L31P in both cases, resulting in the major processing products MBP<sub>mut</sub>-...CPFLSLF<sup>36</sup> and MBP<sub>mut</sub>-...CPFL<sup>33</sup> for both wt proteases (Fig. 5e, and Supplementary Fig. 4c-e). Upon co-expression with *rseP wt*, we found small amounts of three additional cleavage products, MBP<sub>mut</sub>-...CPF<sup>32</sup>, MBP<sub>mut</sub>-...CPFLS<sup>34</sup> (Fig. 5e, and Supplementary Fig. 4c) and MBP<sub>mut</sub>-...CPFLSLFS<sup>37</sup> (Supplementary Table 2). We do not know whether these truncated MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P proteins are also RseP processing products or whether the two major RseP processing products are further degraded by other *E. coli* proteases. MBP<sub>mut</sub>-...CPFL<sup>33</sup> can no longer bind heme by bis-thiolate ligation, since the loss of the C-terminal pentapeptide L<sup>35</sup>FSFL<sup>39</sup> already leads to mono-thiolate ligated heme<sup>25</sup> (Fig. 2). MBP<sub>mut</sub>-...CPFLSLF<sup>36</sup>, which lacks the second serine residue Ser37 important for (heme promoted) dimerization<sup>25</sup>, could still dimerize to a small extent and bind heme by bis-thiolate ligation (compare absorbance of the processing products of RseP at 450 nm, Fig. 4c).

Small amounts of the major processing product of RseP, MBP<sub>mut</sub>-...SRRCPFLSLF<sup>36</sup>, were also identified for the “Empty



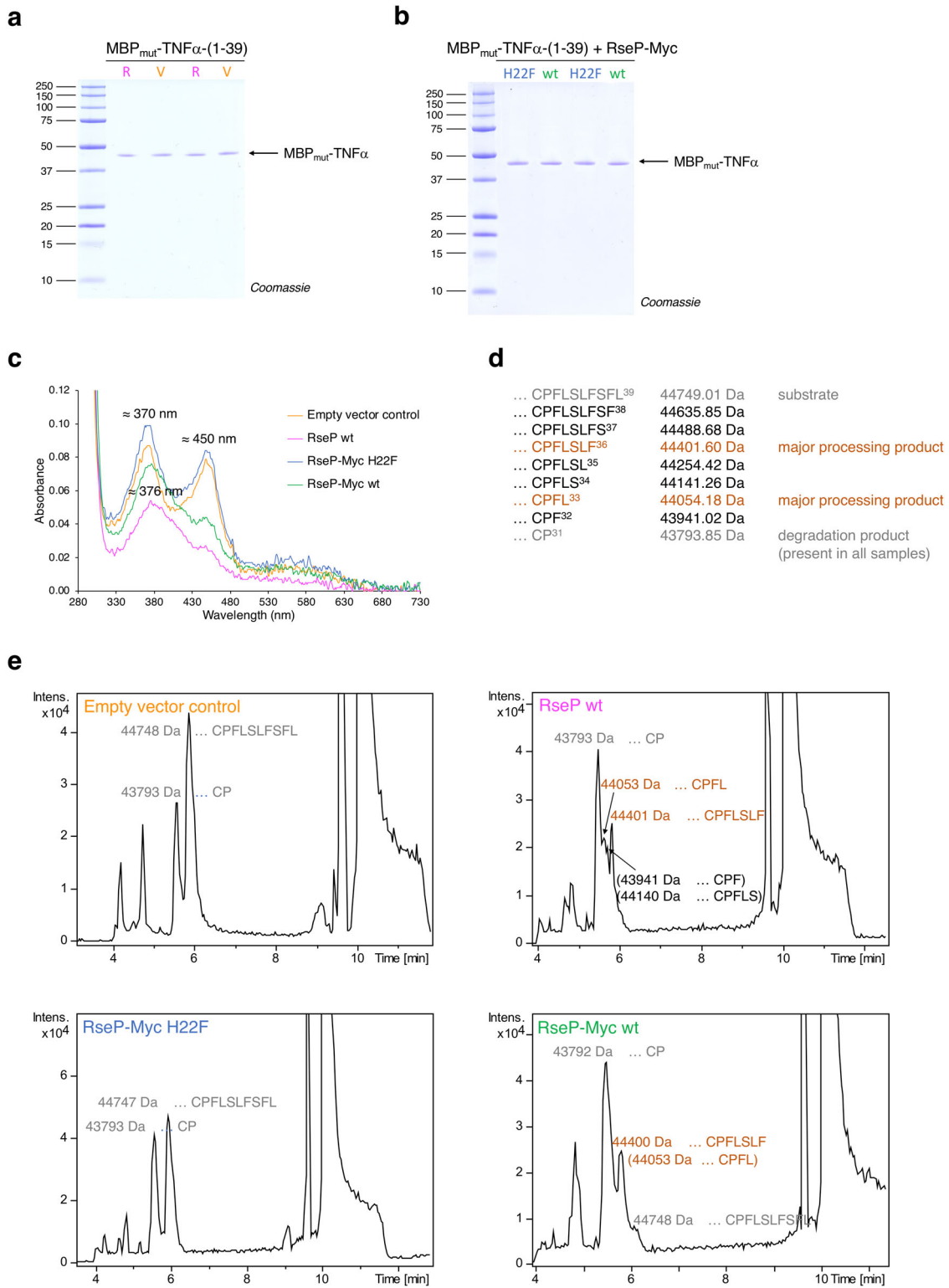
**Fig. 4 Characterization of the RseP cleavage product in vitro.** **a** Purification of membrane-associated MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins by amylose affinity chromatography from IPTG induced T7 cells co-expressing either genes for MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and RseP wt (left column) or genes for MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and mutant RseP H22F (right column). The purified MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins (which correspond to the product and substrate of the RseP reaction, respectively) were then analyzed **b** by SDS-PAGE, **c** by UV/Vis spectroscopy [blue trace, + RseP wt; orange trace, + RseP H22F; showing the full spectra (left absorbance axis) and the same spectra magnified (right absorbance axis)] and **d** by gel filtration on a Superose 6 Increase column [mAU = milli absorbance units; absorbance is monitored at 280 nm (blue trace), 371 nm (orange trace), and 451 nm (grey trace)]. As expected, the heme binding dimers and processing products (indicated by arrows) show different  $A_{371\text{nm}}/A_{451\text{nm}}$  absorbance ratios. It can be assumed that heme binding to monomeric TNF $\alpha$  processing products significantly alters their shape/structure and thus their elution volume and apparent molecular weight. **e** UV/Vis spectroscopy of gel filtration fractions containing heme binding MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins (blue trace, + RseP wt, elution volume 15.4 ml - 16.2 ml; orange trace, + RseP H22F, elution volume 15.4 ml - 15.8 ml of the gel filtration runs shown in Fig. d). Purification of MBP<sub>mut</sub>-TNF $\alpha$  proteins coproduced with either RseP wt or with RseP H22F (**a-c**) was performed four times and representative results are shown. **b-e** belong to the same experiment. Gel filtration (**d** and **e**) of the purified MBP<sub>mut</sub> fusion proteins was performed once.

vector control” but not for “RseP-Myc H22F” by LC-MS/MS analysis (Supplementary Table 2). It is likely that minor cleavage of the substrate protein by chromosomally encoded RseP (and other *E. coli* proteases) occurs, which can be detected by MS/MS after tryptic digest of purified MBP fusion proteins. This processing of MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was not detected in

intact mass analysis (shown in Fig. 5e and Supplementary Fig. 4a).

In summary, mass spectrometric analysis confirms that RseP cleaves the TNF $\alpha$  substrate within the TMD residues 31-39. The different heme binding modes of substrate and processing product(s) determined by UV/Vis spectroscopy are demonstrably





related to C-terminal truncation of the substrate. The mass spectrometric analysis suggests a processive cleavage mechanism of RseP and also gives insights into the specificity of RseP.

**Combining error-prone mutagenesis of RseP with iCliPSpy.** To identify residues in RseP that are critical for protease activity and/or substrate binding, we combined the *in vivo* assay iCliPSpy with error-prone mutagenesis of the *rseP* gene (Fig. 6a, b). We selected seven green *E. coli* clones co-producing an inactivated RseP

enzyme and the substrate MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P for sequencing of the *rseP* gene. In six out of seven clones, we found premature stop codons and frameshift mutations in addition to several amino acid exchanges (Fig. 6c), and therefore excluded them from further analysis. In contrast, the mutant designated RseP G1 contained eleven non-synonymous mutations, but neither premature stop codons nor frameshift mutations. We mapped the eleven amino acid exchanges onto the AlphaFold wt structure of RseP and found that the Y69H substitution is located

**Fig. 5 Determination of RseP cleavage sites.** **a** and **b** Membrane-associated MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins from IPTG-induced *E. coli* T7 Express cells were purified by amylose affinity chromatography and analyzed by SDS-PAGE. **a** MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was co-expressed with pCOLADuet-1-rseP wt (R) or the empty vector control pCOLADuet-1 (V), or **b** with pCOLADuet-1-rseP-Myc wt (wt) or pCOLADuet-1-rseP-Myc H22F (H22F). Comparable amounts of purified proteins were run twice and alternately on SDS-PAGE to detect very small differences in molecular weight. **c** UV/Vis spectra of MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins purified from these different genetic backgrounds (orange trace, empty vector control; magenta trace, + RseP wt; blue trace, + RseP-Myc H22F; green trace, RseP-Myc wt). **d** Calculated average molecular weights of C-terminally truncated MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P fusion proteins (only the C-terminal residues of the MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins are shown here and in **e**). **e** Purified, putative unprocessed (left panels) and processed (right panels) MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins (compare **a** and **b**) were analyzed by LC-MS. The average molecular weights of the eluted MBP<sub>mut</sub>-TNF $\alpha$  fusion proteins obtained by deconvolution of the ESI-MS spectra (compare Supplementary Fig. 4a-d) are given. The substrate MBP<sub>mut</sub>-... CPFLSLF<sup>39</sup> and the degradation product MBP<sub>mut</sub>-... CP<sup>31</sup> (found in all samples) are shown in grey, RseP major processing products MBP<sub>mut</sub>-... CPFLSLF<sup>36</sup> and MBP<sub>mut</sub>-... CPFL<sup>33</sup> in brown and additional processing products in black font. For MBP<sub>mut</sub>-TNF $\alpha$  preparations from “Empty vector control” and “RseP-Myc H22F”, we found a protein eluting at 9.1 min with an average molecular weight of about 45448 Da min, which was higher than expected for a MBP<sub>mut</sub>-TNF $\alpha$  fusion protein and was therefore not further characterized. Compounds eluting at retention times above 9.5 min (corresponding to 90% acetonitrile) were contaminants not further characterized. **a** to **e** belong to one experiment, and the processing specificity of RseP wt was confirmed in this experiment by analysis of RseP-Myc wt. The C-terminal sequences of the substrate MBP<sub>mut</sub>-... CPFLSLF<sup>39</sup> fusion protein and processing product MBP<sub>mut</sub>-... CPFLSLF<sup>36</sup> were confirmed by LC-MS/MS analysis (Supplementary Fig. 4e).

within the edge of the  $\beta$ -MRE at a position close to the active-site within the substrate binding groove (Fig. 6d). The edge structure was shown to be important for both protease activity and substrate binding<sup>19,20</sup>. For example, the Y69P substitution in RseP significantly impairs the cleavage of the model substrate HA-MBP-RseA148<sup>19</sup>. However, the introduction of a Pro residue can significantly alter the structure of the  $\beta$ -MRE<sup>19,20</sup>. To further investigate the role of Tyr69 in protease function without altering the overall structure of the  $\beta$ -MRE, we tested whether the single substitution Y69H, in which Tyr69 is exchanged for a more hydrophilic histidine residue, reduces RseP activity *in vivo*. Assuming that the color of colonies on agar plates reflects the proteolytic activity of RseP, we found a decreased *in vivo* activity of RseP Y69H compared to RseP wt (Fig. 6e), which was more pronounced for RseP G1 and the active-site variant RseP H22F (Fig. 6e). Next, we focused on a more detailed characterization of residues Tyr69 and Tyr428 by MD simulations and site-directed mutagenesis. Residue Tyr428 of TMD 4 was included in our studies because it is localized on the opposite side of Tyr69 at the putative substrate binding groove, and Tyr69 and Tyr428 together may work in gating substrates<sup>22</sup> into the binding groove (Fig. 6d).

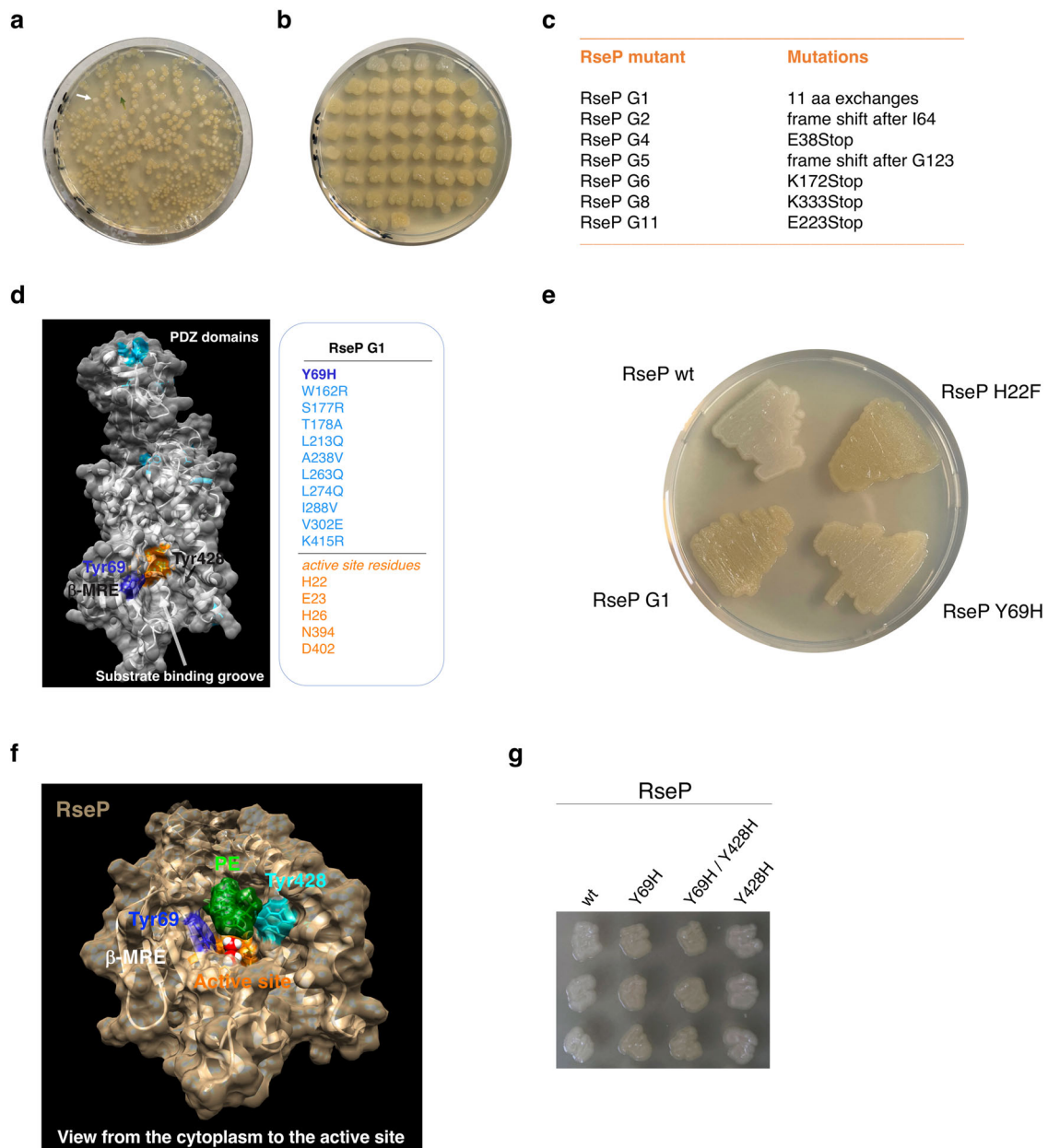
**MD simulations of membrane-embedded RseP.** To further investigate the role of Tyr69 in the function of RseP, we performed MD simulations of RseP based on the structure predicted by AlphaFold<sup>29,30</sup> (with a Zn<sup>2+</sup> ion manually added to the active site), as the pdb file 7W6X for the crystal structure of full-length RseP<sup>22</sup> was not publicly available. To mimic the native environment, RseP was embedded in an artificial membrane resembling the lipid composition of *E. coli* membranes (see Supplementary Table 3 and Supplementary Note 3) and MD simulations (Fig. 6f and Supplementary Fig. 5a-e) were performed on a time scale of 2 microseconds. The structures obtained showed only minor changes in the RseP active site-containing membrane core compared to the AlphaFold prediction used for the simulations (and compared to the RseP crystal structure). However, the MD simulations indicated that the periplasmic tandem PDZ domain is flexible (Supplementary Fig. 5e). The MD simulations of membrane-embedded RseP predict that Tyr69 and Tyr428 are localized at the cytosol/membrane interface adjacent to the active site of RseP as part of the substrate binding groove (Fig. 6f and Supplementary Fig. 5). The physiological importance of a lipid molecule placed by MD simulation next to/in the substrate binding groove, however, is less clear. In the case of the MD simulation shown in Fig. 6f and Supplementary Fig. 5 this lipid was phosphatidylethanolamine (PE), in other simulations we found cardiolipin, suggesting that overall hydrophobicity is the

major determinant for insertion into the substrate binding pocket of RseP.

In our *in vivo* assay (Fig. 6g), the Y428H substitution had no significant impact on the proteolytic activity of RseP. Furthermore, the double substitution Y69H/Y428H was indistinguishable from the single substitution Y69H (Fig. 6g), suggesting that Tyr428 neither plays a role in binding the substrate MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P nor is important for catalytic activity. To support this result, MBP fusion proteins purified from *E. coli* cells co-overexpressing MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P with wt or mutant *rseP* genes were analyzed by SDS-PAGE (Supplementary Fig. 6a) and UV/Vis spectroscopy (Supplementary Fig. 6b). When MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P was co-expressed with *rseP* Y69H or with *rseP* Y69H/Y428H, heme remained bound to the reporter *via* bis-thiolate ligation (Supplementary Fig. 6b), confirming that intramembrane proteolysis by RseP is impaired when Tyr69 is exchanged for His. However, RseP Y428H behaved like the wt enzyme, processing the C-terminal residues of the reporter substrate (compare Fig. 5) and shifting heme binding towards a monomeric, thiolate-bound form.

## Discussion

Recently, we observed that the N-terminal domain of TNF $\alpha$  comprising residues 1-39, TNF $\alpha$ -(1-39), dimerizes and binds heme *via* two cysteine residues, whereas the proteolytically processed peptide TNF $\alpha$ -(1-34) binds heme *via* one cysteine residue, as the dimerization motif S<sup>34</sup>XXS<sup>37</sup> within the TMD residues 31-39 of TNF $\alpha$  is disrupted<sup>25</sup>. In the present study, we exploit these spectroscopic properties of TNF $\alpha$  to develop a simple chromophore-based *in vivo* assay to monitor I-CliP activity. We chose *E. coli* RseP as a benchmark enzyme, which is well characterized by both *in vitro* and *in vivo* assays<sup>23</sup>. Our *in vivo* studies show that co-expression of TNF $\alpha$ -(1-39)-L31P with active *rseP* wt results in white *E. coli* colonies, while co-expression with inactive RseP H22F results in green colonies. After purification of the MBP-tagged reporter protein we confirmed by UV/Vis spectroscopy and mass spectrometry that TNF $\alpha$ -(1-39)-L31P is indeed cleaved by RseP. Mass spectrometric analysis showed that RseP engages in the trimming of MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P at L/F-S peptide bonds and releases the tripeptides S<sup>37</sup>FL<sup>39</sup> and S<sup>34</sup>LF<sup>36</sup>, whereas the native substrate RseA is cleaved by RseP between Ala108 and Cys109<sup>13</sup>. However, the low substrate specificity of RseP and other I-CLiPs is well known<sup>23,35</sup>. Interestingly, tripeptide trimming is also observed for the processing of amyloid  $\beta$ -peptide 49 by  $\gamma$ -secretase, and in this case the trimming mechanism has recently been studied in detail using a novel peptide Gaussian accelerated molecular dynamics method<sup>39</sup>. In



**Fig. 6 Error-prone mutagenesis of *rseP* in combination with the iClIPSpy assay.** **a** Mutant *rseP* genes were co-transformed with pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P. The transformation plates were incubated overnight at 37 °C and then at room temperature for two days (as an example, a white colony is indicated by a white arrow and an ochre-green clone by a green arrow). **b** White (presumably active RseP proteases; first row of the plate) and green clones (inactive RseP) were transferred to a new plate and incubated again for one day at room temperature. **c** From seven green clones expressing inactive RseP proteases, the plasmids pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and pCOLADuet-1-*rseP* were purified together and the sequence of mutant *rseP* genes was determined. **d** The eleven non-synonymous amino acid substitutions in RseP G1 mutant were mapped to the AlphaFold structure of RseP wt (substituted amino acid residues are shown in light blue). Tyr69 (Y69, dark blue) is the only mutated residue located near the predicted active site (shown in orange);  $\beta$ -MRE, membrane-reentrant  $\beta$ -loop. **e** The Y69H substitution found in RseP G1 was introduced as a single mutation in RseP and analyzed together with RseP wt, RseP H22F or RseP G1 using the iClIPSpy assay. Cells were incubated at room temperature for four days (overnight incubation at 37 °C followed by incubation at room temperature resulted in less colored RseP Y69H cells). **f** MD simulations of the AlphaFold RseP wt structure (with Zn<sup>2+</sup> manually curated into the active site) in a membrane environment visualized with Chimera. View from the cytoplasm to the active site [composed of His22, His26 and Asp402 complexing the Zn<sup>2+</sup> ion (shown in black, but mostly hidden by two water molecules) and the residues Glu23 and Asn394] shown in orange. Tyr69 at the cytosol/membrane interface is shown in dark blue, Tyr428 in cyan;  $\beta$ -MRE, membrane-reentrant  $\beta$ -loop. Membrane lipids have been omitted (shown in Supplementary Fig. 5c), except for 1-hexadecanoyl-2-(9Z-hexadecenoyl)-sn-glycero-3-phosphoethanolamine (PYPE, in green), which is adjacent to/inside the binding groove. **g** Comparison of the in vivo activities of RseP wt, RseP Y69H, RseP Y69H/Y428H and RseP Y428H using the iClIPSpy assay. Cells shown in **g** had been incubated at room temperature for two days. Error-prone mutagenesis of *rseP* was performed once, in vivo analysis of RseP Y69H activity (**e**) was performed twice. In vivo assay for RseP wt, RseP Y69H, RseP Y69H/Y428H and RseP Y428H was repeated once with the same result as shown in **g**.

the future, MD simulations may also help to determine the mechanism of RseP more precisely and provide better insight into the mechanism of I-CLiPs in general.

The crystal structures of the tandem PDZ domain of RseP<sup>16</sup> and of the transmembrane core domain of an S2P metalloprotease from *Methanocaldococcus jannaschii*<sup>40</sup> led to an AlphaFold structure prediction<sup>29,30</sup> with very high confidence for RseP. We used this prediction for mapping mutations (Fig. 6d) and for MD simulations of RseP (Fig. 6f and Supplementary Fig. 5), because the pdb file for the crystal structure of RseP<sup>22</sup> was only available as we were working on the revised version of the manuscript but not in the beginning of our studies. Later, it also turned out that the AlphaFold structure prediction for the membrane core of RseP is very similar to the published crystal structure of RseP (Supplementary Fig. 5f), so that we did not repeat the MD simulations using the crystal structural data. Thus, our approach led to an overall correct picture of the structure of membrane-embedded RseP without knowledge of the crystal structure. Although the physiological importance of a lipid molecule placed next to/into the substrate binding groove by MD is unclear, this may be an important discovery as the lipid has to be exchanged for the substrate protein during catalysis. Future work will show whether this bound lipid is part of a lipid/substrate exchange during RseP-mediated cleavage of proteins.

Using the iCLiPSpy *in vivo* assay, we identified the inactive RseP G1 mutant with eleven non-synonymous mutations, including the substitution Y69H, by error-prone mutagenesis. We focused on characterizing the single mutant RseP Y69H, which has not been previously described. Although the AlphaFold structures of RseP wt and RseP Y69H are nearly identical except for the Y69H exchange, we found decreased proteolytic cleavage of the substrate TNF $\alpha$ -(1-39) for RseP Y69H. The finding of decreased activity of RseP Y69H is consistent with the importance of the edge strand (residues Gly67 to Val70) for substrate binding by RseP, as it had been previously shown that proline substitutions of the edge residues Gly67, Gly68, Tyr69, and Val70 perturb the  $\beta$ -strand conformation of the  $\beta$ -MRE and abolish the proteolytic activity of RseP<sup>19,20,22</sup>. It is tempting to speculate that Tyr69 engages in a hydrophobic contact to a side chain of the unwound substrate peptide within the substrate binding groove and that the Y69H exchange reduces the affinity for the substrate. The crystal structure of RseP now shows that the bound inhibitor batimastat is in close contact to the side chains of Phe44 (which is part of the GFG motif<sup>20</sup>) and Tyr69. How the other mutations within the G1 mutant contribute to the impairment of RseP activity is less clear.

When the photoreactive hydrophobic amino acid, p-benzoyl-L-phenylalanine (pBPA) was incorporated at position 69, 71 or 74 of RseP (e.g. RseP Y69pBPA), RseP retained its proteolytic activity and the substrate RseA was photo-crosslinked to RseP, leading to a model in which the  $\beta$ -MRE binds to substrate TMDs and stabilizes their extended conformation<sup>19,20,22</sup>. In the recently proposed mechanism for substrate accommodation and cleavage in RseP, it is assumed that RseP is in equilibrium between a gate-open and gate-closed conformation, that the substrate is bound by the edge residues in the gate-open conformation, and that in a final step the transmembrane segment of the bound substrate is extended by the edge strand to promote proteolytic cleavage<sup>22</sup>.

Although the obtained results were largely confirmative, our studies convincingly demonstrate the power of the combination of error-prone mutagenesis with the iCLiPSpy assay for the characterization of RseP and possibly other I-CLiPs at the molecular level.

It is not known what fraction of the substrate protein MBP<sup>mut</sup>-TNF $\alpha$ -(1-39)-L31P forms a dimeric complex with heme and what proportion is monomeric, since we neither know the heme-

binding constant nor the equilibrium constant for (heme-promoted) dimerization of TNF $\alpha$ -(1-39). Also, the dimerization mechanism itself is not elucidated, but we favor a model in which heme is first bound by monomeric TNF $\alpha$ -(1-39) and then this protein-heme complex binds a second TNF $\alpha$ -(1-39) monomer assisted by the SXXS dimerization motif. We assume (considering the sizes of the peptide binding groove and the active-site) that monomeric TNF $\alpha$ -(1-39) (without ligated heme) and not the dimeric (heme binding) one is finally cleaved by RseP. This does not exclude that initially present dimeric TNF $\alpha$ -(1-39) is bound by RseP and that beta-strand addition dissociates the dimer before cleavage. RseP-catalyzed cleavage of the dimerization motif within monomeric TNF $\alpha$ -(1-39) will result in decreased amounts of peptide TNF $\alpha$ -(1-39) available for (heme promoted) dimerization and increased amounts of cleaved peptides binding heme by only one Cys residue.

In summary, we present iCLiPSpy, a chromophore-based assay for *in vivo* analysis of intramembrane proteolysis. In addition, combining iCLiPSpy with error-prone mutagenesis can help in targeted protein design, for example to increase protease activity. In principle, iCLiPSpy can be extended to other I-CLiPs (that cleave type II transmembrane proteins) as well as to other protein substrates of the THOMAS (Thiolate-Heme Oligomeric type II trans-Membrane proteins with heme binding mode adjusted by SPPL2a/b) protein family<sup>25</sup>. For example, CD74-NTF also binds heme *via* bis-thiolate ligation<sup>25</sup> and can be used to expand the repertoire of substrates to study I-CLiPs with different substrate specificities. The investigation of other I-CLiPs, such as SPPL2a/b proteases, using iCLiPSpy depends on their expression in *E. coli* and moreover, *in vivo* detection of heme binding substrate proteins such as CD74-NTF is only possible at high expression levels (changing the color of cells). Further research and development is needed to transfer this approach to an eukaryotic expression system such as yeast. In combination with the over-expression of I-CLiP reporters such as TNF $\alpha$ -(1-39)-L31P or CD74-NTF, green yeast colonies would turn into colorless colonies due to the activity of I-CLiPs. Such a system would be adaptable for high-throughput screening approaches aimed at discovering inhibitors or activators of I-CLiPs, which is of great interest for proteases being explored as biomedical targets.

In pathogenic bacteria, site-2 proteases play an important role in sensing host signals and regulating the expression of virulence genes during infection<sup>41</sup>. In *Vibrio cholerae*, for example the I-CLiP YaeL (a homolog of RseP) is involved in the inactivation of the membrane localized transcription factor TcpP *via* a two-step RIP pathway. TcpP activates transcription of *toxT*, which encodes the direct activator of toxin and pilus genes<sup>42</sup>. In *Streptococcus pneumoniae*, RseP cleaves the competence protein ComM to control its quantity. ComM protects *S. pneumoniae* from CbpD-mediated lysis during competence<sup>43,44</sup>. The iCLiPSpy assay can provide a simple yet powerful tool for the identification/detection and characterization of these and other I-CLiPs from pathogenic bacteria and opens up the possibility of searching for small molecule modulators of protease activity in an *in vivo* model system.

## Methods

**Expression plasmids.** MBP<sup>mut</sup>-TNF $\alpha$ -(1-39)-L31P, MBP<sup>mut</sup>-TNF $\alpha$ -(1-34)-L31P, His-MBP-TNF $\alpha$ -(1-39)-L31P and *E. coli* *rseP* wt, *rseP*-Myc wt and mutant genes, respectively, optimized for expression in *E. coli* were synthesized by Invitrogen GeneArt (Thermo Fisher Scientific). MBP<sup>mut</sup> contains substitutions increasing the affinity for amylose (A312V, I317V), reducing the surface entropy (D82A, K83A, E172A, N173A, and K239A) and increasing the crystal-packing interactions between neighboring MBP molecules (A215H/K219H)<sup>31</sup>. MBP only contains the substitution A312V (numbering of amino acid residues refers to the sequence of periplasmic MBP lacking the signal peptide). To minimize differences in expression rates of *rseP* wt, *rseP* H22F, *rseP* Y69H, *rseP* Y69H/Y428H, *rseP* Y428H and *rseP*-

*Myc wt*, *rseP-Myc H22F* genes due to different codon usage, the sequences only differ in the exchanged codons. Cloning into *NcoI/XhoI* sites of expression vectors pETDuet-1 (ColE1 replicon, ampicillin resistance) and pCOLADuet-1 (COLA replicon, kanamycin resistance; compatible with pETDuet-1), respectively, was performed using standard molecular biology techniques. The DNA sequence of all cloned genes was confirmed using appropriate primers. DNA sequences of all constructed plasmids (listed in Supplementary Table 1) are given in Supplementary Note 1, so that the exact amino acid sequences of all proteins used in this study are also available. Due to the cloning procedure the second residue (Leu; codon CTG) of all RseP variants is exchanged for Val (codon GTG). RseP L2V and RseP L2V H22F (and so on) are referred to as RseP wt and RseP H22F, respectively, in this paper. Transformation and expression were performed with NEB (New England Biolabs) 5a and T7 Express *E. coli* cells, respectively.

**In vivo cleavage of MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P by RseP.** T7 Express *E. coli* cells co-transformed with pTK1130 [= pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P] and pCOLADuet-1 encoded *rseP* genes were streaked from bacterial glycerol stocks onto selection plates [LB agar substituted with 2 g glucose l<sup>-1</sup>, ampicillin (100  $\mu$ g ml<sup>-1</sup>) and kanamycin (50  $\mu$ g ml<sup>-1</sup>), but containing no  $\delta$ -aminolevulinic acid hydrochloride<sup>45</sup>], incubated overnight at room temperature or 37 °C. Photos of the plates were taken after incubation for several additional days at room temperature. Expression of MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and *rseP* occurred because the T7 expression system is leaky.

**Purification of membrane-associated MBP fusion proteins.** As recently described, we used the MBP tag for amylose affinity purification of the MBP-TNF $\alpha$  fusions in combination with *n*-dodecyl- $\beta$ -D-maltopyranoside (DDM, Avanti) as detergent for membrane protein solubilization and subsequent chromatography<sup>25</sup>. *E. coli* T7 Express *E. coli* cells transformed with the appropriate MBP-TNF $\alpha$  encoding pETDuet-1 plasmids (see Supplementary Table 1) were grown in LB medium substituted with 2 g glucose l<sup>-1</sup> and ampicillin (100  $\mu$ g ml<sup>-1</sup>) at 37 °C. Co-expression experiments of pTK1130 [= pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P] and pCOLADuet-1 encoded *rseP* genes, were also performed in *E. coli* T7 Express cells. Cells were grown in this case in LB medium substituted with 2 g glucose l<sup>-1</sup>, ampicillin (100  $\mu$ g ml<sup>-1</sup>) and kanamycin (50  $\mu$ g ml<sup>-1</sup>). For both kinds of experiments, over-expression was induced with 50  $\mu$ M isopropyl 1-thio- $\beta$ -D-galactopyranoside (IPTG, Sigma-Aldrich) at an A<sub>600</sub> of about 0.5, and cultures were then further shaken at 37 °C for about 2 h. Cells from a one liter culture were harvested, resuspended in 30 ml TN buffer (50 mM Tris/HCl pH 7.4 at room temperature and 150 mM NaCl) containing 1 mM phenylmethylsulfonyl fluoride (PMSF, Sigma-Aldrich), frozen in liquid nitrogen and stored at -80 °C. For the experiments shown in Figs. 2d-f, 4, and Supplementary Fig. 3, 0.6 mM  $\delta$ -aminolevulinic acid hydrochloride was added before induction with IPTG<sup>45</sup>. However,  $\delta$ -aminolevulinic acid hydrochloride was not used in all other experiments. Thawed cells were again substituted with 1 mM PMSF and then disrupted by sonication. The membrane pellet obtained after 30 to 45 min centrifugation at 40,000x g and 4 °C was resuspended in 15 ml TN buffer containing 2% of the detergent DDM. Then, membrane proteins were solubilized at 4 °C for about 30 to 45 min in a 50 ml tube using a tube roller mixer. The supernatant obtained after centrifugation for 30 min at 40,000x g and 4 °C was subjected to amylose (NEB) affinity chromatography at room temperature. Amylose affinity columns were equilibrated and washed with TN buffer containing 0.2% DDM. MBP fusion proteins were eluted with 10 mM maltose (Sigma-Aldrich) in equilibration buffer and UV/Vis spectra of the eluted protein solutions were recorded at room temperature in a 1 cm Hellma quartz absorption cuvette using the NanoDrop 2000c spectrophotometer from Thermo Fisher Scientific with elution buffer as blank<sup>25</sup>.

**Gel filtration.** For gel filtration experiments, eluted MBP fusion proteins were concentrated using Vivaspin 500 columns at 4 °C and then stored at 4 °C. Gel filtration was carried out as described<sup>25</sup> at 4 °C using an ÄKTA pure system from GE Healthcare and a Superose 6 Increase 10/300 GL column for high-resolution size exclusion chromatography. 100  $\mu$ l of protein aliquots were subjected to the column equilibrated in TN buffer containing 0.2% DDM at a flow rate of 0.4 ml min<sup>-1</sup>. The elution was followed by absorbance at 280 nm and simultaneously for the detection of heme at additional wavelengths (such as 371 nm and 451 nm), and 0.4 ml fractions were collected. The void volume of the used column was determined with Blue dextran 2000 (GE Healthcare) to be 8.5 ml.

**Mass spectrometric analysis of RseP processing products.** Intact masses of MBP-TNF $\alpha$  fusion proteins were determined using reversed-phase (RP) ultra high-performance liquid chromatography (UHPLC) coupled to electrospray ionization mass spectrometry (RP-ESI-MS). Briefly, RP separation was performed on a 1290 UHPLC (Agilent Technologies) equipped with a PLRP-S column (1.0  $\times$  50 mm, 5  $\mu$ m, 1000 Å; Agilent Technologies) at 80 °C and 0.35 ml/min. Solvent A was 0.1% formic acid (FA) in water. Solvent B was 0.1% FA in acetonitrile. Total run time was 17 min and the gradient was as follows. 0–0.5 min: 20% B, 0.5–8.5 min: linear increase from 20% to 50% B, 8.5–9.0 min: linear increase from 50% to 90% B, 9.0–11.0 min: 90% B, 11.0–11.5 min: linear decrease from 90% to 20% B, and 11.5–17 min: re-equilibration at 20% B. MS analysis was performed (min 4 to 13) in

positive polarity scan mode on an Impact II ESI-QqTOF mass spectrometer (Bruker Daltonics), in a mass range from *m/z* 600 to 7000. Following ESI source parameters were used: end plate offset 500 V, capillary voltage 4500 V, nebulizer gas 3.0 bar, dry gas 12.0 l/min, dry temperature 300 °C. High mass calibration was performed using cesium fluoroheptanoate in enhanced quadratic calibration mode. Mass spectra were processed in DataAnalysis software version 4.4 (Bruker Daltonics). In brief, the protein is getting detected with different charge states in parallel, that are then deconvoluted to obtain the molecular weight of the protein. Masses were determined using maximum entropy charge deconvolution in a range of 10,000 – 50,000 Da. To confirm RseP cleavage specificity in a complementary analysis, tryptic peptides of MBP-TNF $\alpha$  fusion proteins were sequenced by LC-MS/MS (for a detailed setup of this peptide analysis see Supplementary Note 2). Since no extracted ion chromatograms (XIC) of the intact mass analysis could be obtained due to the nature of the analysis, quantification was performed by LC-MS/MS on the peptide level (compare Supplementary Table 2).

**Error-prone PCR.** Error-prone PCR<sup>28</sup> was performed using the JBS (Jena Bioscience) Error-prone kit with plasmid pTK1121 (= pColA-Duet-1-*rseP* wt) as template for the *rseP* gene and using the forward and reverse PCR primers TK154 (5'-AACCTTAATAAGGAGATATACCATGG-3') and TK155 (5'-GGTTTCTTT ACCGACTCGAGTTATTA-3') (*NcoI* and *XhoI* restriction sites are underlined). The rate of mutagenesis of this method is in the range of 0.6–2.0% and the mutations are randomly distributed throughout the amplified PCR product<sup>28</sup>. The PCR product was purified using the Qiagen PCR purification kit, double digested with *NcoI* and *XhoI* and then isolated by agarose gel electrophoresis using the Qiagen gel extraction kit. The *NcoI/XhoI rseP (error prone)* PCR fragment was then ligated with *NcoI/XhoI* double-digested pCOLA-Duet-1 and the ligation mixture was transformed into *E. coli* NEB5a cells. An aliquot of the transformation mixture was used to inoculate 50 ml LB + glucose medium. Plasmids of cells grown overnight at 37 °C were then purified using the Nucleobond PC100 Midi kit. The purified plasmid pColA-Duet-1-*rseP (error prone)* was subsequently co-transformed with pTK1130 [= pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P] into NEB T7 Express *E. coli* cells. The diluted transformation mixture was spread onto selection plates containing ampicillin and kanamycin, incubated overnight at 37 °C and then for two days at room temperature. White and green clones were then transferred to a fresh selection plate, incubated overnight at 37 °C and then for one additional day at room temperature. Plasmid DNA [pETDuet-1-MBP<sub>mut</sub>-TNF $\alpha$ -(1-39)-L31P and pColA-Duet-1-*rseP (error prone)*] of green clones was purified and the sequence of the mutated *rseP* genes was determined using the pCOLA-Duet-1 specific primers TK160 (5'-CTCCTGCATTAGGAAATTAATACG-3') and TK161 (5'-GTGCTTC TCAAATGCCTGAGG-3') and in addition by using several different *rseP* specific primers.

**SDS-PAGE and immunoblotting.** As described<sup>25</sup>, proteins were separated using Tricine-SDS-polyacrylamide (10%) gel electrophoresis under reducing (1x sample buffer containing 10 mM dithiothreitol, DTT) conditions<sup>46</sup>. Molecular weights of standard proteins (all blue prestained protein standard, Bio-Rad) are indicated on the side of the gel pictures. For documentation, destained gels were scanned using an Epson Perfection V850 Pro scanner with 1200 d.p.i. resolution. Proteins were electrophoretically transferred to PVDF membrane (Immobilion-FL, Merck Millipore) after SDS-PAGE by semi-dry blotting. Myc-tagged proteins were detected with a monoclonal anti-Myc Tag antibody (Sigma-Aldrich SABI305535, clone 9E10, diluted 1:500) as first antibody, an Alexa 680 goat-anti-mouse antibody (Life Technologies, Thermo Fischer Scientific, 0.8  $\mu$ g ml<sup>-1</sup>) as secondary antibody and then using the Odyssey Imaging System (Li-Cor Biosciences).

**AlphaFold and MD simulations.** We used the AlphaFold<sup>29,30</sup> predicted structure of *E. coli* K12 RseP (with a Val residue in position 2) for our MD simulations. Into the predicted structure we manually placed a Zn<sup>2+</sup> ion coordinated by the side chains of residues H22, H26 and D402. The modified structure was then embedded in silico into a described *E. coli* model membrane system containing about 75% cardiolipin and about 25% PE with multiple types of acyl chains<sup>47</sup> using the CHARMM GUI membrane builder interface<sup>48,49</sup>. Six MD simulations for membrane-embedded RseP wt, were performed each for 2  $\mu$ s using GROMACS<sup>50</sup> software and CHARMM36m forcefield<sup>51</sup> (for a detailed setup of the methodology compare Supplementary Table 3 and Supplementary Note 3). To choose the statistically most representative structures for illustration, the obtained trajectories were analyzed by clustering using the GROMOS method<sup>52</sup> applying an RMSD cutoff of 0.15 nm. All pdb files were visualized with UCSF Chimera<sup>53</sup>.

**Statistics and reproducibility.** No statistical analyses were performed. Wt and mutant proteins were always purified in parallel using the same purification procedure including the same buffers and detergent solutions (and using cells grown under the same conditions). The precise numbers of purifications are indicated in the figure legends.

**Reporting summary.** Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

All underlying data of this study are available in the main article and Supplementary Information. The source data behind Figs. 2f, 4c–e, 5c are provided as Supplementary Data 1 and the DNA sequences of all constructed plasmids are included in Supplementary Note 1. The uncropped version of Fig. 2c is included as Supplementary Fig. 7. The mass spectrometry proteomics data have been deposited with the ProteomeXchange Consortium via the PRIDE<sup>54</sup> partner repository with the dataset identifier PXD038785 and 10.6019/PXD038785. All other data are available from the corresponding authors upon reasonable request.

## Code availability

No custom or in-house software was used in the analysis of the MD simulations. AlphaFold2 model and output files as well as MD simulations parameter files are publicly available at Zenodo: <https://doi.org/10.5281/zenodo.7612846><sup>55</sup>.

Received: 7 June 2022; Accepted: 2 March 2023;

Published online: 18 March 2023

## References

- Kühnle, N., Dederer, V. & Lemberg, M. K. Intramembrane proteolysis at a glance: from signalling to protein degradation. *J. Cell. Sci.* **132**, jcs217745 (2019).
- Wolfe, M. S. Unraveling the complexity of  $\gamma$ -secretase. *Semin Cell Dev. Biol.* **105**, 3–11 (2020).
- Bradley, J. R. TNF-mediated inflammatory disease. *J. Pathol.* **214**, 149–160 (2008).
- 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).
- Lichtenthaler, S. F., Lemberg, M. K. & Fluhrer, R. Proteolytic ectodomain shedding of membrane proteins in mammals—hardware, concepts, and recent developments. *EMBO J.* **37**, e99456 (2018).
- McCarthy, A. J., Coleman-Vaughan, C. & McCarthy, J. V. Regulated intramembrane proteolysis: emergent role in cell signalling pathways. *Biochem. Soc. Trans.* **45**, 1185–1202 (2017).
- Dartigalongue, C., Lofler, H. & Raina, S. EcfE, a new essential inner membrane protease: its role in the regulation of heat shock response in *Escherichia coli*. *EMBO J.* **20**, 5908–5918 (2001).
- Kanehara, K., Akiyama, Y. & Ito, K. Characterization of the *yaeL* gene product and its S2P-protease motifs in *Escherichia coli*. *Gene* **281**, 71–79 (2001).
- Akiyama, Y., Kanehara, K. & Ito, K. RseP (YaeL), an *Escherichia coli* RIP protease, cleaves transmembrane sequences. *EMBO J.* **23**, 4434–4442 (2004).
- Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z. & Gross, C. A. DegS and YaeL participate sequentially in the cleavage of RseA to activate the  $\sigma^E$ -dependent extracytoplasmic stress response. *Genes Dev.* **16**, 2156–2168 (2002).
- Alba, B. M., Zhong, H. J., Pelayo, J. C. & Gross, C. A. *degS* (*hhoB*) is an essential *Escherichia coli* gene whose indispensable function is to provide  $\sigma^E$  activity. *Mol. Microbiol.* **40**, 1323–1333 (2001).
- Kanehara, K., Ito, K. & Akiyama, Y. YaeL (EcfE) activates the  $\sigma^E$  pathway of stress response through a site-2 cleavage of anti- $\sigma^E$ , RseA. *Genes Dev.* **16**, 2147–2155 (2002).
- Li, X. et al. Cleavage of RseA by RseP requires a carboxyl-terminal hydrophobic amino acid following DegS cleavage. *Proc. Natl Acad. Sci. USA* **106**, 14837–14842 (2009).
- Hizukuri, Y. & Akiyama, Y. PDZ domains of RseP are not essential for sequential cleavage of RseA or stress-induced  $\sigma^E$  activation in vivo. *Mol. Microbiol.* **86**, 1232–1245 (2012).
- Hizukuri, Y. et al. A structure-based model of substrate discrimination by a noncanonical PDZ tandem in the intramembrane-cleaving protease RseP. *Structure* **22**, 326–336 (2014).
- Inaba, K. et al. A pair of circularly permuted PDZ domains control RseP, the S2P family intramembrane protease of *Escherichia coli*. *J. Biol. Chem.* **283**, 35042–35052 (2008).
- Kanehara, K., Ito, K. & Akiyama, Y. YaeL proteolysis of RseA is controlled by the PDZ domain of YaeL and a Gln-rich region of RseA. *EMBO J.* **22**, 6389–6398 (2003).
- Koide, K., Maegawa, S., Ito, K. & Akiyama, Y. Environment of the active site region of RseP, an *Escherichia coli* regulated intramembrane proteolysis protease, assessed by site-directed cysteine alkylation. *J. Biol. Chem.* **282**, 4553–4560 (2007).
- Akiyama, K. et al. Roles of the membrane-reentrant beta-hairpin-like loop of RseP protease in selective substrate cleavage. *Elife* **4**, e08928 (2015).
- Akiyama, K., Hizukuri, Y. & Akiyama, Y. Involvement of a conserved GFG motif region in substrate binding by RseP, an *Escherichia coli* S2P protease. *Mol. Microbiol.* **104**, 737–751 (2017).
- Miyake, T., Hizukuri, Y. & Akiyama, Y. Involvement of a Membrane-Bound Amphiphilic Helix in Substrate Discrimination and Binding by an *Escherichia coli* S2P Peptidase RseP. *Front Microbiol.* **11**, 607381 (2020).
- Imaizumi, Y. et al. Mechanistic insights into intramembrane proteolysis by *E. coli* site-2 protease homolog RseP. *Sci. Adv.* **8**, eabp9011 (2022).
- Hizukuri, Y., Akiyama, K. & Akiyama, Y. Biochemical Characterization of Function and Structure of RseP, an *Escherichia coli* S2P Protease. *Methods Enzymol.* **584**, 1–33 (2017).
- Saito, A. et al. Post-liberation cleavage of signal peptides is catalyzed by the site-2 protease (S2P) in bacteria. *Proc. Natl Acad. Sci. USA* **108**, 13740–13745 (2011).
- Kupke, T., Klare, J. P. & Brügger, B. Heme binding of transmembrane signaling proteins undergoing regulated intramembrane proteolysis. *Commun. Biol.* **3**, 73 (2020).
- Horiuchi, T., Mitoma, H., Harashima, S., Tsukamoto, H. & Shimoda, T. Transmembrane TNF- $\alpha$  structure, function and interaction with anti-TNF agents. *Rheumatol. (Oxf.)* **49**, 1215–1228 (2010).
- Fluhrer, R. et al. A gamma-secretase-like intramembrane cleavage of TNF $\alpha$  by the GxGD aspartyl protease SPPL2b. *Nat. Cell Biol.* **8**, 894–896 (2006).
- Cadwell, R. C. & Joyce, G. F. Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**, 28–33 (1992).
- Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
- Varadi, M. et al. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444 (2022).
- Bokhove, M. et al. Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins. *J. Struct. Biol.* **194**, 1–7 (2016).
- Briand, L. et al. A self-inducible heterologous protein expression system in *Escherichia coli*. *Sci. Rep.* **6**, 33037 (2016).
- Zhang, Z. et al. High-level production of membrane proteins in *E. coli* BL21(DE3) by omitting the inducer IPTG. *Micro. Cell Fact.* **14**, 142 (2015).
- Hanna, D. A. et al. Heme dynamics and trafficking factors revealed by genetically encoded fluorescent heme sensors. *Proc. Natl Acad. Sci. USA* **113**, 7539–7544 (2016).
- Koide, K., Ito, K. & Akiyama, Y. Substrate recognition and binding by RseP, an *Escherichia coli* intramembrane protease. *J. Biol. Chem.* **283**, 9562–9570 (2008).
- Douchin, V., Bohn, C. & Boulloc, P. Down-regulation of porins by a small RNA bypasses the essentiality of the regulated intramembrane proteolysis protease RseP in *Escherichia coli*. *J. Biol. Chem.* **281**, 12253–12259 (2006).
- Barr, I. et al. DiGeorge critical region 8 (DGCR8) is a double-cysteine-ligated heme protein. *J. Biol. Chem.* **286**, 16716–16725 (2011).
- Faller, M., Matsunaga, M., Yin, S., Loo, J. A. & Guo, F. Heme is involved in microRNA processing. *Nat. Struct. Mol. Biol.* **14**, 23–29 (2007).
- Bhattarai, A. et al. Mechanism of tripeptide trimming of amyloid  $\beta$ -peptide 49 by  $\beta$ -Secretase. *J. Am. Chem. Soc.* **144**, 6215–6226 (2022).
- Feng, L. et al. Structure of a site-2 protease family intramembrane metalloprotease. *Science* **318**, 1608–1612 (2007).
- Schneider, J. S. & Glickman, M. S. Function of site-2 proteases in bacteria and bacterial pathogens. *Biochim Biophys. Acta* **1828**, 2808–2814 (2013).
- Teoh, W. P., Matson, J. S. & DiRita, V. J. Regulated intramembrane proteolysis of the virulence activator TcspP in *Vibrio cholerae* is initiated by the tail-specific protease (Tsp). *Mol. Microbiol.* **97**, 822–831 (2015).
- Marquart, M. E. Pathogenicity and virulence of *Streptococcus pneumoniae*: Cutting to the chase on proteases. *Virulence* **12**, 766–787 (2021).
- Straume, D., Stamsas, G. A., Salehian, Z. & Havarstein, L. S. Overexpression of the fratricide immunity protein ComM leads to growth inhibition and morphological abnormalities in *Streptococcus pneumoniae*. *Microbiol. (Read.)* **163**, 9–21 (2017).
- Verderber, E. et al. Role of the hemA gene product and delta-aminolevulinic acid in regulation of *Escherichia coli* heme synthesis. *J. Bacteriol.* **179**, 4583–4590 (1997).
- Schägger, H. & Jagow, G. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368–379 (1987).
- Wu, E. L. et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J. Comput. Chem.* **35**, 1997–2004 (2014).

48. Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One* **2**, e880 (2007).
49. Jo, S., Lim, J. B., Klauda, J. B. & Im, W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. *Biophys. J.* **97**, 50–58 (2009).
50. Lindahl, E., Hess, B. & van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Mod.* **7**, 306–317 (2001).
51. Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
52. Daura, X. et al. Peptide Folding: When simulation meets Experiment. *Angew. Chem. Int. Ed.* **38**, 236–240 (1999).
53. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
54. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543–D552 (2022).
55. Lolicato, F. & Beck, R. AlphaFold2 model and Molecular dynamics simulations files for “In vivo characterization of the bacterial intramembrane-cleaving protease RseP using the heme binding tag-based assay iChIPSpy”. <https://doi.org/10.5281/zenodo.7612846>. (2022).
56. Zhang, L. & Guarente, L. Heme binds to a short sequence that serves a regulatory function in diverse proteins. *EMBO J.* **14**, 313–320 (1995).

## Acknowledgements

This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 112927078-TRR83, and by the Ministry of Education and Research (BMBF) – project Drugs4Future (13FH8I051A) within the “FH-Impuls” framework. CSC-IT Centre for Science Ltd. (Espoo, Finland) is acknowledged for excellent computational resources. The authors gratefully acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research, and the Arts Baden-Württemberg (MWK) and the DFG through grant INST 35/1314-1 FUGG and INST 35/1503-1 FUGG. For the publication fee we acknowledge financial support by DFG within the funding programme „Open Access Publikationskosten“ as well as by Heidelberg University.

## Author contributions

T.K. formed the concept for the in vivo intramembrane-cleaving protease assay, designed the overall research in coordination with B.B. and performed all experiments apart from analytical mass spectrometry which was done by R.G., F.R. and C.H. R.B. and F.L. did the MD simulations. T.K. prepared Figs. 1 to 6 and Supplementary Figs. 1 to 7 (apart from

Supplementary Fig. 5e which was prepared by RB). Funding acquisition: B.B., C.H. and W.N. T.K. and B.B. wrote the manuscript. Review and editing: R.G., F.R., F.L., R.B., C.H. and W.N.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-04654-z>.

**Correspondence** and requests for materials should be addressed to Thomas Kupke or Britta Brügger.

**Peer review information** *Communications Biology* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Ingrid Span and Gene Chong. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023