THE EMBO JOURNAL

A conserved signaling pathway activates bacterial CBASS immune signaling in response to DNA damage

Rebecca Lau, Eray Enustun, Yajie Gu, Justin Nguyen, and Kevin Corbett
DOI: 10.15252/embj.2022111540

Corresponding author(s): Kevin Corbett (kcorbett@ucsd.edu)

Review Timeline:

Submission Date:	27th Apr 22
Additional correspondence (Editor):	31st May 22
Additional correspondence (Author):	1st Jun 22
Editorial Decision:	3rd Jun 22
Revision Received:	16th Aug 22
Editorial Decision:	6th Sep 22
Revision Received:	7th Sep 22
Accepted:	8th Sep 22

Editor: Hartmut Vodermaier

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

Thank you for submitting your manuscript on a new bacterial defense system to The EMBO Journal. I have now received the reports of three expert referees, copied below for your information. As you will see, the referees acknowledge the overall interest of the topic and of your analyses, but also (especially referees 2 and 3) raise a number of substantive concerns that would appear to affect the eventual suitability of this work for our broad-readership journal. Since it is not clear if, and to which extent, these issues might be easily addressable during a revision, I would in this case appreciate discussing with you how you would envision responding to the referees' points should you be given the opportunity to revise this work for The EMBO Journal. Therefore, please carefully consider the constructive criticisms of the referees and send back a brief point-by-point response outlining how they might be addressed/clarified. With your proposal in hand, I would be happy to further discuss via email or Zoom call, either later over the course this week or following my return to the office on June 13th.

\---------------------------------

Referee #1 (Report for Author)

The manuscript by Lau et al., describes detailed biochemical and structural characterisation of a two-part signalling system controlling expression of CBASS phage defence systems. The data robustly demonstrate that a metallopeptidase, CapP, cleaves a HTH-containing transcriptional repressor, CapH. Furthermore, this pairing is prevalent throughout CBASS loci and the authors also show they are associated with other known forms of phage defence system, and many systems that are predicted to act in phage defence.

This is likely to be of interest to a wide readership, and the high quality of the data support the conclusions made in the manuscript. The intriguing (and perhaps in some ways disappointing) result is that deletion of CapP, whilst preventing over-expression of CBASS components, actually had no impact on phage defence. This is a strange disconnect but I think the authors dealt with it admirably within the first two paragraphs of the discussion and with Figure 7. I do not suggest that it is an issue. Instead, it is a curiosity that allows the authors to propose a more complex model for investigation in future studies. The structural models have been built very well (with enviable resolution!).

I suggest the following minor corrections:
1. Intro first paragraph, last sentence. Co-operation between systems has recently been shown by Picton et al, NAR, 2021. Please edit statement to reflect these data. https://academic.oup.com/nar/article/49/19/11257/6389623
2. Results second paragraph. "Vast majority" - please give actual numbers.
3. First paragraph of "CapH binds the promoter region of its associated CBASS system". In Figure 2A, you have the alpha regions labelled but don't mention them until later. I think they need introducing earlier on, here. Also, how do you know they are "highly conserved" residues. Please explain, or demonstrate by showing CONSURF outputs as supp data.
4. Fig. S2A - has I99M been modelled at half occupancy for two conformers? It looks like it, unless I am mistaken. If so, please add comment to fig legend to make this clear.
5. End of "CapH oligomerization is required for DNA binding". Dimers and tetramers are invoked, but how do the authors envision the final DNA-bound complex looking? In the 22 bp site, are there inverted repeats? Will all 4 HTH domains bind? I do not want to suggest EMSAs, as the current data are sufficient for the conclusions. However, some discussion of a binding model would be helpful, or indeed EMSAs to show the number of binding steps to get full saturation, depending on probes

used.
6. When discussing CapP you state there is a 3-stranded beta sheet and then mention an extra strand. Perhaps clarify to say the extra strand acts as the 4th strand of an extended sheet.
7. When introducing Fig 6A the text specifically states, "CapP-mediated CapH cleavage...". I suggest removing the two words "CapP-mediated", as whilst it is likely true (fig 5A), without doing the E98Q control it cannot be stated.
8. Fig. 3B, Fig S2B, please remove the little coloured lines on the peaks, they look scruffy, yuck! Neater labels can be applied.
9. Fig. S4C legend - adjust as I do not think E98Q was used in this panel.

Referee #2 (Report for Author)

I. General summary and opinion about the principal significance of the study, its questions and findings.
In recent years a large number of novel bacterial defense systems have been identified, and there is considerable interest in how these defense systems are regulated. This paper identifies a two component regulatory system comprising the CapP and CapH genes that are linked to a subset of CBASS defense systems (408 cases out of ~6000 CBASS systems analyzed). The authors show that CapH binds to the promoter region of the adjacent CBASS system (presumably as a tetramer) and represses transcription of the CBASS system. They demonstrate that CapP is a protease that cleaves CapH, and that CapP proteolytic activity is stimulated by single-stranded DNA. Cleavage of CapH in turn relieves CapH-mediated transcriptional repression of CBASS genes. The authors show that capP was not required for CBASS restriction of a lytic phage (λ cI-) of the E. coli MS115-1 CBASS system (they had previously shown the four core CBASS genes of this system did restrict λ cI- in this system). The authors propose a model in which CapP/CapH primarily functions to restrict lysogenic phage infections. In this model, DNA damage leads to production of ssDNA, which binds CapP and activates its proteolytic cleavage of CapH (thus leading to high levels expression of the core CBASS genes).

This a thorough and technically sound study that provides strong support for the author's model. There is substantial interest in the field about defense system regulation, and this report provides a clear demonstration of one such mechanism. The paper is well written and easy to follow, and would be of high interest to the EMBO audience.

I. Specific major concerns essential to be addressed to support the conclusions

1. The biochemical and x-ray studies are overall well done, but the SEC-MALS analysis of full-length CapH in Figure 3B needs to be repeated with untagged protein. This analysis should also be done for the untagged CapH (I99M) mutant (Figure S2B); consider showing an overlay of the two proteins in Figure S2b. This is an important experiment since CapH forming a tetramer is a substantial component of the model presented by the authors.

The SEC-MALS analyses for the SUMO-CapHCTD and SUMO-CapHCTD do not need to be repeated since the authors have crystal structures of these domains.

2. There needs to be more description of the DNA target sequence recognized by CapH. Ideally the authors would present binding studies with an expanded DNA target containing additional sequence

flanking CapH's 22 bp target. If CapH indeed forms a tetramer, it seems possible that it may bind on a larger footprint than the 22 bp shown in Figure 1D (which may result in higher binding affinity). Gel shifts could be used if the fluorescence polarization studies are not possible with longer constructs. The authors may also want to comment on the nature of the 22 bp target (eg does it contain any palindromic or tandemly repeated sequences?).

II. Minor concerns that should be addressed
1. Update Methods to provide description of purification of tagged protein constructs.
2. Figure 3D and S2C legends: clarify which proteins constructs (tagged or untagged) were used in the DNA binding studies.
3. Do the authors have any insight to where ssDNA may bind on capP? For example, does an electrostatic analysis of CapP indicate any regions of the protein that would be compatible with DNA binding?

4. Data should be shown for the experiment in which DNAse-treated boiled lysates decreased CapP stimulation (this is described as data not shown).

5. The legend of Figure 7 starts: "Upon DNA damage induced by lysogenic phage infection...". Please clarify if this is correct (i.e that lysogenic phage induce DNA damage) or whether DNA damage in general (i.e. not linked to lysogenic phage infection) is the source of ssDNA that activates CapP. The text describes that DNA damage is a signal that induces lysogenic phage to become lytic.

III. Any additional non-essential suggestions for improving the study (which will be at the author's/editor's discretion)

No other suggestions.

Referee #3 (Report for Author)

The authors describe an interesting characterization of two regulators, a transcriptional repressor CapH and a metallopeptidase CapP, that control the expression of a CBASS phage immune system derived from E. coli MS115-1. Through structural and biochemical analyses, the authors provide evidence that CapH is a helix-turn-helix (HTH) DNA binding protein that serves as the CBASS operon repressor, and CapP is a protease that cleaves the repressor when bound to ssDNA. The regulators are homologues to SOS response regulators from Deinococcus deserti, studied in detail (e.g. PMID 25170972, 31598697), whereby IrrE was found to cleave the repressor DdrO. The regulation explored here is also very similar to the classical RecA-LexA SOS response discovered almost 50 years ago.
The conceptual advance provided by this study is limited as similar regulatory systems were already described. Furthermore, the CapH-CapP regulation of CBASS is shown to be dispensable for phage defense, and therefore the biological function of this regulation is not clear. I also find major flaws in the experimental design and interpretations as detailed below.
Major concerns:
1) The biological function of CapH-CapP is unknown and surprisingly, the authors found that the repression or activation of the operon has no effect on phage infection. This alarming result suggests that phage defense is not the actual function of the system. The authors showed a CapH-CapP-mediated increase in CBASS expression in response to DNA damage, a typical cue that activates

the SOS response, but did not link the increase to any phenotypic effect such as resistance to DNA damage.
The authors do not use the native host, E. coli MS115-1, to study the system impact, and instead express it artificially in a non-host bacterium. Using the natural host might help in revealing the actual function of the system.
2) The authors show that the regulation of CBASS operon does not influence phage infection, yet they suggest that this is a defense operon. They further claim that the CBASS basal expression, undetectable by Western blot, is sufficient to provide maximal defense. This is an unsubstantiated claim, and the authors need to monitor the expression of the system by other methods (e.g. RT-PCR for all genes in the operon). It will be also beneficial to use a cGAS inactive mutant or expressing a catalytically dead NucC to validate that the effect is directly mediated by the system. In addition, it is critical to test bacterial growth and viability, with and without the CBASS carrying plasmid to rule out an indirect influence on infection.
3) Based on figure 1I, the authors concluded that cells harboring WT and ΔcapP CBASS have similar DNA degradation kinetics. This is very unclear to me; why do cells having such a difference in the nuclease expression display similar abortive infection kinetics? Figure 1I, is inconclusive, the images are not clear, very few cells are shown for the CBASS carrying strains, and the DAPI staining is heterogeneous from the beginning. The authors should quantify the data, provide clearer images, and use additional methodologies such as DNA extraction from infected cells or nick end labeling (TUNEL) to substantiate their claim.
4) The authors observed high nucC levels in the absence of capH. Does it impact cell viability? I would expect such high levels of a nuclease to have an impact on cell growth and viability.
In line with this, does the increase in nucC during DNA damage cause host DNA degradation? Could nucC be a repair enzyme rather than a non-specific nuclease?
5) The regulation observed is very similar to the classical RecA-LexA SOS response. Yet, the authors do not draw a clear line between the systems and their components; i.e. following binding to ssDNA, RecA is activated and facilitates the LexA repressor cleavage culminating in SOS gene activation.
6) The promoter region investigated in this study is not defined. Two promoters driving opposing transcription are located in the intergenic region between capH-cdnC. Please define and provide details.
7) Do additional E. coli genes contain the putative CapH binding motive described in Fig 2C?
8) Why only fragments of the CapH were crystalized and not the entire protein?
9) The DNA binding and oligomerization states of CapH (Fig 3) should be substantiated by gel shift analysis.
10) Fig 1G: The authors should show host viability and infection kinetics simply by OD600 measurements. The Y axis was defined as "plaque forming units per mL of purified phage", could the authors describe the procedure of phage purification?
11) The results for ΔcapH should be included in Figure 1E-1I.
12) Figure 2F, R44A still acts as a repressor as appears in the gel. This seems inconsistent with the text (p6: "We found that mutation of Ser32 or Arg44 on the predicted DNA-binding face of CapH eliminated detectable DNA binding").
13) The figure legends and the methodologies lack essential details such as the bacterial strains and their genotypes, MOI, plasmid construction.
14) The introduction lacks important information such as reviewing the SOS response.
15) p6: I think "CinR" should be "SinR".

Thanks for your note from earlier this week, and for handling the review of this manuscript. After carefully reading and considering these reviews with the first author, Rebecca Lau, I'm confident that we will be able to address all of the substantive concerns raised by the reviewers. In many cases, this will involve simply rewriting parts of the manuscript to better clarify our findings and how they point toward a more nuanced picture of anti-phage immunity than is currently appreciated by most of the field. In other cases, these comments will spur us to perform some experiments that were admittedly lacking in the original manuscript.

In terms of new experiments, one major addition that we will provide in a revision is a more comprehensive survey of the CBASS promoter region in question. Spurred by the reviewers' comments, we have already annotated the most likely forward and reverse operator sequences (-35, -10, and TSS) in this 182 bp region. While we are confident that the site we identified is valid, we appreciate the reviewers' questions about site size and will perform careful binding assays with panels of oligos spanning the entire region to identify the full binding site. We will use both fluorescence polarization assays as in our original submission, and EMSA assays which have the potential to show binding of one versus two CapH dimers through a supershift.

In another set of control experiments, we plan to include RT-qPCR data showing expression levels of CdnC and NucC in our system, to quantitatively compare the basal expression level to that induced by either phage infection or DNA damage. We have worked on this assay extensively in the past and found it to be somewhat irreproducible, which is why we did not include the data in the original manuscript. Nonetheless, we will work to optimize this assay to provide a control for our Western blots and verify that there is measurable expression of the system in unperturbed cells.

Below we briefly address each comment and question of the three reviewers, with referee comments in black text and our responses in blue text.

Referee #1

The manuscript by Lau et al., describes detailed biochemical and structural characterisation of a two-part signalling system controlling expression of CBASS phage defence systems. The data robustly demonstrate that a metallopeptidase, CapP, cleaves a HTH-containing transcriptional repressor, CapH. Furthermore, this pairing is prevalent throughout CBASS loci and the authors also show they are associated with other known forms of phage defence system, and many systems that are predicted to act in phage defence.

This is likely to be of interest to a wide readership, and the high quality of the data support the conclusions made in the manuscript. The intriguing (and perhaps in some ways disappointing) result is that deletion of CapP, whilst preventing over-expression of CBASS components, actually had no impact on phage defence. This is a strange disconnect but I think the authors dealt with it admirably within the first two paragraphs of the discussion and with Figure 7. I do not suggest that it is an issue. Instead, it is

a curiosity that allows the authors to propose a more complex model for investigation in future studies. The structural models have been built very well (with enviable resolution!).

I suggest the following minor corrections:
1. Intro first paragraph, last sentence. Co-operation between systems has recently been shown by Picton et al, NAR, 2021. Please edit statement to reflect these data.
https://academic.oup.com/nar/article/49/19/11257/6389623

We thank the referee for pointing this out, and we will add this reference.

2. Results second paragraph. "Vast majority" - please give actual numbers.

We apologize for the vagueness of this statement. Of the 408 CBASS systems identified with an associated CapP gene, 15 are apparent CapH-CapP fusions and therefore do not have a separate CapH gene. In 70 cases, there is no CapH gene annotated but there is an un-annotated ORF in the genome sequence. Therefore, all instances of CapP have either a fused or neighboring CapH gene. We will clarify this point in the revised version.

3. First paragraph of "CapH binds the promoter region of its associated CBASS system". In Figure 2A, you have the alpha regions labelled but don't mention them until later. I think they need introducing earlier on, here. Also, how do you know they are "highly conserved" residues. Please explain, or demonstrate by showing CONSURF outputs as supp data.

We will highlight these C-terminal alpha-helical elements in CapH as requested, and note that they were predicted by PHYRE modeling prior to structure determination. Also, we will include in the fiigure a segment of our comprehensive sequence alignment of CapH genes, highlighting the conservation of the residues in question.

4. Fig. S2A - has I99M been modelled at half occupancy for two conformers? It looks like it, unless I am mistaken. If so, please add comment to fig legend to make this clear.

Yes, this residue was modeled at half occupancy for two conformers. We will add this information to the figure legend.

5. End of "CapH oligomerization is required for DNA binding". Dimers and tetramers are invoked, but how do the authors envision the final DNA-bound complex looking? In the 22 bp site, are there inverted repeats? Will all 4 HTH domains bind? I do not want to suggest EMSAs, as the current data are sufficient for the conclusions. However, some discussion of a binding model would be helpful, or indeed EMSAs to show the number of binding steps to get full saturation, depending on probes used.

The reviewer's comment is much appreciated; in the revised version, we will better address our model for how binding is achieved. This will be significantly aided by our planned experiments to identify the full site, and planned EMSA assays to determine whether one or two CapH dimers bind the site.

Preliminarily, we do not detect any tandem repeats or palindromes in this region; we will mention this finding in this section in the revised version.

6. When discussing CapP you state there is a 3-stranded beta sheet and then mention an extra strand. Perhaps clarify to say the extra strand acts as the 4th strand of an extended sheet.

We will adjust the wording as requested.

7. When introducing Fig 6A the text specifically states, "CapP-mediated CapH cleavage...". I suggest removing the two words "CapP-mediated", as whilst it is likely true (fig 5A), without doing the E98Q control it cannot be stated.

We will adjust the wording as requested.

8. Fig. 3B, Fig S2B, please remove the little coloured lines on the peaks, they look scruffy, yuck! Neater labels can be applied.

The reviewer may have misunderstood that the colored lines are in fact molecular weight estimates (right-hand vertical axis). We will add further clarification to the figure and figure legend to make this point more clear.

9. Fig. S4C legend - adjust as I do not think E98Q was used in this panel.

We will re-word this legend.

Referee #2

I. General summary and opinion about the principal significance of the study, its questions and findings. In recent years a large number of novel bacterial defense systems have been identified, and there is considerable interest in how these defense systems are regulated. This paper identifies a two component regulatory system comprising the CapP and CapH genes that are linked to a subset of CBASS defense systems (408 cases out of ~6000 CBASS systems analyzed). The authors show that CapH binds to the promoter region of the adjacent CBASS system (presumably as a tetramer) and represses transcription of the CBASS system. They demonstrate that CapP is a protease that cleaves CapH, and that CapP proteolytic activity is stimulated by single-stranded DNA. Cleavage of CapH in turn relieves CapH-mediated transcriptional repression of CBASS genes. The authors show that capP was not required for CBASS restriction of a lytic phage (λ cI-) of the E. coli MS115-1 CBASS system (they had previously

shown the four core CBASS genes of this system did restrict λ cI- in this system). The authors propose a model in which CapP/CapH primarily functions to restrict lysogenic phage infections. In this model, DNA damage leads to production of ssDNA, which binds CapP and activates its proteolytic cleavage of CapH (thus leading to high levels expression of the core CBASS genes).

This a thorough and technically sound study that provides strong support for the author's model. There is substantial interest in the field about defense system regulation, and this report provides a clear demonstration of one such mechanism. The paper is well written and easy to follow, and would be of high interest to the EMBO audience.

I. Specific major concerns essential to be addressed to support the conclusions

1. The biochemical and x-ray studies are overall well done, but the SEC-MALS analysis of full-length CapH in Figure 3B needs to be repeated with untagged protein. This analysis should also be done for the untagged CapH (I99M) mutant (Figure S2B); consider showing an overlay of the two proteins in Figure S2b. This is an important experiment since CapH forming a tetramer is a substantial component of the model presented by the authors.

The SEC-MALS analyses for the SUMO-CapHCTD and SUMO-CapHCTD do not need to be repeated since the authors have crystal structures of these domains.

We feel strongly that the use of an MBP tag, which is known to be monomeric, did not affect the findings of the noted SEC-MALS experiments. Nonetheless, we will perform a new set of SEC-MALS experiments with untagged full-length wild-type CapH and the I99M mutant.

2. There needs to be more description of the DNA target sequence recognized by CapH. Ideally the authors would present binding studies with an expanded DNA target containing additional sequence flanking CapH's 22 bp target. If CapH indeed forms a tetramer, it seems possible that it may bind on a larger footprint than the 22 bp shown in Figure 1D (which may result in higher binding affinity). Gel shifts could be used if the fluorescence polarization studies are not possible with longer constructs. The authors may also want to comment on the nature of the 22 bp target (eg does it contain any palindromic or tandemly repeated sequences?).

As noted above, we will further dissect the promoter region by testing CapH binding to a panel of overlapping 40mer DNAs spanning the entire 182 bp promoter region. We will then narrow the search to identify the most preferred site, which will likely include the identified 22 bp region and potentially flanking regions. Once the full site is identified, and we have the results of EMSA assays that may show a supershift if two CapH dimers bind, we will develop and describe a more complete model for DNA binding by CapH.

II. Minor concerns that should be addressed

1. Update Methods to provide description of purification of tagged protein constructs.

All proteins were purified using the same purification strategy, as they all encoded a His6-tag fused either to the protein itself or upstream of the MBP or SUMO tag. We will add a note about the tags to the first sentence of the Methods section.

2. Figure 3D and S2C legends: clarify which proteins constructs (tagged or untagged) were used in the DNA binding studies.

We will add this information.

3. Do the authors have any insight to where ssDNA may bind on capP? For example, does an electrostatic analysis of CapP indicate any regions of the protein that would be compatible with DNA binding?

Based on the known functions of the GAF domain in other proteins, it is reasonable to assume that ssDNA binds to this domain. We have been unable to isolate mutants in the proposed ssDNA binding site that are soluble enough to be purified for in vitro testing, limiting our ability to predict where ssDNA binds or how it alters CapP structure to activate its peptidase activity. We will add further discussion of these points to the revised manuscript.

4. Data should be shown for the experiment in which DNAse-treated boiled lysates decreased CapP stimulation (this is described as data not shown).

We will add this to the supplementary data.

5. The legend of Figure 7 starts: "Upon DNA damage induced by lysogenic phage infection...". Please clarify if this is correct (i.e that lysogenic phage induce DNA damage) or whether DNA damage in general (i.e. not linked to lysogenic phage infection) is the source of ssDNA that activates CapP. The text describes that DNA damage is a signal that induces lysogenic phage to become lytic.

We will change the wording to "Upon DNA damage…" since the reviewer is correct that lysogenic phage also induce DNA damage, and we (as yet) do not know the source of the activating DNA damage.

Referee #3

The authors describe an interesting characterization of two regulators, a transcriptional repressor CapH and a metallopeptidase CapP, that control the expression of a CBASS phage immune system derived from E. coli MS115-1. Through structural and biochemical analyses, the authors provide evidence that CapH is a helix-turn-helix (HTH) DNA binding protein that serves as the CBASS operon repressor, and CapP is a protease that cleaves the repressor when bound to ssDNA. The regulators are homologues to

SOS response regulators from Deinococcus deserti, studied in detail (e.g. PMID 25170972, 31598697), whereby IrrE was found to cleave the repressor DdrO. The regulation explored here is also very similar to the classical RecA-LexA SOS response discovered almost 50 years ago.
The conceptual advance provided by this study is limited as similar regulatory systems were already described. Furthermore, the CapH-CapP regulation of CBASS is shown to be dispensable for phage defense, and therefore the biological function of this regulation is not clear. I also find major flaws in the experimental design and interpretations as detailed below.
Major concerns:
1) The biological function of CapH-CapP is unknown and surprisingly, the authors found that the repression or activation of the operon has no effect on phage infection. This alarming result suggests that phage defense is not the actual function of the system. The authors showed a CapH-CapP-mediated increase in CBASS expression in response to DNA damage, a typical cue that activates the SOS response, but did not link the increase to any phenotypic effect such as resistance to DNA damage.

It is well established by multiple studies that CBASS systems are anti-phage immune systems. The system under study also shows anti-phage activity, either when the four core CBASS genes are expressed ectopically (as in our earlier studies published in *Mol Cell*) or when they under the control of CapH and CapP (this work). The protective effect of the system is markedly lower when under CapH/CapP control compared to ectopic expression, which suggests that *primary* phage defense is not the main function of this particular CBASS system (we suggest that it is a *secondary*, or backup defense system). We will clarify these points in a revised manuscript.

The reviewer is correct that altering CapH or CapP to eliminate DNA damage-mediated expression activation does not affect the (relatively modest) anti-phage effect of the full CBASS system. However, we take strong issue with the reviewer's contention that this result is "alarming" - indeed, we think of it as an opportunity to explore more complex signaling in the bacterial defense field. We present a compelling model for how CapH and CapP could enable a linked CBASS system (or another CapH/CapP-controlled defense system) to serve as a secondary or backup defense, after activation of a DNA-targeting restriction-modification or CRISPR-Cas pathway. While testing this model is outside the scope of the current work, the concept of synergy between defense systems is a new concept in the field and will spur significant interest and follow-up research.

Finally, to the reviewer's suggestion that this pathway may confer resistance to DNA damage: this pathway activates a destructive nuclease, not a DNA repair pathway.

The authors do not use the native host, E. coli MS115-1, to study the system impact, and instead express it artificially in a non-host bacterium. Using the natural host might help in revealing the actual function of the system.

The native host strain encodes an array of anti-phage defense systems: AbiH, Gabija, Kiwa, Lamassu, Type III restriction modification, Class IE CRISPR, CBASS (our system), and Retron (as identified by the DefenseFinder server). We have attempted lambda-red genome engineering in this strain to generate

tagged or mutant versions of the CBASS operon, without success. This is perhaps not surprising, given the array of defense systems in the genome.

From a broader perspective, while we appreciate that using a native host is preferable in many cases, CBASS systems are known to be self-contained and portable. For our work, using the system in isolation, away from other confounding effects, is preferable.

The reviewer is correct that the genomic context of this system is likely important. We are currently pursuing work to test cooperation or synergy between this CBASS system and other defense systems in *E. coli* MS115-1. However, these experiments will take significant time and effort, and are outside the scope of the current structure and biochemical mechanism-focused work.

2) The authors show that the regulation of CBASS operon does not influence phage infection, yet they suggest that this is a defense operon. They further claim that the CBASS basal expression, undetectable by Western blot, is sufficient to provide maximal defense. This is an unsubstantiated claim, and the authors need to monitor the expression of the system by other methods (e.g. RT-PCR for all genes in the operon).

This is simply an unfair criticism. Our claim that basal expression of the CBASS system is sufficient for phage protection is substantiated by the data presented in **Figure 1G-I**. Nonetheless, we do plan to add qRT-PCR assays to measure CdnC and NucC mRNA levels, in the presence and absence of phage infection or DNA damaging agents. We will include these data in a revised version of the manuscript.

It will be also beneficial to use a cGAS inactive mutant or expressing a catalytically dead NucC to validate that the effect is directly mediated by the system. In addition, it is critical to test bacterial growth and viability, with and without the CBASS carrying plasmid to rule out an indirect influence on infection.

We will include plaque assays for the cGAS-dead and NucC-dead versions of the operon in the revised version. We have also performed growth curves with several strains under study and can include these in the revised version.

3) Based on figure 1I, the authors concluded that cells harboring WT and ΔcapP CBASS have similar DNA degradation kinetics. This is very unclear to me; why do cells having such a difference in the nuclease expression display similar abortive infection kinetics?

Our data shows that the substantial increase in CBASS expression does not occur until ~90 minutes after phage infection. As the data in **Figure 1I** (and our prior studies with this system) shows, cell death occurs earlier than this (around 60 minutes) when the cells harbor the NucC-containing CBASS system. As we state in the manuscript, this result means that the basal level of NucC is sufficient to mediate genome degradation and cell death.

Figure 1I, is inconclusive, the images are not clear, very few cells are shown for the CBASS carrying strains, and the DAPI staining is heterogeneous from the beginning. The authors should quantify the data, provide clearer images, and use additional methodologies such as DNA extraction from infected cells or nick end labeling (TUNEL) to substantiate their claim.

We will provide clearer images, and quantify those images as we did in our prior work, in the revised version. Since this manuscript is not about the mechanism of NucC (which we previously studied and is by now well established), we do not consider it necessary to perform further assays to demonstrate the mechanism of NucC-dependent cell death by DNA degradation.

4) The authors observed high nucC levels in the absence of capH. Does it impact cell viability? I would expect such high levels of a nuclease to have an impact on cell growth and viability.

The reviewer is correct that in strains lacking CapH, the resulting high levels of NucC are apparently toxic. This is the reason we could not clone a Δ*capH* mutant version of the native CBASS system, prompting us to build the GFP reporter system where we could delete *capH* without associated toxicity. It is not clear why cells can live with a Δ*capP*- Δ*capH* mutant system, while those with a Δ*capH* system are apparently inviable; nonetheless, cells with the Δ*capP*- Δ*capH* system grow significantly more slowly than those with the wild-type system, likely because of the high NucC levels.

This point highlights a question as to why we did not observe toxicity in our prior studies of this system, where we cloned the four core CBASS genes into a Lac-inducible vector (prior to our discovery of CapH and CapP). Using Western blots, we find that the level of CBASS expression upon *capH* deletion is significantly higher than even the highest IPTG-inducible levels in our other plasmid. We can add this information to the revised manuscript.

In line with this, does the increase in nucC during DNA damage cause host DNA degradation? Could nucC be a repair enzyme rather than a non-specific nuclease?

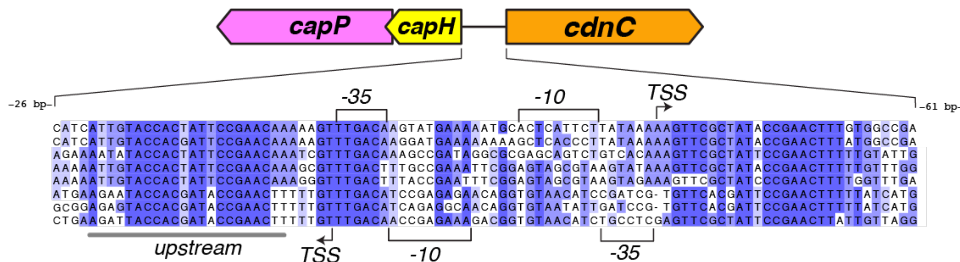
This is a good suggestion, and we will perform microscopy and cell growth assays after induction of DNA damage to observe the effects on cells. There is no chance that NucC is a repair enzyme: we previously showed that it is a highly active, nonspecific nuclease that degrades DNA to ~50 bp pieces.

5) The regulation observed is very similar to the classical RecA-LexA SOS response. Yet, the authors do not draw a clear line between the systems and their components; i.e. following binding to ssDNA, RecA is activated and facilitates the LexA repressor cleavage culminating in SOS gene activation.

We will certainly add some more background about the SOS response to the paper's introduction, and we will mention the mechanistic parallels the reviewer notes in our discussion. We feel it important to point out that the mechanisms of these two pathways are very different, even if they are both activated by single-stranded DNA.

6) The promoter region investigated in this study is not defined. Two promoters driving opposing transcription are located in the intergenic region between capH-cdnC. Please define and provide details.

This is also a good point. We took a closer look at the region in question, and have identified the -35, -10, and TSS for both the forward (core CBASS) and reverse (capH and capP) directions (using the BPROM and SAPPHIRE servers). The site labeled "upstream" in the figure below is our identified CapH binding site; as noted above, we will perform additional DNA binding assays to better determine the full extent of this site.



7) Do additional *E. coli* genes contain the putative CapH binding motive described in Fig 2C?

We will search the genome of the host strain (*E. coli* MS115-1) to identify any likely CapH binding sites, after defining the full site in the CBASS promoter. While we doubt that we will identify anything as CBASS systems are known to be portable and self-contained, this may turn up interesting regulation that we had not anticipated.

8) Why only fragments of the CapH were crystallized and not the entire protein?

We attempted to crystallize the full length protein, but it did not form crystals. We suspect that the flexible linker region between the two domains prevents organized packing.

9) The DNA binding and oligomerization states of CapH (Fig 3) should be substantiated by gel shift analysis.

We will perform EMSAs for the revised version. These may highlight a supershift if two CapH dimers bind the region in question.

10) Fig 1G: The authors should show host viability and infection kinetics simply by OD600 measurements. The Y axis was defined as "plaque forming units per mL of purified phage", could the authors describe the procedure of phage purification?

We will include phage purification protocols in the methods. As noted above, we have also performed growth curves after infection for several strains, and will include this data in the revised version.

11) The results for ΔcapH should be included in Figure 1E-1I.

As noted above, we could not generate a Δ*capH* version of the full CBASS system due to toxicity (we will make this point clearer in the revised manuscript).

12) Figure 2F, R44A still acts as a repressor as appears in the gel. This seems inconsistent with the text (p6: "We found that mutation of Ser32 or Arg44 on the predicted DNA-binding face of CapH eliminated detectable DNA binding").

The apparent disconnect between in vitro and in vivo phenotypes for R44A reflects a common phenomenon: while DNA binding of the R44A mutant is defective enough to be undetectable in vitro, it apparently does bind DNA sufficiently to suppress CBASS expression in cells. This may arise partly from our use of a (possibly) incomplete binding site in our DNA binding assays, which we will address as described above. Otherwise, we will better explain this disconnect in the revised text.

13) The figure legends and the methodologies lack essential details such as the bacterial strains and their genotypes, MOI, plasmid construction.

We will include this information.

14) The introduction lacks important information such as reviewing the SOS response.

As noted above, our data do not support a functional parallel with the SOS response beyond the very broad strokes concept of ssDNA activated proteolysis. Nonetheless, it is an interesting conceptual parallel that we will be sure to address in the revised manuscript; if not in the introduction, then in the discussion.

15) p6: I think "CinR" should be "SinR".

We thank the reviewer for catching this error, and will correct this name in the revised version.

1st Editorial Decision **3rd Jun 2022**

Thank you for your detailed tentative response to the referee reports, and proposal for revising this work for The EMBO Journal. I have now had a chance to consider these plans, and found them potentially well-suited for addressing the main concerns raised by our three reviewers. I shall therefore be happy to consider a revised manuscript further for EMBO Journal publication, and would like to herewith formally invite you to prepare and resubmit a new version, modified and extended as proposed in your draft response. I should remind you that it is our policy to allow only a single round of (major) revision, making it important to carefully revise and answer all points raised to the referees' satisfaction at this point. As usual, any competing manuscript published during the course of your revision will have no negative impact on our final decision on your study. Furthermore, we are always open to discussing possible extensions of the default revision time - please do not hesitate to get in touch with us should this become necessary, or if you would like to discuss any other points related to this revision.

Responses to Reviewer Comments

We thank the reviewers for their insightful comments and questions. In the attached revised manuscript, we have added significant new data and clarified our results throughout, resulting in a more approachable and logically sound manuscript. Spurred by the reviewers' questions, we have added major new data in several areas:

First and most importantly, we have added a comprehensive analysis of the *E. coli* MS115-1 CBASS promoter region. We now show predicted promoter sequences (-35, -10, and TSS) for both the top strand (driving CBASS core gene expression) and bottom strand (driving *capH* and *capP* expression) in our updated schematic in **Fig 2**. We have added a comprehensive survey of CapH binding across this region, using a panel of 40 bp DNAs, which refines the previously identified CapH binding site and also reveals a second binding site in the promoter region. The previously identified CapH binding site originally showed modest, non-cooperative binding by CapH. We now find that expanding the target site to 40 bp (from 22 bp in the original submission) results in strong, cooperative binding to this site (now termed "Site 1"). In addition, this analysis revealed a second CapH binding site ("Site 2") containing a palindromic sequence, which also shows strong, cooperative binding by CapH. This data is presented in the revised **Fig 2** and **Fig EV3**.

Second, we now include qRT-PCR data showing the expression level of the *cdnC* gene in our system, first to show that CBASS core genes are indeed expressed at low levels in unperturbed cells, and second to quantitatively compare the basal expression level to that induced by either phage infection or DNA damage. This analysis is shown in **Fig EV1E** of the revised manuscript.

Third, we now provide quantitation for the light-microscopy images showing the effect of CBASS on phage-infected cells (**Fig EV2A-B**), and added new bacterial growth curves for both uninfected and phage-infected cells (**Fig EV2C**). These data show that the six-gene *E. coli* MS115-1 CBASS system is not toxic to host cells on its own, and that the modest anti-phage protection we observe in plaque assays is attributable to CBASS/NucC action.

In addition to these major changes, we have made minor changes throughout the text to address individual reviewer questions and clarify the text; these changes are shown in red text in the attached manuscript. Below, we address each comment and question of the three reviewers, with reviewer comments in black text and our responses in blue text.

Referee #1

The manuscript by Lau et al., describes detailed biochemical and structural characterisation of a two-part signalling system controlling expression of CBASS phage defence systems. The data robustly demonstrate that a metallopeptidase, CapP, cleaves a HTH-containing transcriptional repressor, CapH. Furthermore, this pairing is prevalent throughout CBASS loci and the authors also show they are associated with other known forms of phage defence system, and many systems that are predicted to act in phage defence.

This is likely to be of interest to a wide readership, and the high quality of the data support the conclusions made in the manuscript. The intriguing (and perhaps in some ways disappointing) result is that deletion of CapP, whilst preventing over-expression of CBASS components, actually had no impact on phage defence. This is a strange disconnect but I think the authors dealt with it admirably within the first two paragraphs of the discussion and with Figure 7. I do not suggest that it is an issue. Instead, it is a curiosity that allows the authors to propose a more complex model for investigation in future studies. The structural models have been built very well (with enviable resolution!).

I suggest the following minor corrections:

1. Intro first paragraph, last sentence. Co-operation between systems has recently been shown by Picton et al, NAR, 2021. Please edit statement to reflect these data.

We thank the referee for pointing this out; we have added this reference.

2. Results second paragraph. "Vast majority" - please give actual numbers.

We have now rewritten this paragraph to clarify our findings. Of the 408 CBASS systems identified with an associated *capP* gene, 15 are apparent CapH-CapP fusions and therefore do not have a separate *capH* gene. All others (393 systems) encode both CapH and CapP, though in 70 cases the *capH* gene is not properly annotated in the genome sequence. Therefore, all instances of *capP* have either a fused or neighboring *capH* gene.

3. First paragraph of "CapH binds the promoter region of its associated CBASS system". In Figure 2A, you have the alpha regions labelled but don't mention them until later. I think they need introducing earlier on, here. Also, how do you know they are "highly conserved" residues. Please explain, or demonstrate by showing CONSURF outputs as supp data.

We have rewritten this paragraph to better introduce the overall structure of CapH, including the C-terminal alpha-helical domain. To address the question about conservation of Ser32, Arg40, and Arg44, we now include a new **Fig EV3G** that shows a sequence logo from 56 related CapH proteins. This shows that Ser32 and Arg44 are well conserved among CapH proteins, but Arg40 is not as conserved.

4. Fig. S2A - has I99M been modelled at half occupancy for two conformers? It looks like it, unless I am mistaken. If so, please add comment to fig legend to make this clear.

Yes, this residue was modeled at half occupancy for two rotamers. We have now added this information to the figure legend.

5. End of "CapH oligomerization is required for DNA binding". Dimers and tetramers are invoked, but how do the authors envision the final DNA-bound complex looking? In the 22 bp site, are there inverted repeats? Will all 4 HTH domains bind? I do not want to suggest EMSAs, as the current data are sufficient for the conclusions. However, some discussion of a binding model would be helpful, or indeed EMSAs to show the number of binding steps to get full saturation, depending on probes used.

Based on our identification of optimal binding sites for CapH that show cooperative binding, and our finding that I99M does not eliminate the observed cooperativity of binding, we conclude that the predominant binding mode for CapH on DNA is as a homodimer. This is made more clear in the revised text.

We did perform EMSA assays with wild-type CapH and the I99M mutant (**Appendix Fig S1C**), which revealed two shifted bands for wild-type CapH and one shifted band for CapH(I99M). Since the second shifted band appears only at high CapH concentration, and the overall affinity of CapH(I99M) for both DNAs is similar to that of wild-type CapH (**Fig 3E**), we conclude that CapH tetramer

formation likely occurs only at relatively high protein concentration and plays at most a minor role in DNA binding.

6. When discussing CapP you state there is a 3-stranded beta sheet and then mention an extra strand. Perhaps clarify to say the extra strand acts as the 4th strand of an extended sheet.

We have slightly adjusted the wording of this sentence.

7. When introducing Fig 6A the text specifically states, "CapP-mediated CapH cleavage...". I suggest removing the two words "CapP-mediated", as whilst it is likely true (fig 5A), without doing the E98Q control it cannot be stated.

We have deleted “CapP-mediated” from this sentence.

8. Fig. 3B, Fig S2B, please remove the little coloured lines on the peaks, they look scruffy, yuck! Neater labels can be applied.

These colored lines are in fact molecular weight estimates from the light-scatting analysis (see right-hand vertical axis). We have now added further clarification to the figure and figure legend to make this point more clear.

9. Fig. S4C legend - adjust as I do not think E98Q was used in this panel.

We have removed mention of the E98Q mutant from this figure legend.

Referee #2

I. General summary and opinion about the principal significance of the study, its questions and findings.

In recent years a large number of novel bacterial defense systems have been identified, and there is considerable interest in how these defense systems are regulated. This paper identifies a two component regulatory system comprising the CapP and CapH genes that are linked to a subset of CBASS defense systems (408 cases out of ~6000 CBASS systems analyzed). The authors show that CapH binds to the promoter region of the adjacent CBASS system (presumably as a tetramer) and represses transcription of the CBASS system. They demonstrate that CapP is a protease that cleaves CapH, and that CapP proteolytic activity is stimulated by single-stranded DNA. Cleavage of CapH in turn relieves CapH-mediated transcriptional repression of CBASS genes. The authors show that capP was not required for CBASS restriction of a lytic phage (λ cI-) of the E. coli MS115-1 CBASS system (they had previously shown the four core CBASS genes of this system did restrict λ cI- in this system). The authors propose a model in which CapP/CapH primarily functions to restrict lysogenic phage infections. In this model, DNA damage leads to production of ssDNA, which binds CapP and activates its proteolytic cleavage of CapH (thus leading to high levels expression of the core CBASS genes).

This a thorough and technically sound study that provides strong support for the author's model. There is substantial interest in the field about defense system regulation, and this report provides a clear demonstration of one such mechanism. The paper is well written and easy to follow, and would be of high interest to the EMBO audience.

I. Specific major concerns essential to be addressed to support the conclusions

1. The biochemical and x-ray studies are overall well done, but the SEC-MALS analysis of full-length CapH in Figure 3B needs to be repeated with untagged protein. This analysis should also be done for the untagged CapH (I99M) mutant (Figure S2B); consider showing an overlay of the two proteins in Figure S2b. This is an important experiment since CapH forming a tetramer is a substantial component of the model presented by the authors.

The SEC-MALS analyses for the SUMO-CapHCTD and SUMO-CapHCTD do not need to be repeated since the authors have crystal structures of these domains.

> SEC-MALS analysis of untagged CapH and CapH(I99M) is now presented in the new **Appendix Fig S1A**. These data are not as visually appealing as those from tagged CapH due to the low yield of the proteins, but the data are nonetheless unambiguous. The analysis shows that wild-type CapH has a tendency to form slightly higher oligomers (average of 2.3-mer) compared to the I99M mutant (average of 2.0-mer). The relatively small difference between the two proteins agrees with our newer data showing that the I99M mutation has only a small effect on binding affinity for both Site 1 and Site 2, and supports a model in which CapH tetramer formation plays at most a minor role in determining DNA binding affinity.

2. There needs to be more description of the DNA target sequence recognized by CapH. Ideally the authors would present binding studies with an expanded DNA target containing additional sequence flanking CapH's 22 bp target. If CapH indeed forms a tetramer, it seems possible that it may bind on a larger footprint than the 22 bp shown in Figure 1D (which may result in higher binding affinity). Gel shifts could be used if the fluorescence polarization studies are not possible with longer constructs. The authors may also want to comment on the nature of the 22 bp target (eg does it contain any palindromic or tandemly repeated sequences?).

> Spurred by this comment and other inconsistencies in our original data, we performed a full survey of the CBASS promoter region and, as described above, identified two sites to which CapH shows strong, cooperative binding (**Figs 2 and EV3**).
>
> Our newly-identified Site 2 contains a perfect palindrome, suggesting that two CapH monomers could bind these sites. Site 1, while highly conserved (**Fig EV3A**), does not contain any recognizable palindromes or tandem repeats.

II. Minor concerns that should be addressed

1. Update Methods to provide description of purification of tagged protein constructs.

> All proteins were purified using the same purification strategy, as they all encoded a His_6-tag fused either to the protein itself or upstream of the MBP or SUMO tag. We have added further clarification about the tags to the first sentence of the Methods section.

2. Figure 3D and S2C legends: clarify which proteins constructs (tagged or untagged) were used in the DNA binding studies.

We now note specifically in both the **Methods** and in the Figure legends that all DNA binding assays were done with His$_6$-MBP tagged proteins.

3. Do the authors have any insight to where ssDNA may bind on capP? For example, does an electrostatic analysis of CapP indicate any regions of the protein that would be compatible with DNA binding?

Based on the known functions of the GAF domain in other proteins, it is reasonable to assume that ssDNA binds to this domain. We have been unable to isolate mutants in the proposed ssDNA binding site that are soluble enough to be purified for in vitro testing, limiting our ability to predict where ssDNA binds or how it alters CapP structure to activate its peptidase activity. Examination of the surface charge of CapP does not lend any further insight into the DNA binding site. Since the binding of single-stranded DNA may be primarily through the bases rather than the charged backbone, the lack of a clearly-defined charged patch is not surprising.

4. Data should be shown for the experiment in which DNAse-treated boiled lysates decreased CapP stimulation (this is described as data not shown).

We have removed mention of this experiment from the manuscript. When revisiting the data from this experiment, we realized that we did not observe a significant reduction in CapP stimulation after treatment with DNase. This finding is consistent with our later observation that DNAs as short as 5 bases can stimulate CapP: since our DNase treatment was likely incomplete, there likely remained a significant number of short DNA segments even after treatment. We regret the inclusion of this erroneous statement in the original manuscript.

We do now include the noted data with regard to fractionation of the *E. coli* cell lysate by anion-exchange chromatography, in the new **Fig EV5B**. Since all fractions from a Superdex 75 gel filtration column separation of cell lysate showed cleavage activity, we chose not to show this data.

5. The legend of Figure 7 starts: "Upon DNA damage induced by lysogenic phage infection...". Please clarify if this is correct (i.e that lysogenic phage induce DNA damage) or whether DNA damage in general (i.e. not linked to lysogenic phage infection) is the source of ssDNA that activates CapP. The text describes that DNA damage is a signal that induces lysogenic phage to become lytic.

We have changed the wording to "Upon DNA damage…" to avoid confusion and because we do not know the exact source(s) of DNA damage that activate CapP.

Referee #3

The authors describe an interesting characterization of two regulators, a transcriptional repressor CapH and a metallopeptidase CapP, that control the expression of a CBASS phage immune system derived from E. coli MS115-1. Through structural and biochemical analyses, the authors provide evidence that CapH is a helix-turn-helix (HTH) DNA binding protein that serves as the CBASS operon repressor, and

CapP is a protease that cleaves the repressor when bound to ssDNA. The regulators are homologues to SOS response regulators from Deinococcus deserti, studied in detail (e.g. PMID 25170972, 31598697), whereby IrrE was found to cleave the repressor DdrO. The regulation explored here is also very similar to the classical RecA-LexA SOS response discovered almost 50 years ago.

The conceptual advance provided by this study is limited as similar regulatory systems were already described. Furthermore, the CapH-CapP regulation of CBASS is shown to be dispensable for phage defense, and therefore the biological function of this regulation is not clear. I also find major flaws in the experimental design and interpretations as detailed below.

Major concerns:

1) The biological function of CapH-CapP is unknown and surprisingly, the authors found that the repression or activation of the operon has no effect on phage infection. This alarming result suggests that phage defense is not the actual function of the system. The authors showed a CapH-CapP-mediated increase in CBASS expression in response to DNA damage, a typical cue that activates the SOS response, but did not link the increase to any phenotypic effect such as resistance to DNA damage.

It is well established by multiple studies that CBASS systems are anti-phage immune systems. The system under study shows strong anti-phage activity when the four core CBASS genes are expressed ectopically (as shown in our earlier published studies), and markedly weaker anti-phage activity when they under the control of CapH and CapP (this work). The weaker protective effect of the system when under its native CapH/CapP control compared to ectopic expression suggests that *primary* phage defense is not the main function of this particular CBASS system. We suggest that it is a *secondary*, or backup defense system. In the originally submitted manuscript, these points were perhaps not clear; we have tried to clarify both our findings and our interpretations in the revised manuscript.

The reviewer is correct that altering CapH or CapP to eliminate DNA damage-mediated expression activation does not affect the (modest) anti-phage effect of the full CBASS system. However, we take strong issue with the reviewer's contention that this result is "alarming" - indeed, we think of it as an opportunity to explore the complexities of signaling in bacterial defense. We present a compelling model for how CapH and CapP could enable a linked CBASS system (or another of the many CapH/CapP-controlled defense systems) to serve as a secondary or backup defense, perhaps after activation of a DNA-targeting restriction-modification or CRISPR-Cas pathway. While testing this model is outside the scope of the current work, the concept of synergy between defense systems is a new concept in the field and will spur significant interest and follow-up research.

Finally, to the reviewer's suggestion that this pathway may confer resistance to DNA damage: this pathway activates a destructive nuclease, not a DNA repair pathway. While the related proteins in *Deinococcus* activate expression of DNA repair genes as the reviewer notes, high-level expression of this CBASS system is likely toxic to the host cell – as demonstrated by our inability to clone a deregulated version of the system lacking CapH.

The authors do not use the native host, E. coli MS115-1, to study the system impact, and instead express it artificially in a non-host bacterium. Using the natural host might help in revealing the actual function of the system.

The native host strain encodes a broad array of anti-phage defense systems: AbiH, Gabija, Kiwa, Lamassu, Type III restriction modification, Class IE CRISPR, CBASS (our system), and Retron (as identified by the DefenseFinder server). We have attempted lambda-red genome engineering in this

strain to generate tagged or mutant versions of the CBASS operon, without success. This is perhaps not surprising, given the array of defense systems in the genome.

From a broader perspective, while we appreciate that using a native host is preferable in many cases, CBASS systems are known to be self-contained and portable. The exact system under study is found in over 150 sequenced bacterial strains, including 104 sequenced *E. coli* strains and other Enterobacteria. For our work, using the system in isolation, away from other confounding effects, is preferable.

The reviewer is correct that the genomic context of this system is likely important. We are currently pursuing work to test cooperation or synergy between this CBASS system and other defense systems in *E. coli* MS115-1. However, these experiments will take significant time and effort, and are outside the scope of the current structure and biochemical mechanism-focused work.

2) The authors show that the regulation of CBASS operon does not influence phage infection, yet they suggest that this is a defense operon. They further claim that the CBASS basal expression, undetectable by Western blot, is sufficient to provide maximal defense. This is an unsubstantiated claim, and the authors need to monitor the expression of the system by other methods (e.g. RT-PCR for all genes in the operon).

The data presented in **Figs 1** and **EV2** clearly shows that the (modest) anti-phage activity of the CBASS system is not affected by a deletion of *capP*. Thus, our claim that basal expression of CBASS is sufficient for this level of protection is completely justified. Higher levels of expression (as in our earlier study with this operon in an IPTG-inducible vector) can of course provide stronger protection, but expression increases induced by CapH and CapP are too slow to make a meaningful difference in phage defense in our current experimental setup.

To gain additional information on CBASS expression and the timing of expression changes in response to both phage infection and DNA damage, we have used qRT-PCR as suggested by the reviewer to track *cdnC* mRNA levels compared to a control RNA polymerase gene, *rpoA.* With this assay, we can show that cdnC is expressed at around 2% of the *rpoA* control in unperturbed cells (ΔCt = -4.8), a very low level but not zero. We also observe that phage infection induces measurable changes in *cdnC* mRNA levels as soon as 40 minutes after infection, while the transcriptional response to DNA damage is faster at 20 minutes (**Fig EV1E**). These data suggest that the expression changes in *cdnC* that we observe upon phage infection are due to DNA damage induced by the phage.

It will be also beneficial to use a cGAS inactive mutant or expressing a catalytically dead NucC to validate that the effect is directly mediated by the system. In addition, it is critical to test bacterial growth and viability, with and without the CBASS carrying plasmid to rule out an indirect influence on infection.

We now include plaque assays for the cGAS (CdnC)-dead and NucC-dead versions of the operon in **Fig EV1C-D**, which confirm that the observed protective effect is due to CBASS function. We have also performed growth curves with and without infection, with the data shown in **Fig EV2C**. These data show that none of the tested plasmids cause bacterial viability or growth defects on their own.

3) Based on figure 1I, the authors concluded that cells harboring WT and ΔcapP CBASS have similar DNA degradation kinetics. This is very unclear to me; why do cells having such a difference in the nuclease expression display similar abortive infection kinetics?

> Our data shows that the substantial increase in NucC protein levels does not occur until ~90 minutes after phage infection. As our microscopy data (now in **Fig EV2A-B**) and our prior studies with this system show, cell death occurs earlier at around 60 minutes when the cells harbor the NucC-containing CBASS system. As we state in the manuscript, this result means that the basal level of NucC is sufficient to mediate genome degradation and cell death.

Figure 1I, is inconclusive, the images are not clear, very few cells are shown for the CBASS carrying strains, and the DAPI staining is heterogeneous from the beginning. The authors should quantify the data, provide clearer images, and use additional methodologies such as DNA extraction from infected cells or nick end labeling (TUNEL) to substantiate their claim.

> In the revised manuscript we have now quantified DNA degradation as we did in our prior work (**Fig EV2B**). Since this manuscript is not about the mechanism of NucC (which we previously studied and is by now well established), we do not consider it necessary to perform further assays to demonstrate the mechanism of NucC-dependent cell death by DNA degradation.

4) The authors observed high nucC levels in the absence of capH. Does it impact cell viability? I would expect such high levels of a nuclease to have an impact on cell growth and viability.

> In our original submission, we noted that we were unable to clone a version of the CBASS operon lacking the *capH* gene, likely due to toxicity induced by high-level expression of the core CBASS genes. While we did report results from a Δ*capP*-Δ*capH* plasmid, more recent validation of this construct by whole-plasmid sequencing reveals that it contains a partial deletion of the *cdnC* gene, explaining why this construct was not as toxic to host cells as the Δ*capH* plasmid. Thus, unregulated CBASS core gene expression is unequivocably toxic to host cells.

In line with this, does the increase in nucC during DNA damage cause host DNA degradation? Could nucC be a repair enzyme rather than a non-specific nuclease?

> This is a good suggestion. To gain a quantitative sense of whether NucC expression sensitizes cells to DNA damaging agents (as would be expected from a destructive nuclease), we performed minimum inhibitory concentration (MIC) assays using zeocin (**Fig EV1F**). Surprisingly, our data shows no significant difference in zeocin's MIC in the presence or absence of CBASS (WT or Δ*nucC*). While this data is preliminary and requires follow-up in later studies, it suggests that increased expression of CBASS genes is not sufficient to activate the system: activation may still require a phage trigger to activate second messenger synthesis by the CD-NTase and activation of NucC.
>
> There is no chance that NucC is a repair enzyme: we previously showed that it is a highly active, nonspecific nuclease that degrades DNA to ~50 bp segments.

5) The regulation observed is very similar to the classical RecA-LexA SOS response. Yet, the authors do not draw a clear line between the systems and their components; i.e. following binding to ssDNA, RecA is activated and facilitates the LexA repressor cleavage culminating in SOS gene activation.

We have added mention of the SOS response and its parallels to our observations, to both the introduction and discussion of the revised manuscript. We feel it important to point out that the molecular mechanisms of these two pathways are very different, even if they are both activated by single-stranded DNA.

6) The promoter region investigated in this study is not defined. Two promoters driving opposing transcription are located in the intergenic region between capH-cdnC. Please define and provide details.

As we note above, spurred by this comment and one from Reviewer #2, we more closely examined the CBASS promoter region both bioinformatically and experimentally. We have identified likely promoters for both forward and reverse expression, and define two high-affinity binding sites for CapH that overlap these promoter sequences.

7) Do additional *E. coli* genes contain the putative CapH binding motive described in Fig 2C?

Since we do not yet know the exact binding requirements for CapH, this analysis is not possible.

8) Why only fragments of the CapH were crystallized and not the entire protein?

We attempted to crystallize the full-length protein, but it did not form crystals. We suspect that the flexible linker region between the two domains prevents organized packing. This point is now noted explicitly in the Results section.

9) The DNA binding and oligomerization states of CapH (Fig 3) should be substantiated by gel shift analysis.

Because fluorescence polarization is an equilibrium method for measuring binding affinity of macromolecules, it is superior to EMSAs which are not an equilibrium method. Moreover, our detection of cooperative binding to DNA, and the demonstration that CapH-I99M still shows cooperative binding, indicate that CapH functions primarily as a dimer, rather than a tetramer. We did use EMSA to confirm that wild-type CapH shows a super-shifted species that CapH(I99M) does not show (Figure S4C), but our other data show that CapH tetramerization does not play a major role in DNA binding.

10) Fig 1G: The authors should show host viability and infection kinetics simply by OD600 measurements. The Y axis was defined as "plaque forming units per mL of purified phage", could the authors describe the procedure of phage purification?

We now show growth curves during infection in **Fig EV2C**. We also now include phage purification protocols (which are standard) in the **Methods** section.

11) The results for ΔcapH should be included in Figure 1E-1I.

As noted above, we could not generate a Δ*capH* version of the full CBASS system due to toxicity. The Δ*capH* construct noted in Figure 1 could only be generated in the GFP reporter system.

12) Figure 2F, R44A still acts as a repressor as appears in the gel. This seems inconsistent with the text (p6: "We found that mutation of Ser32 or Arg44 on the predicted DNA-binding face of CapH eliminated detectable DNA binding").

Now that we have better defined the binding sites for CapH, the in vitro and in vivo data for the CapH mutants is more consistent. Only S32A shows a strong effect on binding in vitro, and on reporter expression in vivo.

13) The figure legends and the methodologies lack essential details such as the bacterial strains and their genotypes, MOI, plasmid construction.

We have added new information about plasmid construction, bacterial strain genotypes, and viral MOIs.

14) The introduction lacks important information such as reviewing the SOS response.

As noted above, our data do not support a functional parallel with the SOS response beyond the very broad strokes concept of ssDNA activated proteolysis. Nonetheless, it is an interesting conceptual parallel that we discuss in the revised Discussion section.

15) p6: I think "CinR" should be "SinR".

We thank the reviewer for catching this error, and we have corrected this name.

1st Revision - Editorial Decision 6th Sep 2022

Thank you for submitting your revised manuscript for our consideration. It has now been assessed once more by the original referees 1 and 2. I am pleased to say that in light of their positive overall assessment, we shall be happy to publish the study in The EMBO Journal, following incorporation of a few remaining issues noted by the referees (see comments below).

In addition, please also address the following editorial points:

Referee #1:

The revised manuscript from Lau et al provides substantial further supporting evidence beyond the initial submission. As a result the manuscript is stronger and more robust. It is, in my opinion, ready for publication. Congratulations to the authors.

They have addressed all my comments except one minor point - in the intro, the appropriate ref is Picton et al 2021, as Picton et al 2022 deals with a different topic.

I have no further comments.

Referee #2:

I have carefully read the revised manuscript and responses from the authors and am satisfied with their responses to all three referees. The additional experiments and text revisions substantially improved the manuscript and it should be accepted for publication with some very minor revisions (see below). In my opinion this is a beautiful study.

A few minor comments (these do not change the acceptance of the manuscript):

In response to Reviewer 3: "While we did report results from a ΔcapP-ΔcapH plasmid, more recent validation of this construct by whole-plasmid sequencing reveals that it contains a partial deletion of the cdnC gene, explaining why this construct was not as toxic to host cells as the ΔcapH plasmid."
In figure EV1B, there is data for dcap/dcapH plasmid (lane 3)-is this lane valid--i.e. is the construct correct?

Please edit/update Appendix S1A figure & legends. Also be sure to match the call-outs in the text to the proper panel (e.g one of the "S1A"'s in the results should be S1B).I noticed these issues:
• Panel A appears to be described by figure legend E.
• Data described in figure legends B and C seems to be missing (maybe formatting issue)
• Figure legend D: "fluorescence polarization" should read "EMSA" or "gel shift"

Update methods to reflect sizes of DNA constructs used in fluorescence polarization experiments.

Responses to Editor and Reviewer Comments

Dear Dr. Vodermaier – We were pleased to receive the positive news regarding our manuscript, #EMBOJ-2022-111540R. Below we respond to each of your points, and those of the reviewers, in blue text.

Referee #1 Notes:

The revised manuscript from Lau et al provides substantial further supporting evidence beyond the initial submission. As a result the manuscript is stronger and more robust. It is, in my opinion, ready for publication. Congratulations to the authors.

They have addressed all my comments except one minor point - in the intro, the appropriate ref is Picton et al 2021, as Picton et al 2022 deals with a different topic.

We regret this oversight. We have changed the reference to Picton et al. 2021 as noted.

Referee #2 Notes:

I have carefully read the revised manuscript and responses from the authors and am satisfied with their responses to all three referees. The additional experiments and text revisions substantially improved the manuscript and it should be accepted for publication with some very minor revisions (see below). In my opinion this is a beautiful study.

A few minor comments (these do not change the acceptance of the manuscript):

In response to Reviewer 3: "While we did report results from a ΔcapP-ΔcapH plasmid, more recent validation of this construct by whole-plasmid sequencing reveals that it contains a partial deletion of the cdnC gene, explaining why this construct was not as toxic to host cells as the ΔcapH plasmid."
In figure EV1B, there is data for dcap/dcapH plasmid (lane 3)-is this lane valid--i.e. is the construct correct?

The reviewer is correct; we neglected to remove this lane from the figure when we removed reference to it in the text. We have updated Figure EV1 to remove this lane from the blot.

Please edit/update Appendix S1A figure & legends. Also be sure to match the call-outs in the text to the proper panel (e.g one of the "S1A"'s in the results should be S1B).I noticed these issues:

- Panel A appears to be described by figure legend E.
- Data described in figure legends B and C seems to be missing (maybe formatting issue)
- Figure legend D: "fluorescence polarization" should read "EMSA" or "gel shift"

We failed to update the legend for this figure after significantly revising it. We thank the author for catching these problems, and have updated the legend to fix the noted issues. We also changed the noted mistaken figure callout in the main text.

Update methods to reflect sizes of DNA constructs used in fluorescence polarization experiments.

We have updated the Methods section to properly reflect the DNAs used for the revised submission.

2nd Revision - Editorial Decision 8th Sep 2022

Thank you for submitting your final revised manuscript for our consideration. I am pleased to inform you that we have now accepted it for publication in The EMBO Journal.

EMBO Press Author Checklist

Corresponding Author Name: Kevin D. Corbett
Journal Submitted to: The EMBO Journal
Manuscript Number: EMBOJ-2022-111540R

USEFUL LINKS FOR COMPLETING THIS FORM
The EMBO Journal - Author Guidelines
EMBO Reports - Author Guidelines
Molecular Systems Biology - Author Guidelines
EMBO Molecular Medicine - Author Guidelines

Reporting Checklist for Life Science Articles (updated January
This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your
Please note that a copy of this checklist will be published alongside your article.

Abridged guidelines for figures
1. Data
The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical
- if n<5, the individual data points from each experiment should be plotted. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship guidelines on Data

2. Captions
Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements.
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.
- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple χ2 tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Please complete ALL of the questions below.
Select "Not Applicable" only when the requested information is not relevant for your study.

Materials

Newly Created Materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
New materials and reagents need to be available; do any restrictions apply?	Not Applicable	

Antibodies	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
For **antibodies** provide the following information: - Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number - Non-commercial: RRID or citation	Yes	Methods

DNA and RNA sequences	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Short novel DNA or RNA including primers, probes: provide the sequences.	Yes	Methods

Cell materials	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Cell lines: Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID.	Not Applicable	
Primary cultures: Provide species, strain, sex of origin, genetic modification status.	Not Applicable	
Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination.	Not Applicable	

Experimental animals	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Laboratory animals or Model organisms: Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID.	Not Applicable	
Animal observed in or captured from the field: Provide species, sex, and age where possible.	Not Applicable	
Please detail **housing and husbandry conditions**.	Not Applicable	

Plants and microbes	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Plants: provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens).	Not Applicable	
Microbes: provide species and strain, unique accession number if available, and source.	Yes	Methods

Human research participants	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants.	Not Applicable	

Core facilities	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If your work benefited from core facilities, was their service mentioned in the acknowledgements section?	Yes	Acknowledgements

Design

Study protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI.	Not Applicable	
Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable.	Not Applicable	

Laboratory protocol	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Provide DOI OR other citation details if **external detailed step-by-step protocols** are available.	Not Applicable	

Experimental study design and statistics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Include a statement about **sample size** estimate even if no statistical methods were used.	Yes	Figure legends and Methods
Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described?	Not Applicable	
Include a statement about **blinding** even if no blinding was done.	Not Applicable	
Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established? If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification.	Not Applicable	
For every figure, are **statistical tests** justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared?	Yes	Figure legends and Methods

Sample definition and in-laboratory replication	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
In the figure legends: state number of times the experiment was **replicated** in laboratory.	Yes	Figure legends
In the figure legends: define whether data describe **technical or biological replicates**.	Yes	Figure legends

Ethics

Ethics	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval.	Not Applicable	
Studies involving **human participants**: Include a statement confirming that **informed consent** was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	Not Applicable	
Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained.	Not Applicable	
Studies involving experimental **animals**: State details of **authority granting ethics approva**l (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations.	Not Applicable	
Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why.	Not Applicable	

Dual Use Research of Concern (DURC)	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm	Not Applicable	
If you used a select agent, is the security level of the lab appropriate and reported in the manuscript?	Not Applicable	
If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript?	Not Applicable	

Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

Adherence to community standards	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided.	Not Applicable	
For **tumor marker prognostic studies,** we recommend that you follow the **REMARK** reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	Not Applicable	
For **phase II and III randomized controlled trials**, please refer to the **CONSORT** flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	Not Applicable	

Data Availability

Data availability	Information included in the manuscript?	In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section)
Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section?	Yes	Methods and Data Availability
Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement?	Not Applicable	
Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided?	Not Applicable	
If publicly available data were reused, provide the respective **data citations in the reference list**.	Not Applicable	