Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type-III CRISPR-Cas that degrades the second messenger cA₄

Sharidan Brown¹, Colin C. Gauvin¹,², Alexander A. Charbonneau¹,², Nathaniel Burman¹ and C. Martin Lawrence¹,²*

¹Department of Chemistry and Biochemistry and the ²Thermal Biology Institute Montana State University, Bozeman, MT 59717, USA

*Corresponding author: Martin Lawrence
E-mail: lawrence@montana.edu

Running title: Csx3, a ring nuclease that regulates the cA₄ signal

KEYWORDS: cyclic tetra-adenylate, CARF, RNase, phosphodiesterase, fluorescence lifetime

Cas10 is the signature gene for type-III CRISPR-Cas surveillance complexes. Unlike type-I and type-II systems, type-III systems do not require a PAM, and target nascent RNA associated with transcriptionally active DNA. Further, target RNA recognition activates the cyclase domain of Cas10, resulting in the synthesis of cyclic oligoadenylate (cOA) second messengers. These second messengers are recognized by ancillary Cas proteins harboring CARF domains, and regulate the activities of these proteins in response to invading nucleic acid. Csx3 is a distant member of the CARF domain superfamily previously characterized as a Mn²⁺ dependent deadenylation exoribonuclease. However, its specific role in CRISPR-Cas defense remains to be determined. Here we show Csx3 is strongly associated with type-III systems and that Csx3 binds cyclic tetra-adenylate (cA₄) second messenger with high affinity. Further, Csx3 harbors cyclic oligonucleotide phosphodiesterase activity that quickly degrades this cA₄ signal. In addition, structural analysis identifies core elements that define the CARF domain fold, and the mechanistic basis for ring nuclease activity are discussed. Overall, the work suggests Csx3 functions within CRISPR-Cas as a counter balance to Cas10 to regulate the duration and amplitude of the cA₄ signal, providing an off-ramp from the programmed cell death pathway in cells that successfully cure viral infection.

CRISPR-Cas is a prokaryotic adaptive immune system that is now well known for its repurposed uses in genome editing. Within the prokaryotic domains of life where these systems evolved, a multitude of various classes, types and subtypes are now recognized. In each case, however, the basic mechanism of adaptive immunity is generally described in three distinct stages. Briefly, in stage 1 (adaptation) short fragments of invading DNA known as spacers are inserted into the CRISPR locus. During stage 2 (crRNA biogenesis) the CRISPR locus is transcribed and cleaved into short CRISPR RNAs (crRNAs), and in stage 3 (target interference) crRNA is assembled into nucleoprotein surveillance complexes that survey the cell for complementary DNA (type-I, type-II) or RNA (type-III). Recognition of target DNA or RNA then activates nucleases that destroy the invading nucleic acid, providing immunity. Further, in type-III systems, which recognize RNA associated with transcriptionally active DNA, not only is the RNA degraded, but the DNA as well (1,2).

This process is wholly dependent on a wealth of CRISPR associated (Cas) protein...
machinery. Cas1, Cas2 and Cas4 are generally involved in adaptation (spacer acquisition), Cas6 frequently in crRNA biogenesis, and diverse groups of type and subtype dependent proteins are utilized in the various stage 3 (target interference) surveillance complexes. Class 1 systems generally utilize multiple Cas proteins to assemble their respective surveillance complexes. Among class 1 systems, the type-I and type-III systems are the most common, and can be differentiated by the presence of Cas3 and Cas10, respectively. Class 2 systems, on the other hand, generally utilize a single large protein subunit to assemble the surveillance complex, and among the class 2 systems, the type-II systems which utilize Cas9 are the most common. In addition, a number of “ancillary” proteins not directly implicated in the above processes are also clearly involved. Among these, proteins containing the CRISPR associated Rossmann Fold, or CARF domain (3) are central to this work.

In 2011 Lintner et al. studied two proteins with N-terminal CARF domains, Csa3 and Csx1, concluding these dimeric CARF domain proteins contained a single ligand binding site that spanned the dimer interface, and that the activities of these proteins would be regulated by a 2-fold (pseudo-) symmetric signal, potentially a cyclic oligonucleotide (4). This hypothesis was subsequently echoed and expanded upon by others, as they identified HEPN nuclease domains in Csx1 and Csm6 similar to those in toxin-antitoxin systems, and suggested the unknown signal might activate these non-specific RNases. Non-specific degradation of cellular mRNAs by Csx1 and Csm6, in turn, are expected to deplete the cellular mRNA pool and inhibit protein synthesis. This, in turn, may result in the onset of cell dormancy until the infection can be cured, or even induce programmed cell death (3,5).

In 2015 Yan et al. demonstrated manganese dependent 3’-deadenylase activity for yet another, apparently unrelated, ancillary Cas protein known as Csx3 (6). They also determined crystal structures of Csx3 in the free form (3WZG), in the presence of Mn2+ (3WZH), and in complex with a single stranded RNA fragment (3WZI). Interestingly, the 4-base RNA fragment was found distal from conserved active site histidines on the opposite face of this single domain protein. Topuzlu et al. then recognized Csx3 as a distant member of the CARF domain superfamily, and that the 4-base RNA fragment was bound in a nearly circular fashion in the putative regulatory binding site, suggesting a cyclic 4-base RNA as a regulatory ligand for CARF domain proteins (7).

In beautiful work, this hypothesis was then confirmed by both Kazlauskienė et al. (8) and Niewohner et al. (9). They showed that when type-III surveillance complexes bind target RNA, the Cas10 cytosine domain synthesizes small oligoadenylates, including cyclic tetra- and hexa-adenylate (cA4 and cA6). Further, they also showed these products are recognized by the N-terminal CARF domains of Csx1 and Csm6, where they do indeed activate latent RNase activity within their C-terminal HEPN domains, consistent with the hypothesis that these cOA signals promote cell dormancy, or in the case that the cOA signal is not attenuated, programmed cell death.

Thus, what is the end fate of this cOA; are cellular cOA levels controlled reversibly? The answer to this conundrum was recently elucidated by Athukoralage et al., who found that some archaea express a “ring nuclease” that degrades the cA4 signal into two molecules of A2 (10). Interestingly, this ring nuclease, christened “CRISPR ring nuclease I” (Crn1), also utilizes a CARF domain for cA4 recognition, but with catalytic residues installed to slowly degrade the cA4 signal.
Overall, these advances in understanding the role of cyclic oligonucleotide signals in CRISPR-Cas prompted us to further investigate Csx3 and its potential interaction with the second messenger cA4.

Results

Csx3 associates with Type-III CRISPR-Cas Systems – In early seminal work on CRISPR-Cas, Haft et al. identified 45 different CRISPR associated (cas) gene families (11). These included core cas genes common to multiple CRISPR types, as well as modular groups of cas genes now known to act in concert. However, clear contextual patterns for some cas families were elusive, and these cas genes were assigned to CRISPR subtype X, or csx. Archaeoglobus fulgidus (DSM4304) possesses both Type-IA and Type-IIIB CRISPR-Cas systems, with csx3 clearly embedded within the Type-IIIB cassette (12). And while it has been noted that type-III systems may include ancillary csx3 genes (13), it is not clear if the reverse is true, i.e., that csx3 is most commonly associated with type-III systems, and if so, in what context. We thus identified putative AfCsx3 orthologs using a protein-protein BLAST search (14) against the non-redundant database. The top 46 unique sequences ranged from 84 to 113 residues in length and showed expect values < 6 x 10^-17 with greater than 45% sequence identity to AfCsx3 over at least 75% of the query. Sequences below this threshold included larger, multi-domain proteins that, while containing a Csx3-like domain, are not clear homologs and thus were not included. The sequence data was present in different forms, coming from 17 complete genomes, 7 scaffolds and 23 contigs. While the genomic context was not always clear, especially within contigs, csx3 was present in a type-III A cassette in 14 cases and in a type-IIIB cassette in 13 other cases. In just a single case was csx3 found outside a type-III system, where it was instead adjacent to type-I cassette. In this case, csx3 was at the extreme 5'-end of the contig, and we were unable to analyze the neighboring upstream sequence.

When the genes immediately up- or downstream of csx3 were surveyed, members of the csm6/csxl family of CARF domain HEPN RNases were the most common, followed by cmr6, cmr4 and cas10 (Supplemental Figure 1). In contrast, except for a lone csa8a gene (type-IA), an immediate type-I or type-II neighbor was not observed. These single domain Csx3 orthologs thus show a clear association with type-III CRISPR-Cas systems. Because type-III (Cas10) surveillance complexes synthesize cOA signals when bound to target RNA, the association of Csx3 with type-III systems is consistent with the proposed recognition of cA4.

Conserved sequence motifs in Csx3 – We next looked for conserved sequences motifs in AfCsx3. A multiple sequence alignment performed with Clustal Omega (15) identified twelve strictly conserved residues (Supplemental Figure 1). These include two isolated residues, Pro33 and His80, and three short motifs; GR(G/A)P(L/F)W50, LxHxxH60 and DPR71. When mapped to the AfCsx3 structure, His80 and the LxHxH60 motif localize to the RNase active site identified by Yan et al. (6). In contrast, the GR(G/A)P(L/F)W50 and DPR71 motifs map to the tetra-base RNA binding site on the opposite face of the Csx3 dimer (6), and are structurally equivalent to the conserved motifs previously identified in the CARF domains of both Csa3 and Csx1 (4). Thus, the putative cA4 binding site is strictly conserved among this group of Csx3 orthologs.

Csx3 binds cA4 with nanomolar affinity – The presence of Trp50 among the conserved residues [GR(G/A)P(L/F)W50] in the distal RNA binding site suggested intrinsic tryptophan fluorescence might be used to monitor ligand binding at this site. We utilized fluorescence lifetime spectroscopy
for this purpose, taking advantage of an in-house time-resolved fluorescence spectrophotometer to collect subnanosecond-resolved fluorescence emission waveforms. This approach clearly demonstrates a very tight interaction between Csx3 and cA4 (Figure 1, Supplemental Figure 2). We measure an apparent $K_d$ for cA4 of $55 \pm 4.5$ nM. Notably, this value is less than the concentration of the Csx3 dimer (150 nM) used in the assay, explaining the near stoichiometric (linear) signal at cA4 concentrations below 150 nM. Ideally, the assay would utilize Csx3 concentrations below the $K_d$, but the sensitivity of the assay does not currently allow this. The apparent $K_d$ thus provides only an upper limit, and could be an underestimate of the true affinity. For perspective, a single cA4 molecule per prokaryotic cell, approximate diameter of 1 µm, gives a concentration of ~100 nM. The affinity for cA4 is thus highly relevant, allowing recognition even at concentrations corresponding to a single molecule per cell. Clearly, Csx3 harbors a physiologically relevant cA4 binding site.

We next asked how the affinity for cA4 compares to the affinity for one of the poly-A tailed oligonucleotides originally utilized by Yan et al. For this, we chose the shortest of these, 5'-GGGAAAAGAAAAAA-3' [Supplemental Table 1, “A6” of Yan et al. (6)]. Here, the Csx3 fluorescence lifetime binding assay reveals a much lower affinity, giving a $K_d = 1.9 \pm 0.2$ µM (Figure 1, Supplemental Figure 2). The significantly lower affinity for “A6” provides further evidence that cA4 is indeed a physiologically relevant ligand.

The Csx3 fold – To elucidate the structure of Csx3/cA4 complex and conformational changes in in the Csx3 homodimer upon binding, we crystallized Csx3 in the presence of cA4. Csx3 crystallized in space group I222 with a single subunit in the asymmetric unit, and usable diffraction data to 1.8 Å resolution. The structure unexpectedly revealed a 7th C-terminal β-strand not present in previously modeled structures of AfCsx3 that runs anti-parallel to the first strand (Figure 2). Further, the additional β-strand is preceded by a C-terminal α-helix (α3) that while present in the previously determined RNA bound structure (3WZI), is absent from the earlier apo- (3WZG) and Mn$^{2+}$-bound (3WZH) structures (6). Thus, in contrast to previous crystal structures, the structural core of the Csx3 protomer is a mixed 7-stranded β-sheet with $\beta_7\uparrow-\beta_1\downarrow-\beta_2\uparrow-\alpha_1-\beta_3\uparrow-\alpha_2-\beta_4\uparrow-\beta_5\downarrow-\beta_6\uparrow-\alpha_3$ topology, in which the α-helices form right-handed crossovers between the indicated β-strands (α3 connects β6 to β7). This places α1 and α2 on one face of the sheet, and α3 on the opposite side. In the context of the dimer, α1 and α2 are buried at the subunit interface, while α3 is surface exposed.

It was unclear whether the presence of the additional β-strand was due to the higher resolution of the data set, or whether the C-terminal tail is a dynamic structural element. To address this, we downloaded the 3WZI, 3WZG and 3WZH models and structure factors from the protein data bank. In addition to these 3 models, we generated 3 additional models in which coordinates for residues 98-104 from our superposed I222 structure were included. For 3WZI, where the C-terminal tail is modeled, we also deleted the corresponding residues (98-104) of the original structure. For each of the six models, the bulk of the model is already well refined, and the only difference is whether the I222 β-sheet is present or absent. We then refined each of the 6 structures against the corresponding set of structure factors. In each case, the resulting $R$-factors were significantly lower (lower) for the model lacking β7. Specifically, the respective $R_{work}/R_{free}$ values for 3WZG, 3WZH and 3WZI in the absence of the I222 tail are 20.57%/23.63%, 21.71%/25.69%,
20.97%/22.81%, respectively, but, when the I222 tail is included in the refinement, the R-factors rise to 21.20%/24.66%, 22.61%/26.74% and 24.45%/26.29%. The presence of the β7 tail is thus inconsistent with the observed structure factors. In addition, the resulting maps also lack density for the I222 tail, and when the I222 tail is included in the models, instead shows strong negative difference density. This clearly suggests the C-terminal tail is indeed mobile, and while present as a seventh beta-stand in the new I222 structure reported here, adopts alternate conformations in the previous structures.

Further examination of this I222 structure revealed that β7 is anchored not only by hydrogen bond interactions as it runs antiparallel to β1 of the same subunit, but that it is also stabilized by a parallel β-strand interaction with β6 of a neighboring subunit (Figure 3). Interestingly, this results in formation of an extended, inter-subunit β-sheet running the length of the crystal. Visual observation and PISA interface analysis software (16), however, indicate this inter-subunit crystal contact is insufficient to support formation of a stable protein interface in solution. While such open symmetry is common in filamentous protein assemblies, it’s atypical for enzyme complexes, which more commonly show closed point group symmetry. Regardless, the dynamic nature of the C-terminal tail and its ability to mediate inter-dimer interactions is reminiscent of 3D domain swapping (17) and may indicate the ability of β1, β6 and the C-terminal tail to mediate additional protein-protein interactions within the cell.

Although Csx3 was co-crystallized with cA4, density for cA4 was not observed in the RNA binding pocket, or elsewhere. While Csx3 is a Mn2+ dependent deadenylase, our crystallization conditions did not contain added divalent cations. Thus, given the extreme affinity of Csx3 for cA4 and the previous structure of Yan et al. with an unidentified RNA fragment (3WZI), this was unexpected. On the other hand, crystallization conditions did not include chelators and trace metals could have enabled cA4 degradation by the RNase active site. Alternatively, self-regulating metal independent ring nuclease activity has been seen for select CARF domains (18-20), suggesting Csx3 might also harbor metal independent, cA4 specific ring nuclease activity at the cA4 binding site, as opposed to the metal dependent RNase active site present on the opposite face of the protein.

Ring nuclease activity - For these reasons, we examined the ability of Csx3 to degrade cA4. The activity assay utilized thin layer chromatography to separate unlabeled cA4 (Biolog Life Science/Axxora) from potential degradation products. Using this assay, we observed unexpectedly high levels of ring nuclease activity upon the addition of 200 μM Mn2+ (Figure 4), where the thin layer chromatography was consistent with diadenylate (A2) as the major product (10). At times trace activity levels were observed in samples lacking added Mn2+, perhaps due to low-level copurification of Csx3 with divalent cations, and potentially explaining the lack of bound cA4 in the crystal structure. In general, this activity was inhibited upon addition of EDTA. Importantly, when the progress curve in the presence of 200 μM Mn2+ was modeled as first order exponential decay, the average observed rate constant (kobs) from 5 independent experiments was 4.95 ± 0.43 min⁻¹. For perspective, the rate constant for CRISPR ring nuclease Crn1 identified by Athukoralage et al. in 2018 is 0.089 min⁻¹ (21,22), more than an order of magnitude slower than Csx3 in the presence of Mn2+. From this perspective Csx3 demonstrates robust cyclic oligonucleotide phosphodiesterase, or “ring nuclease” activity. This suggests cA4 is indeed a physiologically relevant substrate and that
Csx3 functions within CRISPR-Cas to regulate the amplitude and duration of the cA4 signal.

Discussion

Defining elements of the Cyclic oligoAdenylate Recognition Fold (CARF) -

With confirmation that Csx3 does indeed bind cyclic tetra-adenylate with high affinity, we can now compare and contrast the structure of this more distant member of the CARF domain superfamily with the canonical CARF domain fold. This is well represented by the relatively simple *S. solfataricus* Csa3 CARF domain (4), which adopts the classic dinucleotide binding domain fold with one simple modification; the right handed helical crossover between $\beta_5$ and $\beta_6$ is lost. The two $\beta$-strands are instead connected by a reverse turn, with $\beta_6$ running antiparallel rather than parallel to $\beta_5$. This modification has significant consequences. As Csa3 forms the homodimer, the helices connecting $\beta_4$ and $\beta_5$ in each subunit pack against and parallel to each other at the center of the dimer interface, with each assuming the place of the “missing” helix in the neighboring subunit. The $\beta_5$-$\beta_6$ reverse turn and loss of the connecting helix is thus responsible for creation of the CARF domain dimer interface (Figure 5). And because dimer formation is inherent to recognition of the 2-fold symmetric cOA signals, this is a defining characteristic of the CARF domain fold. In short, the $\beta_4$$\alpha_4$$\beta_5$$\beta_6$ super-secondary structural element in the Csa3 CARF domain lies at the heart of the fold. Critically, these structural elements in Csa3 (Csm6 and Csx1 as well) are topologically analogous to the $\beta_3$$\alpha_2$$\beta_4$$\beta_5$ elements in Csx3, and within the Csx3 dimer they play the same central role at the dimer interface as the cognate elements in Csa3 and other CARF domain proteins.

As the CARF domain acronym (CRISPR Associated Rossmann Fold) went into use before the ligands for this domain were fully identified, it does not well convey the function of this fold. Now that it is understood to bind small cyclic oligonucleotides, it is tempting to give it a more meaningful name. However, the CARF acronym is now in such wide use that renaming it could be difficult. We note however, that this domain clearly serves as a Cyclic oligo-Adenylate Recognition Fold, and this mnemonic could help some to remember its function. One should keep in mind, however, that it also recognizes small linear polynucleotides (6), and it is difficult to imagine that nature has not also adapted this fold for the recognition of other bases.

The C-terminal tail – The previous structures of Csx3 by Yan et al. show a disordered C-terminal tail. In the structure reported here, the C-terminal tail is instead captured in an ordered conformation. The structure shows a well-defined C-terminal $\alpha$-helix that traces a path along the “back side” of the subunit to position the extreme C-terminus to add an additional $\beta$-strand along the outside edge of the first $\beta$-strand (Figure 2). Thus, we see that the C-terminal tail not only has a propensity to form these secondary structural elements, but to utilize them, especially $\beta_7$, to interact with other secondary structural elements within Csx3. However, the dynamic nature of the tail suggests that within the cell, the C-terminal tail might instead be utilized to interact with other cellular components, potentially as a member of a larger complex that might modulate the activity of Csx3. Importantly, in this conformation both $\beta_1$ and $\beta_7$ are available for intersubunit interactions.

Recognition of cA4 - The core $\beta_3$$\alpha_2$$\beta_4$$\beta_5$ structural motif is functionally important as well. The conserved GR(G/A)(L/F)W$^{50}$ and DPR$^{71}$ motifs identified above lie in the $\beta_3$$\alpha_2$ and $\beta_4$$\beta_5$ loops, respectively, and play a significant role in construction of the Csx3 RNA binding pocket (Figure 5). Similarly, conserved
residues are also found at these positions in Csa3 and Csxl, where they are also involved in construction of the conserved ligand binding site (4).

Additional insight into recognition of cA4 by Csx3 can be gained by re-examining the RNA density in the 3WGI structure determined by Yan et al (6). The 3WGI crystals were grown in the presence of a 37 nucleotide single-stranded RNA molecule, but Yan et al. report that the RNA was degraded over the time course for crystal growth. The 4 ordered nucleotides in the CARF domain binding site were thus modeled as a linear 4-base RNA fragment.

However, perhaps because the density arises from the presence of RNA fragments bound in overlapping conformations with different register, the electron density map clearly shows a cyclic ring of electron density. This suggested the density might be used to inform efforts to model the Csx3/cA4 interaction. Indeed, when we build cA4 into this density, our minimally refined cA4 bound model gives R_work and R_free values of 19.2 and 22.4% respectively, which compares very favorably to the original model with the linear 4-base RNA fragment.

As expected, this model shows the conserved GR(G/A)P(L/F)W50 and DPR71 motifs in the β3α2 and β4β5 loops play a central role in cA4 recognition. In addition to these elements, the β2α1 loop is also prominent, corresponding to the conserved β1α1 loop of Csa3. Analysis of the modeled interaction shows that cA4 buries greater than 800 Å² of protein surface. Base A1 (and A3 in the second subunit) is nestled in a hydrophobic pocket between Arg46, Tyr68 and Pro70, with potential hydrogen bonds from the main chain NH of Gly45, the His15 side chain and Arg46 to N1, N6 and N7 of the adenine ring, respectively. Base A2 (and A4) stacks on top of Pro48 and the edge of Trp50, with N6 donating a hydrogen bond to the main chain carbonyl of Ile25. The phosphate groups in this model lack substantial interactions, though Arg46 and Arg71 could seemingly interact strongly with a relatively minor reorientation of their side chains. Notably, the phosphate groups adjacent to the Arg46 side chains are elevated above the plane of the cA4 ring while those adjacent to Arg71 are more deeply buried. Finally, strictly conserved Asp69 hydrogen bonds to the 2’-OH of the ribose of A2 (and A4).

Structural basis for ring nuclease activity - Yan et al. quite logically proposed a tethered substrate model for Csx3 in which the RNA was bound in the RNA binding site on one face of the Csx3 dimer, and then extended into the active site on the opposite side of the subunit for cleavage. However, this model cannot explain the degradation of cA4, which is far too small. The CRISPR associated HEPN RNases such as Csx1 and Csm6 are two domain proteins that utilize an N-terminal CARF domain to bind cyclic oligoadenylates that then activates the C-terminal HEPN RNase domains. In some cases, these CARF domains have catalytic residues installed in the cA4 binding site that slowly degrade the cA4 signal in a metal independent mechanism (18-20). In this sense, the activities of these ancillary Cas proteins are “self-limiting”, much like the classic picture for slow GTP hydrolysis by G-proteins.

We might consider a similar model for Csx3. The crystal structure suggests Asp69 in the cA4 binding site could potentially act as a general base to deprotonate water or the ribose 2’-OH, catalyzing nucleophilic attack on the neighboring phosphorus to cleave cA4, resulting in a product with 3’ phosphate groups. However, this mechanism fails to explain the effect of Mn²⁺ on the reaction, which instead points to the histidine and
metal rich RNase active site on the other face of the molecule (6). One alternative solution is to invoke a metal independent RNase active site, though this is inconsistent with the data of Yan et al. (6) and our own ring nuclease data. Another possibility is a mechanism that utilizes both sites. Whether this might involve activation of the metal dependent active site (6) by binding of cA4 at the allosteric regulatory binding site, or some other mechanism is not clear. The structure of the RNase active site seems largely unchanged by the presence of the 4-base RNA fragment, but this fragment was a degradation product, which would not be expected to activate the CARF domain. In any event, we currently lack a satisfying explanation for this activity and additional work is clearly needed.

Utility of unlabeled cA4 ring nuclease and fluorescence lifetime binding assays – Previous work on CARF domain proteins has employed radiolabeled substrates and ligands to measure both ring nuclease activity and binding affinity for cA4. Our work, in contrast, has utilized unlabeled cA4 for both the ring nuclease and fluorescence lifetime binding assays. The “cold” ring nuclease assay reported here is easily adapted for the study of ring nuclease activity in general, and thus of general use to the field. And while the fluorescence lifetime binding assay is specific to proteins containing a tryptophan residue in the ligand binding site, Trp50 is strictly conserved among the Csx3 orthologs identified in this study and is thus applicable to Csx3 in general. Notably, the assay benefits from the stoichiometry of the interaction. With a single cA4 molecule bound across the dimer interface, the change in the fluorescence waveform arises from two copies of Trp50 for each cA4 bound. This label free binding assay could be utilized for the development of CARF domain agonists and antagonists, which may have potential therapeutic applications in pathogenic bacteria carrying type III CRISPR-Cas systems or other CARF domain proteins.

Physiological Specificity for cA4 – In a significant advance, Csx3 was previously described as a manganese dependent deadenylation ribonuclease, with specificity for poly-A tails in single stranded RNAs (6). While a rate constant for this activity was not reported, we note the assay conditions utilized a 2:1 ratio of Csx3 (20 μM) to substrate RNAs (10 μM), and that under these conditions, significant uncleaved substrate is still seen at the 5-minute time point. In contrast, we have instead used cA4 in excess, with a Csx3:cA4 stoichiometry of 1:15 (10 μM Csx3/150 μM cA4), and under these conditions, cA4 is largely degraded within 5 minutes. This simple comparison suggests Csx3 shows at least 20 times greater specific activity for cA4 than it does for small poly-A tailed RNAs. Further, the dissociation constant for the poly-A tailed RNA reported here indicates Csx3 shows ~50 fold greater affinity for cA4 than this small poly-A RNA. Considering both the 50-fold lower Kd and the 20-fold greater catalytic activity, the overall catalytic efficiency of Csx3 is at least 1,000-fold greater for cA4 than it is for the poly-A tailed RNA substrate. These are strong arguments that cA4 is indeed the physiologically relevant substrate, and that Csx3 functions within the cell as a cyclic oligonucleotide phosphodiesterase, or ring nuclease.

Additional evidence for a physiological role for Csx3 as a ring nuclease comes from consideration of the activities seen for other CRISPR associated ring nucleases (10). Specifically, Csx3 ring nuclease activity is an order of magnitude greater than that seen for Crn1, and is approaching the level of activity seen for virally encoded anti-CRISPR (Acr) ring nucleases (21). If anything, the activity is perhaps higher than expected (23), and its activity may need to be regulated at some level if present in the cell at significant
concentrations. In this regard, the ability of the dynamic C-terminal tail to mediate interactions within a larger complex may be extremely relevant.

Overall then, Csx3 possesses the ability to recognize cA4 at levels as low as a single molecule per cell and is also well endowed with ring nuclease activity. Once an infection is cured, and Cas10 ceases production of cA4, these attributes would allow it to quickly and effectively scrub the cell of cA4. Csx3 thus has all the attributes needed to serve as an effective counterbalance to cA4 production by Cas10, allowing it to regulate the duration and amplitude of the cA4 signal. In cases where the infection is cured, the cell can then return to normal metabolic activity, providing an off-ramp from the programmed cell death pathway. Importantly, ring nucleases have yet to be identified in many organisms harboring type III systems. The assignment of this activity to Csx3 thus represents a significant advance in our understanding of CRISPR-Cas in cells utilizing type III CRISPR-Cas systems.

**Experimental procedures**

**Gene Annotations** - Protein accessions were matched to their respective genome accessions. In the case of “non-redundant” Csx3 sequences, a representative genome with the highest level of completion was analyzed. In cases where a contig or scaffold were the only available information, analysis was restricted to the assembled read. Genome accessions were downloaded from the GenBank database with annotations from NCBI. In each case, csx3 was already annotated. Genomes were then opened and viewed by eye using the Unipro UGENE software (24). For genomes in which the neighboring genes were already annotated, those annotations were used. In the event that neighboring genes lacked annotation, the amino-acid annotation was analyzed with HHpred (25). Identification was considered conclusive only when the E-value was <0.001 and probability was greater than 95%.

**Csx3 expression and purification** - The native coding sequence for AfCsx3 with a minimal N-terminal His6-tag (Met-His6) was purchased from Thermo Fisher Scientific and moved into Gateway plasmid pDEST14 using previously described protocols (26) to give the pDEST14-AfCsx3 expression vector. For protein expression, *E. coli* BL21 (DE3)-pRIL (Stratagene) was transformed with pDEST14-AfCsx3. Typically a single colony was used to inoculate a 5 ml overnight LB starter culture (100 μg/ml ampicillin and 34 μg/ml chloramphenicol), which was then used to inoculate ZYP-5052 autoinduction media (27) (1:1000) and grown at 20°C for 48 hours. Cells were harvested by centrifugation at 4,000 x g for 30 min and 5-7 g pellets were stored at -20°C until needed. Cell pellets were then thawed and resuspended in 5 mL lysis buffer (20 mM Tris-Cl pH 7.5; 500 mM NaCl; 5% glycerol; 1 mM PMSF) per gram of cell pellet and lysed by French Press. The lysate was centrifuged at 25,000 x g for 30 minutes and the supernatant applied to a Ni-NTA affinity column, which was then washed with lysis buffer plus 20 mM imidazole and eluted with lysis buffer plus 500 mM imidazole. Fractions containing Csx3 were then combined and applied to a Superdex S-75 column equilibrated with 10 mM Tris-Cl pH 7.5, 500 mM NaCl, 5% glycerol. Protein concentrations were determined by Bradford assay (28) using Protein Assay Reagent (Bio-Rad) and BSA as a standard. The purity and molecular weight of Csx3 were confirmed by SDS-PAGE. Standard yields were approximately 25 mg of pure AfCsx3 from a 5 g cell pellet. Fractions containing pure Csx3 were combined and aliquots at 1 mg/ml were stored at -20°C.

**Fluorescence lifetime binding assays** - cA4 was titrated (0 – 750 nM) across a 96-well quartz microplate (Hellma) containing...
150 nM Csx3 dimer in 20 mM Tris-Cl pH 7.5, 100 mM NaCl, and 1.0% glycerol (final concentrations) with total well volumes of 350 µl, with three technical replicates per plate for each experiment. Binding assays were also performed with Poly-A Tail RNA (5'-GGGAAAAGAAAAAA-3') as described above using 300 nM Csx3 dimer and 0 – 20 µM RNA. Wells were excited at λ = 295 nm and time-resolved emission of individual wells followed at λ = 350 nm using a NovaFluor fluorescence lifetime spectrometer [Fluorescence Innovations, (29,30)]. Waveforms were recorded over a total decay time of 128 ns at 2 ps intervals. The fluorescence lifetime data were analyzed as described in supporting information of Schlick et al. (31) and more recently, and perhaps more succinctly, in supplemental of Bernhard et al. (32). Briefly, the waveforms collected from each well were fit to a linear combination of the free (\(W_f\)) and complexed (\(W_c\)) waveforms, yielding the fraction of protein in complex with ligand at each specific cA4 concentration. The data were then fit to the “one-site specific binding” function in GraphPad Prism (\(Y = B_{max}X/(K_d + X)\)).

**Cyclic oligonucleotide phosphodiesterase assays** – “Ring nuclease” assays were adapted from those of Athukoralage et al., but instead utilized unlabeled cA4 (10). Briefly, cA4 (c-tetraAMP, BioLog Lifescience Institute, final concentration 150 µM) was added to reactions containing 10 µM AfCsx3 in 10 mM Tris-Cl pH 7.5, 200 µM MnCl₂, 500 mM NaCl and 5% glycerol (final concentrations) plus SUPERase•In inhibitor and incubated at 50 °C. 10 µl aliquots were removed from the reaction mixture and quenched with equal volume phenol-chloroform at the desired time points. Upon quenching, 20 µl HzO was added to each aliquot to bring the sample volume up to 30 µl aqueous and 10 µl organic. 20 µl of the aqueous portion were then removed and placed in a SpeedVac for 30 minutes or until entirely evaporated. Samples were then re-
suspended in 20 µl methanol, centrifuged in a microcentrifuge at 14,000 RPM for 10 minutes to concentrate any precipitate, and the top 15 µl was removed and again placed in a SpeedVac for 30 minutes or until evaporated, re-suspended in 5 µl methanol, and spotted onto a glass-backed F254 silica gel TLC plate. The TLC plate was developed in 70% ethanol and 0.2 M ammonium bicarbonate (pH 9.3) until the mobile phase reached approximately 75% of the plate height. Plates were dried and imaged with an F254 lamp and imaging box. The intensities of the residual cA4 bands were quantitated using ImageJ software and normalized to time = 0. Using GraphPad Prism, the data were then fit to progress curves as described by Sternberg et al. (40) to determine the first order rate constants (kobs).

**Data availability**

The Csx3 model and associated X-ray diffraction data (structure factors) have been deposited in the Protein Data Bank (6VJG). All additional data are contained within the manuscript and supplemental information.

**Funding and additional information**

Research reported in this publication was supported by the National Science Foundation grant MCB-1413534 to CML. SB acknowledges support by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM103474. Fluorescence assays were collected at Montana State University on instruments funded by the National Science Foundation grant MRI-0821802. X-ray diffraction data were collected in the Macromolecular Diffraction Laboratory at Montana State University supported by the MJ Murdock Charitable Trust and the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Eiger 16M detector on 24-ID-E is funded by a NIH-ORIP HEI grant (S10OD021527). This research also used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

**Conflict of interest**

The authors declare that they have no conflicts of interest with the contents of this article.

**References**

1. Sorek, R., Lawrence, C. M., and Wiedenhoff, B. (2013) CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu Rev Biochem* **82**, 237-266
2. Hille, F., Richter, H., Wong, S. P., Bratovic, M., Ressel, S., and Charpentier, E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell* **172**, 1239-1259
3. Makarova, K. S., Anantharaman, V., Grishin, N. V., Koonin, E. V., and Aravind, L. (2014) CARF and WYL domains: ligand-binding regulators of prokaryotic defense systems. *Front Genet* **5**, 102
4. Lintner, N. G., Frankel, K. A., Tsutakawa, S. E., Alsbury, D. L., Copie, V., Young, M. J., Tainer, J. A., and Lawrence, C. M. (2011) The Structure of the CRISPR-Associated Protein Csa3 Provides Insight into the Regulation of the CRISPR/Cas System. *J Mol Biol* **405**, 939-955
5. Makarova, K. S., Anantharaman, V., Aravind, L., and Koonin, E. V. (2012) Live virus-free or die: coupling of antivirus immunity and programmed suicide or dormancy in prokaryotes. *Biol Direct* **7**, 40
6. Yan, X., Guo, W., and Yuan, Y. A. (2015) Crystal structures of CRISPR-associated Csx3 reveal a manganese-dependent deadenylation exoribonuclease. *RNA Biol* **12**, 749-760

7. Topuzlu, E., and Lawrence, C. M. (2016) Recognition of a pseudo-symmetric RNA tetranucleotide by Csx3, a new member of the CRISPR associated Rossmann fold superfamily. *RNA Biol* **13**, 254-257

8. Kazlauskiene, M., Kostiuk, G., Venclovas, C., Tamulaitis, G., and Siksny, V. (2017) A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. *Science* **357**, 605-609

9. Niewoehner, O., Garcia-Doval, C., Rostol, J. T., Berk, C., Schwede, F., Bigler, L., Hall, J., Marraffini, L. A., and Jinek, M. (2017) Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers. *Nature* **548**, 543-548

10. Athukoralage, J. S., Rouillon, C., Graham, S., Gruschow, S., and White, M. F. (2018) Ring nucleases deactivate type III CRISPR ribonucleases by degrading cyclic oligoadenylate. *Nature* **562**, 277-280

11. Haft, D., Selengut, J., Mongodin, E., and Nelson, K. (2005) A Guild of 45 CRISPR-Associated (Cas) Protien Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. *Plos Computational Biology* **1**, 474-483

12. Pourcel, C., Touchon, M., Villeri, O., Vernaudet, J. P., Couvin, D., Toffano-Nioche, C., and Vergnaud, G. (2020) CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and cas genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res* **48**, D535-D544

13. Shah, S. A., Alkhnbashi, O. S., Behler, J., Han, W., She, Q., Hess, W. R., Garrett, R. A., and Backofen, R. (2019) Comprehensive search for accessory proteins encoded with archaeal and bacterial type III CRISPR-cas gene cassettes reveals 39 new cas gene families. *RNA Biol* **16**, 530-542

14. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389-3402

15. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A. R. N., Potter, S. C., Finn, R. D., and Lopez, R. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* **47**, W636-W641

16. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**, 774-797

17. Bennett, M. J., Schlunegger, M. P., and Eisenberg, D. (1995) 3D domain swapping: a mechanism for oligomer assembly. *Protein Sci* **4**, 2455-2468

18. Athukoralage, J. S., Graham, S., Gruschow, S., Rouillon, C., and White, M. F. (2019) A Type III CRISPR Ancillary Ribonuclease Deprades Its Cyclic Oligoadenylate Activator. *J Mol Biol* **431**, 2894-2899

19. Garcia-Doval, C., Schwede, F., Berk, C., Rostol, J. T., Niewoehner, O., Tejero, O., Hall, J., Marraffini, L. A., and Jinek, M. (2020) Activation and self-inactivation mechanisms of the cyclic oligoadenylate-dependent CRISPR ribonuclease Csm6. *Nat Commun* **11**, 1596

20. Jia, N., Jones, R., Yang, G., Ouerfelli, O., and Patel, D. J. (2019) CRISPR-Cas III-A Csm6 CARF Domain Is a Ring Nuclease Triggering Stepwise cA4 Cleavage with ApA>p Formation Terminating RNase Activity. *Mol Cell* **75**, 944-956 e946
21. Athukoralage, J. S., McMahon, S. A., Zhang, C., Gruschow, S., Graham, S., Krupovic, M., Whitaker, R. J., Gloster, T. M., and White, M. F. (2020) An anti-CRISPR viral ring nuclease subverts type III CRISPR immunity. *Nature* **577**, 572-575

22. Rouillon, C., Athukoralage, J. S., Graham, S., Gruschow, S., and White, M. F. (2018) Control of cyclic oligoadenylate synthesis in a type III CRISPR system. *Elife* **7**

23. Athukoralage, J. S., Graham, S., Rouillon, C., Gruschow, S., Czekster, C. M., and White, M. F. (2020) The dynamic interplay of host and viral enzymes in type III CRISPR-mediated cyclic nucleotide signalling. *Elife* **9**

24. Okonechnikov, K., Golosova, O., Fursov, M., and team, U. (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166-1167

25. Zimmermann, L., Stephens, A., Nam, S. Z., Rau, D., Kubler, J., Lozajic, M., Gabler, F., Soding, J., Lupas, A. N., and Alva, V. (2018) A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. *J Mol Biol* **430**, 2237-2243

26. Menon, S. K., Eilers, B. J., Young, M. J., and Lawrence, C. M. (2010) The crystal structure of D212 from sulfobolus spindle-shaped virus ragged hills reveals a new member of the PD-(D/E)XK nuclease superfamily. *J Virol* **84**, 5890-5897

27. Studier, F. W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**, 207

28. Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem.* **72**, 248-254

29. Petersen, K. J., Peterson, K. C., Muretta, J. M., Higgins, S. E., Gillispie, G. D., and Thomas, D. D. (2014) Fluorescence lifetime plate reader: resolution and precision meet high-throughput. *Rev Sci Instrum* **85**, 113101

30. Muretta, J. M., Kyrychenko, A., Ladokhin, A. S., Kast, D. J., Gillispie, G. D., and Thomas, D. D. (2010) High-performance time-resolved fluorescence by direct waveform recording. *Rev Sci Instrum* **81**, 103101

31. Schlick, K. H., Lange, C. K., Gillispie, G. D., and Cloninger, M. J. (2009) Characterization of protein aggregation via intrinsic fluorescence lifetime. *J Am Chem Soc* **131**, 16608-16609

32. Bernhard, S. P., Goodman, C. K., Norton, E. G., Alme, D. G., Lawrence, C. M., and Cloninger, M. J. (2020) Time-dependent fluorescence spectroscopy to quantify complex binding interactions. *Submitted*

33. Otwinowski, Z., and Minor, W. (1997) [20] Processing of X-ray diffraction data collected in oscillation mode. in *Methods in Enzymology*, Academic Press. pp 307-326

34. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658-674

35. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, D. N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **D66**, 213-221

36. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **66**, 486-501
37. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D-Struct. Biol. 66, 213-221

38. Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. Sect. D-Struct. Biol. 66, 12-21

39. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. http://www.pymol.org.

40. Sternberg, S. H., Haurwitz, R. E., and Doudna, J. A. (2012) Mechanism of substrate selection by a highly specific CRISPR endoribonuclease. RNA 18, 661-672

Abbreviations and Nomenclature
CRISPR, clustered regularly interspaced short palindromic repeats; Cas, CRISPR associated genes/proteins, CARF, CRISPR associated Rossmann fold; cA4, cyclic tetra-adenylate; cOA, cyclic oligoadenylate; Crn1, CRISPR ring nuclease 1; HEPN, Higher Eukaryotes and Prokaryotes, Nucleotide binding; PAM, Protospacer Adjacent Motif
### Table 1 - Data Collection and Model Refinement

| Data Set        | AfCsx3       |
|-----------------|--------------|
| Wavelength (Å)  | 0.979        |
| Space Group     | I222         |
| Cell Constants (Å) | 53.45, 64.54, 70.11 |
| (α, β, γ; °)    | 90, 90, 90   |
| Resolution Range (Å) | 46.25 - 1.80 (1.83 -1.80) |
| Unique Reflections | 10,996 (510) |
| Average Redundancy | 4.8 (4.1) |
| I/σa            | 4.74 (2.87)  |
| Completeness (%) | 99.5 (96.4) |
| R_pim (%)       | 3.2          |
| R_mere (%)      | 3.4 (4.4)    |
| CC1/2           | 0.999 (0.917)|

#### Model

|                    |               |
|--------------------|---------------|
| R_work ab (%)      | 18.7 (24.6)   |
| R_free ab (%)      | 21.6 (27.0)   |
| Real Space CC (%)  | 87.10         |
| Mean B Value (overall; Å²) | 45.68 |
| Coordinate Error (based on maximum likelihood, Å) | 0.19 |
| RMSD from ideality: |               |
| Bonds (Å)          | 0.01          |
| Angles (°)         | 0.82          |
| Ramachandran Plot: |               |
| Most Favored (%)   | 98.04         |
| Additional Allowed (%) | 1.96 |
| PDB Accession Code | 6VJG          |

*Numbers in parenthesis refer to the highest resolution shell.

**R_work = Σ||Fo|−|Fc||/Σ|Fo| where Fo and Fc are the observed and calculated structure factor amplitudes used in refinement. R_free is calculated as R_work, but using the "test" set of structure factor amplitudes that were withheld from refinement (9.97%).

**Correlation coefficient (CC) is agreement between the model and 2mFo-DFc electron density map.
Figure 1. AfCsx3 binds cA4 with high affinity. A) The fluorescence lifetime of AfCsx3 Trp50 was used to follow cA4 binding. The mean waveform at each cA4 concentration was modeled as a linear combination of the cA4 bound and free waveforms, yielding the fraction bound. Each point is the mean of three independent experiments, each with three technical replicates. Error bars represent the standard deviation of the mean. The data were then fit to the one site single binding function in GraphPad Prism giving $K_d = 55 \pm 4.5$ nM. Importantly, the indicated cA4 concentrations along the abscissa represent total ligand concentration. Free ligand concentration is significantly less at the lower concentrations of cA4 used in this assay. Thus, the apparent $K_d$ represents only an upper limit, and may well be an underestimate of the affinity. B) Note that the ligand concentrations in panel B are $\mu$M rather than nM as in panel A. The fluorescence lifetime binding assay shows significantly lower affinity for the small linear RNA with a poly-A tail, giving a $K_d$ of $1.9 \pm 0.2$ $\mu$M.
Figure 2. The AfCsx3 Fold. Previous structures of Csx3 reported a 6-stranded mixed β-sheet with 2 α-helices. The high-resolution crystal structure described here revealed an ordered C-terminus with a third alpha helix (α3, dark blue) followed by an additional β-strand (β7, violet). This C-terminal extension crosses the outside, solvent exposed surface of the β-sheet, allowing β7 to add antiparallel to β1. The fold is thus a mixed seven-stranded β-sheet with β7↑-β1↓-β2↑-α1-β3↑-α2→β4↑-β5↓-β6↑-α3 topology with the α-helices providing right-handed crossovers between the indicated β-strands. In a dimer, α1 and α2 are buried at the subunit interface while α3 is exposed on the surface. Conserved motifs GRXPXW, LXHXXH and DPR are found in the β3-α2, α2-β4, and β4-β5 loops. Note the LXHXXH RNase active site motif that lies at the “bottom” of the protomer, while the GRXPXW and DPR motifs involved in cA4 recognition are at the top.
Figure 3. The dynamic C-terminal tail. A) The Csx3 dimer packs in sheets or planes within the crystal. B) The inter-dimer crystal contacts within a plane are anchored primarily by a parallel β-strand hydrogen bond interaction between β7 (violet) of one subunit and β6 (light blue) of the neighboring dimer. The ordered β7-strand of the C-terminal tail thus serves to cross link dimers within the crystal as it “inserts” between β1 (green) of one subunit and β6 of the neighbor. This interaction leads to the formation of continuous extended β-sheets that running the length of the crystal. Visual observations and PISA interface analysis software indicate these interactions are not substantial enough to support the formation of a stable protein interface in solution. This open symmetry is thus a likely artifact of crystallization conditions and is not indicative of interactions in vivo. That said, the dynamic C-terminal tail clearly has the ability to form inter-protein interactions which might mediate protein-protein interactions within the cell.
Figure 4. Af Csx3 ring nuclease activity. A) TLC analysis showing a representative time course for the degradation of unlabeled cA₄ (150 μM) incubated with Af Csx3 (10 μM dimer) at 50 °C in the presence of 200 μM Mn²⁺. Lanes 1-9 are time points at 0, 0.5, 1, 3, 5, 10, 20, 30, and 60 minutes. Lane 10 was a 60 minute incubation in the presence of Mn²⁺ in the absence of Csx3. Lane 11 was loaded with 1.5 nanomoles of cA₄ standard. B) Data from five independent cA₄ ring nuclease activity replicates were averaged, and the data fit to a first order exponential decay model yielding an apparent first order rate constant, \( k_{obs} = 4.95 \pm 0.43 \text{ min}^{-1} \).
Figure 5. Structural definition of the CARF domain core. Superposition of the $\beta_3$-$\alpha_2$-$\beta_4$-$\beta_5$ sequence of Csx3 (A) onto the $\beta_4$-$\alpha_4$-$\beta_5$-$\beta_6$ sequence of *S. solfataricus* Csa3 CARF domain (B) (RMSD 2.9 Å) reveals core secondary structures shared by the two proteins (green). Csa3 and Csx3 each contain this contiguous $\beta\alpha\beta\beta$ super secondary structures in which the first two $\beta$-strands are connected by a right-handed helical crossover followed by a reverse turn into the 3rd $\beta$-strand. The core helix in this motif is an essential part of the dimer interface, it packs against the two beta-strands connected by the reverse turn, $\beta_4\beta_5$ in Csx3 (left) and $\beta_5\beta_6$ in Csa3 (right). In the classic dinucleotide binding domain, these beta-strands are instead connected by a right-handed helical crossover, in which the connecting helix occupies the same space as the helix from the adjacent Csa3 or Csx3 subunit. The reverse turn connecting these $\beta$-strands is thus a critical structural feature of the CARF domain dimer interface. In addition, Csx3 places $\beta_2$ and $\alpha_1$ in positions equivalent to the discontinuous $\beta_1$ and $\alpha_3$ secondary structural elements of Csa3 (grey).
Figure 6. RNA density in the 3WZI Csx3 structure accommodates cA4. A) Stereo image. An omit map was generated using the 3WZI model with RNA atoms deleted and contoured at 1.5 $\sigma$ (green mesh). The omit map was then used to inform modeling the cA4/Csx3 interaction. Csx3, oriented as in panel 5A, is represented as a surface model with nitrogen atoms in blue, oxygen atoms in red, carbon atoms colored grey, and cA4 is represented in sticks with carbon atoms colored yellow. B) Csx3 secondary structure elements shown in cartoon representation with select side chains shown as sticks. cA4 is also shown in stick representation with carbons in beige.
Csx3 is a cyclic oligonucleotide phosphodiesterase associated with type-III CRISPR-Cas that degrades the second messenger cA4
Sharidan Brown, Colin C. Gauvin, Alexander A. Charbonneau, Nathaniel Burman and C. Martin Lawrence

J. Biol. Chem. published online August 21, 2020

Access the most updated version of this article at doi: 10.1074/jbc.RA120.014099

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts