PLOS BIOLOGY

RESEARCH ARTICLE

Widespread prevalence of a methylation-dependent switch to activate an essential DNA damage response in bacteria

Aditya Kamat[1], Ngat T. Tran[2], Mohak Sharda[1], Neha Sontakke[1], Tung B. K. Le[2]*, Anjana Badrinarayanan[1]*

1 National Centre for Biological Sciences (TIFR), Bengaluru, India, **2** John Innes Centre, Department of Molecular Microbiology, Colney Lane, Norwich, United Kingdom

* tung.le@jic.ac.uk (TL); anjana@ncbs.res.in (AB)

Citation: Kamat A, Tran NT, Sharda M, Sontakke N, Le TBK, Badrinarayanan A (2024) Widespread prevalence of a methylation-dependent switch to activate an essential DNA damage response in bacteria. PLoS Biol 22(3): e3002540. https://doi.org/10.1371/journal.pbio.3002540

Academic Editor: Lotte Søgaard-Andersen, Max Planck Institute for Terrestrial Microbiology: Max-Planck-Institut fur terrestrische Mikrobiologie, GERMANY

Received: October 20, 2023

Accepted: February 6, 2024

Published: March 11, 2024

Data Availability Statement: All relevant data are within the paper and its Supporting information files, except for RNA-seq, ChIP-seq and mass spectrometry data. ChIP-seq and RNA-seq data are deposited in GEO repository, with accession number GSE246227 and GSE246782 respectively. Raw data associated with mass spectrometry experiments are available via ProteomeXchange with identifier PXD049136.

Abstract

DNA methylation plays central roles in diverse cellular processes, ranging from error-correction during replication to regulation of bacterial defense mechanisms. Nevertheless, certain aberrant methylation modifications can have lethal consequences. The mechanisms by which bacteria detect and respond to such damage remain incompletely understood. Here, we discover a highly conserved but previously uncharacterized transcription factor (Cada2), which orchestrates a methylation-dependent adaptive response in *Caulobacter*. This response operates independently of the SOS response, governs the expression of genes crucial for direct repair, and is essential for surviving methylation-induced damage. Our molecular investigation of Cada2 reveals a cysteine methylation-dependent posttranslational modification (PTM) and mode of action distinct from its *Escherichia coli* counterpart, a trait conserved across all bacteria harboring a Cada2-like homolog instead. Extending across the bacterial kingdom, our findings support the notion of divergence and coevolution of adaptive response transcription factors and their corresponding sequence-specific DNA motifs. Despite this diversity, the ubiquitous prevalence of adaptive response regulators underscores the significance of a transcriptional switch, mediated by methylation PTM, in driving a specific and essential bacterial DNA damage response.

Introduction

Information regarding biological processes is primarily encoded in the genomes of living organisms. Functional modalities are further enhanced via modulation of the epigenetic states of DNA [1–4]. For example, methylation of DNA occurs in organisms across all domains of life. In bacteria, physiological DNA methylations at N5-meC, N4-meC, and N6-meA nucleotide positions play diverse functions ranging from defensive roles involving immunity against invasive genetic elements to regulatory roles in cell cycle and transcriptional control [2,5]. In contrast to these physiological modifications, certain types of methylations can be aberrant. DNA methylations on O6-meG and O4-meT positions act as a potent source of mutagenesis, while N3-meA, N1-meA, and N3-meC block DNA replication and transcription [6–9].

Funding: This work was supported by funding to AK (TIFR graduate student fellowship) and to A.B. via the India Alliance Intermediate Grant (Grant number IA/I/21/1/505630) and intramural funding via NCBS-TIFR (Grant number 03/3/2019/R&D-II/DAE/4749). This work was also supported by the Lister Institute fellowship, the Wellcome Trust Investigator grant 221776/Z/2/Z (to T.B.K.L), and by BBSRC funded Institute Strategic Program Harnessing Biosynthesis for Sustainable Food and Health (HBio) (BB/X01097X/1 that supports N.T.T). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: DDR, DNA damage response; HMM, hidden Markov model; MMS, methyl methane sulphonate; MSA, multiple sequence alignment; PTM, posttranslational modification; PYE, peptone yeast extract.

Bacteria cope with this threat of aberrant methylations via eliciting 2 DNA damage responses (DDRs). The first response is the well-characterized and ubiquitously conserved SOS response that consists of 2 regulatory units, RecA and LexA [10]. The transcriptional repressor LexA occupies *lexA* boxes present in the promoter regions of SOS genes and inhibits their expression. Upon DNA damage, an activated RecA nucleoprotein filament triggers the autocleavage of LexA leading to de-repression of the SOS response [11–13]. Three distinct features are hallmarks of this response: (a) It is induced under all forms of DNA damage and is not specific to methylation damage alone [14]; (b) the response has leaky expression even in the absence of damage [15,16]; and (c) the response can result in mutagenesis due to expression of translesion synthesis polymerases [17]. Thus, SOS response induction is a trade-off between its essentiality and its cost.

The SOS response is the primary response of bacteria to DNA damage. However, recent studies indicate that the SOS response is often complemented by SOS-independent DNA damage response pathways [18]. The PafBC response was discovered in *Mycobacterium tuberculosis* and is essential for inducing a majority of the genes under mitomycin C exposure [19,20]. Induction of this response relies on the PafBC heterodimer which is suggested to exhibit an ssDNA-dependent activation similar to the SOS response [21]. The DriD response of *Caulobacter* also exhibits a similar mode of activation [22]. The inhibition of cell division by *didA* (belonging to the DriD regulon) under DNA damage is well characterized [23]. However, recent studies also suggest that DriD may play a role beyond DDR pathway activation [24]. Both DriD and PafBC belong to the WYL-dependent family of transcription factors that is widely conserved across bacteria [25]. These pathways seem to be induced prominently by DNA double strand breaks but their role in the repair of methylation DNA damage remains to be uncovered [23,26].

A temporally delayed DDR subsequent to the SOS response, known as the adaptive response to methylation damage (or the Ada response) has also been reported in *Escherichia coli* [27–29]. This response exhibits key distinguishing features compared to the SOS response: (a) It is activated independent of the SOS response and is specific to methylation damage only [30]; (b) the response is adaptive; i.e., exposure to a sublethal dose of methylation damage encodes for memory that allows cells to adapt and survive formerly lethal levels of DNA methylation damage [27,30]; (c) the response exhibits bi-stability [31], a subpopulation of cells do not induce the response even under continuous exposure to methylation damage; and (d) repair via this response is non-mutagenic [27,32].

The master regulator of this DDR is the methyltransferase EcAda, comprising of an AdaA domain in its N-terminus (which associates with "A" and "B" DNA boxes located upstream of EcAda-regulated promoters) and a C-terminal methyltransferase domain [28]. EcAda expression is tightly regulated, with only 0–2 molecules of Ada present in cells in the absence of damage [31]. Posttranslational modification (PTM) via methylation of a conserved cysteine in N-Ada is crucial for activating EcAda as a transcription factor [33–35]. Once activated, EcAda drives its own expression as well as expression of genes involved in direct repair of methylation lesions.

How prevalent is the risk of methylation damage, which could have resulted in the evolution of a methylation-specific DDR? Instances of encountering methylation DNA damage are not infrequent, and bacteria often face methylation stress. Intrinsic factors, such as insidious metabolic byproducts (e.g., lipid peroxidation) pose a prominent risk for DNA methylation damage [9,28]. Environmental factors including halocarbons and methylating agents produced by bacteria as instruments of inter-microbial warfare also contribute to this damage [36,37]. Additionally, infected mammalian cells subject invading bacteria to methylation stress [38]. More recently, studies have reported instances where physiological DNA

methyltransferases, which are otherwise innocuous, erroneously inflict methylation damage on genomic DNA [39]. Thus, given the prevalence of methylation stress, it is likely that many bacteria have evolved unique systems such as the Ada response to protect their genomes.

Intriguingly, despite the simultaneous discovery of the adaptive response to methylation damage alongside the SOS response [27,40], our knowledge of this pathway is presently restricted only to the *E. coli* paradigm. Thus, the significance and conservation of a methylation-based PTM in the regulation of a damage-specific bacterial DDR remains unknown. Indeed, a limited number of computational studies suggest only sporadic conservation of EcAda across the bacterial kingdom [41–43], and many bacteria, such as *Caulobacter crescentus*, are reported to lack an inducible adaptive response altogether [44,45].

In this work, we discover a methylation-specific DDR in *Caulobacter crescentus*, which showcases key features of an adaptive response. This pathway is regulated by a conserved but uncharacterized transcription factor *ccna_03845* ("Cada2" henceforth). Detailed molecular characterization reveals a novel sequence-specific DNA-binding domain in this protein, as well as a methylation-based PTM required for activation of Cada2 as a transcription factor in a manner that is distinct from its *E. coli* counterpart. Despite the contrasting mechanistic features of EcAda and Cada2, we observe remarkable similarities in the defining characteristics of the downstream response. Phylogenetic distribution of adaptive response regulatory proteins further reveals their widespread prevalence across the bacterial kingdom, with ubiquitous conservation of key residues required for methylation-dependent activation (either in an EcAda-like or Cada2-like form). Collectively, our work highlights the importance of a transcriptional switch mediated by methylation PTM in activating an essential bacterial response to methylation DNA damage.

Results

A methylation-specific DNA damage response in *Caulobacter*

We undertook a comprehensive transcriptomic approach to identify whether *Caulobacter* induces an SOS-independent transcriptional response specific to methylation damage. For this, we treated *Caulobacter* cells with agents that predominantly induce 1 specific form of DNA damage (methyl methane sulphonate (MMS), a causative agent of methylation DNA damage [46], mitomycin-C (MMC) that induces intra-strand crosslinks and mono-adducts [47], and norfloxacin that induces double-strand breaks [48]). We collected samples at 0, 20, and 40 min post damage exposure for RNA sequencing (Figs 1A and S1A). Our analysis revealed several genes that were induced in all conditions ("universal") and a set of genes that were specifically induced under MMS treatment, but in none of the other damaging conditions (methylation-specific) (S1A Fig).

We next asked whether the expression of these genes was independent of the SOS response. We subjected *recA* deleted cells to MMS treatment and carried out RNA sequencing in this background. Comparing wild type to *recA* deletion cells showed that the methylation damage-specific genes were induced in a *recA*-independent manner (SOS-independent) (Fig 1A). In contrast, genes that were induced under all DNA damaging conditions in wild type background remain uninduced in *recA* deletion cells (SOS-dependent) (Figs 1A and S1A).

We identified *ccna_00745* to be the highest induced under methylation damage in the transcriptome analysis (Fig 1A). This gene is co-operonic with a second gene *ccna_00746*, that is also induced in a methylation DNA damage-specific manner (S1A Fig). *ccna_00745* is Bioinformatically predicted to be a 2-oxoglutarate, Fe^{2+}-dependent dioxygenase, similar to *E. coli alkB*, while *ccna_00746* is predicted to possess an Ada-like methyltransferase domain (PF01035) [9].

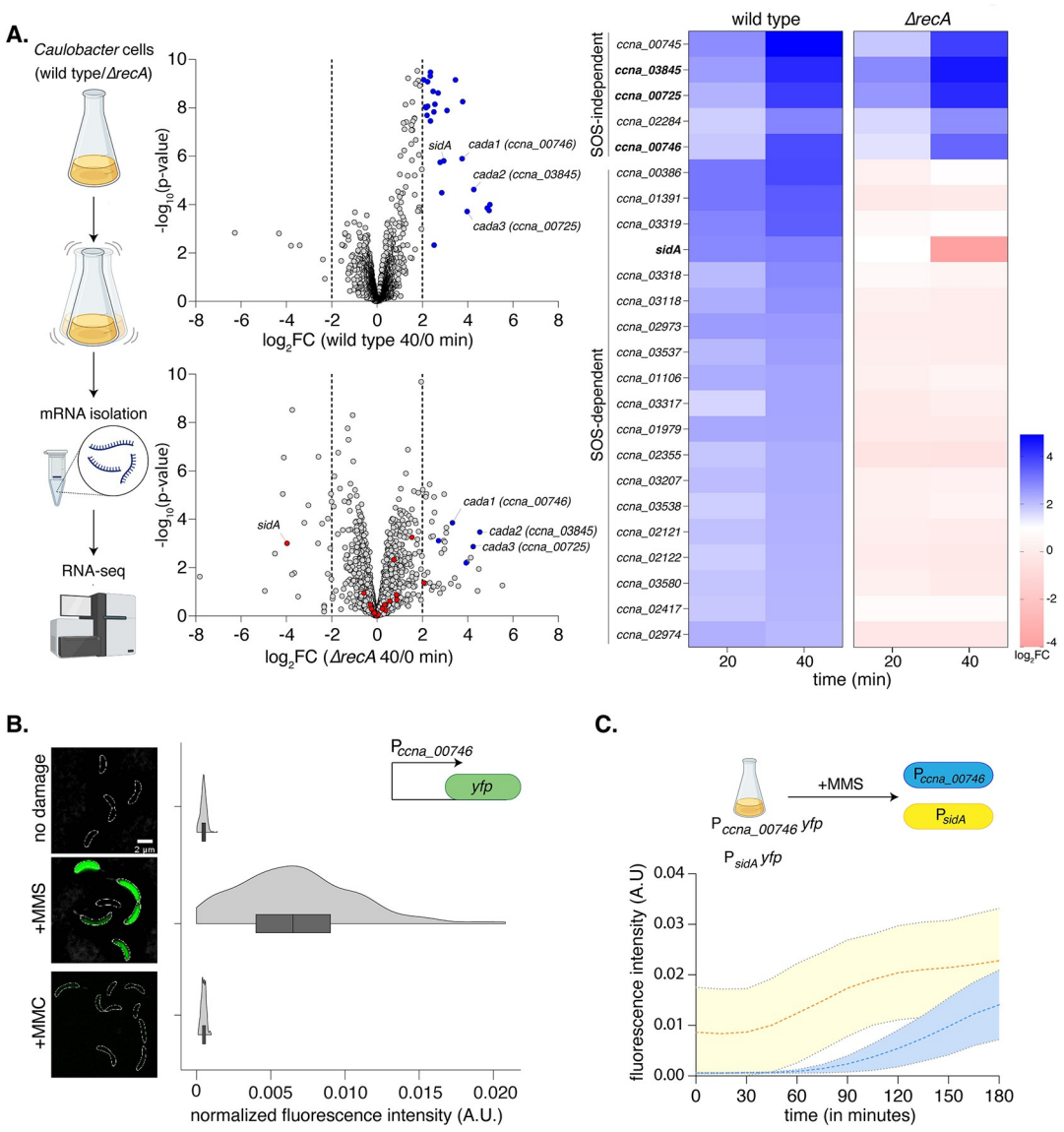

Fig 1. A methylation-specific DNA damage response in *Caulobacter*. (A) [Left] Schematic summarizing the RNA-sequencing experiment is provided. Wild-type and *ΔrecA* (SOS-deficient) cells were exposed to 1.5 mM MMS for 40 min. [Centre] Volcano plots representing differentially expressed genes in comparison to cells exposed to no damage for wild-type (top) and *ΔrecA* (bottom) cells, respectively. Genes up-regulated in wild-type cells ($log_2FC > 2$ and $–log_{10}(p\text{-value}) > 2$) are highlighted in blue, while genes highlighted in red represent genes up-regulated in wild-type but down-regulated in *ΔrecA* cells. [Right] Heat maps represent log_2FC values for individual genes induced in wild-type and *ΔrecA* cells under MMS exposure. **(B)** Wild-type P_{ccna_00746}-*yfp* reporter (schematic inset) was exposed to 1.5 mM MMS or 0.5 μg/ml MMC, respectively for 2 h. Representative cells are shown on the left (scale bar: 2 μm, cell boundaries are marked by white dotted outline, here and in all other images). Violin plots show fluorescence intensity distribution normalized to cell area from single cells ($n = 300$, from 3 biological replicates). The underlying data are available in S1 Data. **(C)** [Top] Schematic of the experimental protocol comparing induction kinetics of the *Caulobacter* SOS and methylation-specific response via time lapse microscopy. [Bottom] Fluorescence intensity normalized to cell area for P_{ccna_00746}-*yfp* and P_{sidA}-*yfp* cells over 3 h of exposure to 1.5 mM MMS. The underlying data are available in S1 Data.

https://doi.org/10.1371/journal.pbio.3002540.g001

To further corroborate the RNA-seq observations, we constructed a fluorescence-based readout for promoter activity of these genes (P_{ccna_00746}-*yfp*). No significant expression of *yfp* was detected in the absence of damage (Fig 1B). We observed YFP expression from this construct only in cells treated with MMS, and not other DNA damaging agents MMC, norfloxacin

and hydroxyurea (HU) (Figs 1B and S1B). We additionally utilized a second methylation damaging agent, streptozotocin (STZ), a naturally occurring antibiotic produced by *Streptomyces achromogenes* var. *streptozoticus* [36,37]. In this case as well, we observed YFP expression from the methylation damage-specific promoter (S1B Fig). These observations were distinct from those seen for a fluorescence reporter for the promoter of *sidA*, a known SOS-dependent gene [49], that has been reported to be induced under other types of DNA damage in a RecA-dependent manner [49,50]. Additionally, the P_{ccna_00746}-*yfp* reporter exhibited methylation damage-dependent induction in cells lacking the *driD* transcription factor as well (S1C Fig). Thus, *Caulobacter* elicits a transcriptional response to methylation damage that is independent of the SOS as well as the DriD-dependent DNA damage responses.

We next investigated the temporal kinetics of the methylation-specific response in relation to the SOS response. For this, we carried out time-lapse imaging of cells carrying either the P_{ccna_00746}-*yfp* or the P_{sidA}-*yfp* reporter after MMS exposure. We observed that expression of *yfp* from the methylation damage-specific promoter expression was temporally delayed when compared to the SOS reporter (Fig 1C). Finally, we tested whether the induction kinetics of this response is adaptive. For this, we first exposed our promoter fusion strain to a low (0.5 mM) dose of MMS. This led to only a modest increase in the expression of P_{ccna_00746}-*yfp* (S1C Fig). Consistent with an "adaptive" response [51], these pre-treated cells were able to induce P_{ccna_00746}-*yfp* significantly faster as compared to untreated cells upon exposure to a higher dose (1.5 mM) of MMS (S1C Fig). Together, the *Caulobacter* methylation damage-specific response showcases key features of an adaptive response to methylation damage as reported in *E. coli*. Based on these observations, we henceforth refer to this response as the *Caulobacter* adaptive response to methylation damage.

Caulobacter adaptive response to methylation damage is regulated by Cada2

Given the striking similarity between the adaptive responses of *Caulobacter* and *E. coli*, we wondered whether the *Caulobacter* response was regulated by an EcAda-like protein [52]. While we could not identify any EcAda-like protein in *Caulobacter*, we observed that 3 adaptive response candidates (*ccna_00746*, *ccna_03845*, and *ccna_00725*) possessed an Ada-like methyltransferase domain (PF01035) in their C-terminus region (Fig 2A). Intriguingly, domains corresponding to N-Ada of EcAda protein (required for forming sequence-specific interactions with the cognate Ada promoters) were split between *ccna_00746* (with A box-binding domain) and *ccna_03845* (with B box-binding domain) (Fig 2A). We thus annotated these as "*cada*" (***C****aulobacter* ***ada****ptive response*) genes *cada1*, *cada2*, and *cada3*.

Next, we asked whether Cada1, Cada2, or Cada3 regulated the adaptive response. We deleted all 3 *cada* genes individually and assessed the activity of the P_{ccna_00746}-*yfp* reporter. We found that only cells lacking *cada2* were unable to induce P_{ccna_00746}-*yfp* under methylation damage (Fig 2B). In support, RNA-sequencing analysis revealed that other MMS-specific, RecA-independent genes were also not induced in cells lacking *cada2*, but were unaffected in cells lacking *cada1* (Fig 2C). Thus, A-box containing Cada1 does not drive gene expression under the adaptive response, while Cada2 possessing solely the B-box binding domain is required for response activation. We next tested cell survival upon exposure to MMS as well as STZ. We found that survival was significantly compromised specifically under STZ damage in cells lacking *cada2* (Figs 2D and S1E), suggesting that the Cada2-dependent adaptive response to methylation damage was an essential DDR pathway in *Caulobacter*. The difference in survival between the 2 methylating agents can be attributed to distinct patterns of methylation modifications induced by STZ and MMS [53]. Thus, it is possible that the extent of essentiality

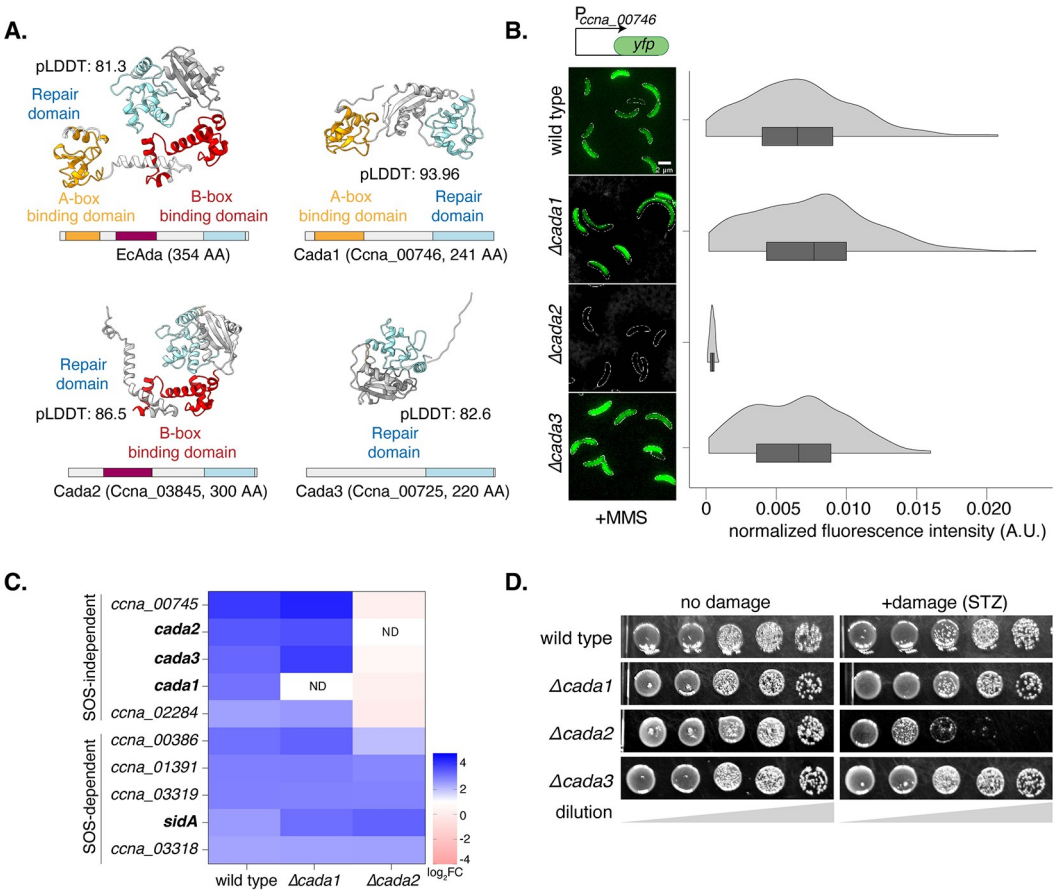

Fig 2. *Caulobacter* adaptive response to methylation damage is regulated by Cada2. (A) Comparative analysis of high-confidence Alphafold structural models of EcAda, Cada1, Cada2, and Cada3 proteins along with their domain organizations. **(B)** [Left] Representative cells showing P_{ccna_00746}-*yfp* reporter induction in wild type (from Fig 1B) and individual deletion strains of the *cada* genes under 1.5 mM MMS damage. [Right] Violin plots showing fluorescence intensity distribution normalized to cell area from single cells (n = 300, from 3 biological replicates). The underlying data are available in S1 Data. **(C)** Heat map of $\log_2$FC values from RNA-seq experiments for genes up-regulated in wild type, *Δcada1* and *Δcada2* cells following 1.5 mM MMS treatment. **(D)** Survival assay of individual deletions of *cada* genes with and without methylating agent (STZ) (5 μg/ml).

https://doi.org/10.1371/journal.pbio.3002540.g002

is nuanced and dependent on the proportions of various base modifications that a given methylating agent induces.

Cada2 associates with adaptive response promoters in a sequence-specific manner

To determine whether Cada2 was directly responsible for the induction of the adaptive response genes under methylation damage, we performed ChIP-sequencing experiments (Fig 3A). For this, cells expressing *cada2-3x-flag* from its endogenous promoter were used. The flag-tagged strain resembled survival of wild-type cells under methylation DNA damage and damage-dependent expression of Cada2 was detected via western blot (S2A and S2B Fig). Using this strain, we carried out ChIP-seq experiments in the presence or absence of methylation damage (Fig 3A). We estimated Cada2-3xFlag enrichment across the *Caulobacter* genome under damage using untagged wild-type cells as control. This allowed us to identify bona fide Cada2-binding sites.

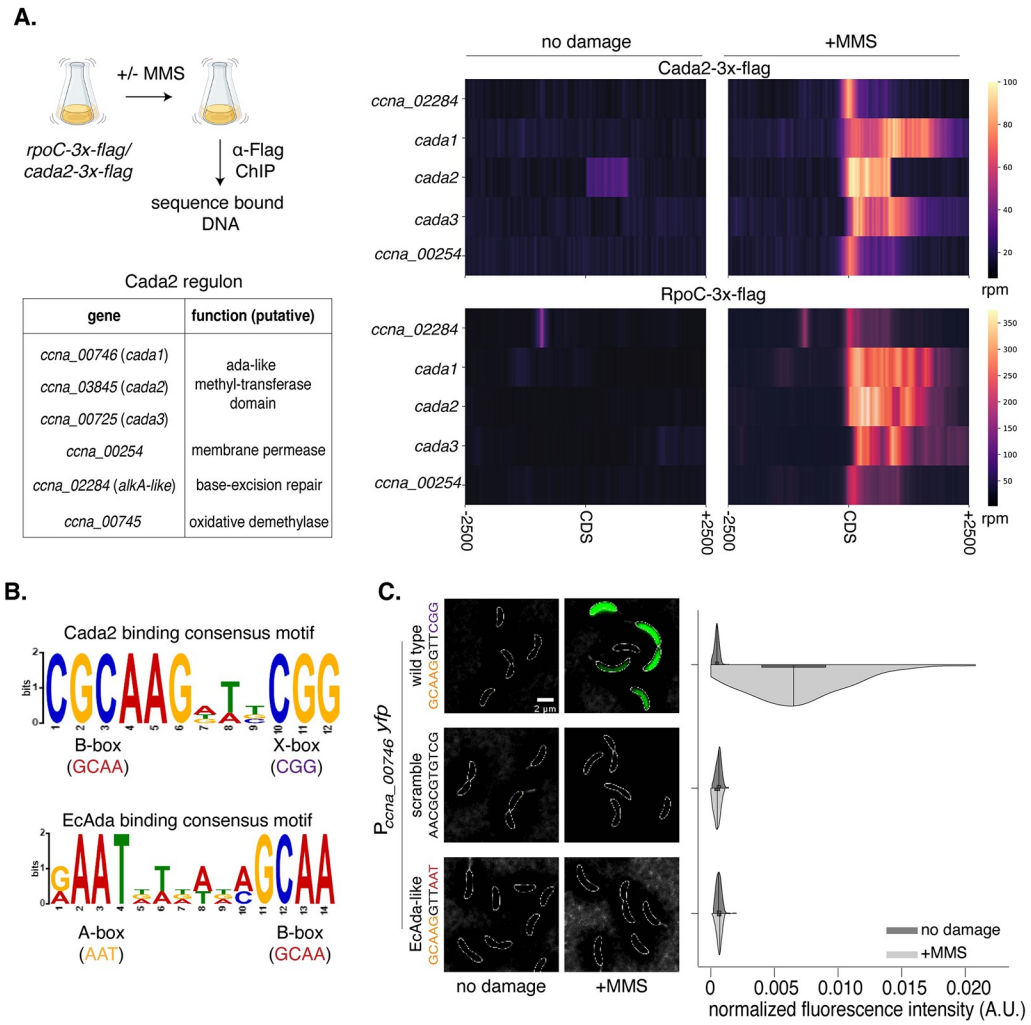

Fig 3. Cada2 associates with adaptive response promoters in a sequence-specific manner. (A) [Top left] Schematic of the ChIP-sequencing protocol. [Bottom left] Table showcasing the Cada2 regulon. Gene names and predicted functions are listed. [Top right] Normalized reads (in rpm) for *Cada2-3x-Flag* ChIP-seq represented ±2.5 kb around the CDS of the Cada2 regulon genes in the presence/absence of 1.5 mM MMS damage. [Bottom right] Data represented in a similar fashion for *RpoC-3x-Flag* ChIP-seq. **(B)** Binding consensus motif for Cada2 and EcAda derived from the promoters of the respective constituents of their regulon. **(C)** Representative cells showing induction of variants of the P_{ccna_00746}-*yfp* reporter (scramble/*EcAda*-like) compared to wild type under 1.5 mM MMS damage. [Bottom right] Half-violin plots show fluorescence intensity distribution normalized to cell area from single cells for the reporter variants in the presence (dark) and absence of damage (light) (n = 300, from 3 biological replicates). Wild-type data are represented again from Fig 1B. The underlying data are available in S1 Data.

https://doi.org/10.1371/journal.pbio.3002540.g003

Under damage, we observed Cada2 enrichment at its own promoter as well as at other promoters of the adaptive response (Figs 3A, S2C, and S2D). This autoregulation of its own expression appears to be a conserved feature between EcAda [54] and Cada2. To further assess the methylation damage-dependent transcription of these genes, we performed ChIP-seq with flag-tagged RNA polymerase subunit C (RpoC-3xflag) construct. We did not observe localization of RpoC to the *cada2*-associated promoters in the absence of methylation damage (Figs 3A, S2C, and S2D). Upon exposure to damage, RpoC signal could be detected at the *cada2* promoter, as well as other adaptive response promoters (Figs 3A, S2C, and S2D). ChIP-seq results overlapped significantly with the Cada2-dependent up-regulated genes identified using RNA-

seq experiments (approximately 83%). Superimposing the 2 datasets allowed us to describe the Cada2 regulon comprising 6 genes (Fig 3A), most of which appear to have direct repair-associated activities.

We carried out MEME analysis of Cada2-bound regions to identify any sequence-specific DNA-binding motifs. This revealed 2 sequences consistent across all promoters of the *Caulobacter* adaptive response (Fig 3B). One of these sequences, "GCAA," was identical to the B-box sequence motif bound by the B-box binding domain of EcAda (Fig 3B). Thus, we annotate it as the B-box motif in case of *Caulobacter* as well. We did not detect an A-box sequence motif ("AAT," bound by EcAda A-box binding domain). We instead identified a second recurrent motif that was GC-rich ("CGG"). We annotated this as "X-box" sequence motif. These 2 sequence motifs were separated by a 3-base pair spacer that exhibited no recurrently conserved sequence, but seemed to be consistently AT-rich. Together, we annotated the B-box and X-box motif, separated by the AT-rich spacer as the Cada2 binding sequence motif (Fig 3B).

To assess whether this sequence is required for Cada2-mediated regulation, we scrambled the complete sequence in our P_{ccna_00746}-*yfp* reporter construct. We found that such a reporter was no longer induced under methylation damage (Fig 3C). We next asked whether an EcAda-like DNA-binding motif comprising of an "A-box" instead of the "X-box" DNA motif could drive promoter activity by Cada2. Here too we observed that our reporter construct carrying an A and B box DNA motif was not responsive to methylation damage (Fig 3C). This suggests that the newly identified B+X-box motif is essential for the induction of the *Caulobacter* adaptive response.

Cada2-like proteins are widespread and encode a novel DNA-binding domain

The absence of the A-box DNA motif as well as the A-box protein domain in case of Cada2 led us to hypothesize that the X-box DNA motif must likely have a cognate DNA binding region in the Cada2 protein. We thus performed a comprehensive computational search for all Cada2-like methyltransferases across the bacterial kingdom. For this, we built a hidden Markov model (HMM) [55] profile from a multiple-sequence alignment of Cada2-like proteins. As comparison, we followed the same process to identify EcAda-like [28,35] and AdaA-like [42] proteins as well. Phylogenomic distribution of these proteins across a curated, nonredundant database of diverse bacterial species showcased that they are abundant and widespread across all major bacterial phyla (Figs 4A, S3D, and S3E). Furthermore, Cada2-like proteins infrequently co-occurred with EcAda-like proteins in the same genome (approximately 7%, S3F Fig).

We zoomed into the DNA-binding (regulatory) regions of EcAda and Cada2-like proteins to identify conserved and unique features. As anticipated, both classes of proteins encoded the B-box binding domain (marked by the conserved "SPHFQR" amino acid residues (Fig 4B)). It is likely that this region associates with the B-box DNA motif that can be found in both EcAda and Cada2 regulons. The 2 proteins deviated with regards to the A-box, with A-box (marked by the 4 cysteine residues) being conserved in all EcAda-like proteins (Fig 4B), but absent in all Cada2-like proteins. Instead, proximal to the putative B-box binding domain, Cada2-like proteins possessed a unique and highly conserved "RLHD" sequence domain (Fig 4B). We highlight here that the conservation of the RLHD domain is even more than that of the B-box. AlphaFold [56,57] model of the Cada2 protein predicted that this RLHD domain falls in a helix-turn-helix domain similar to the B-box binding domain (S3A Fig).

In support of the significance of these putative DNA-binding regions of Cada2, mutation of the conserved arginine (usually associated with DNA binding) in both domains to an alanine

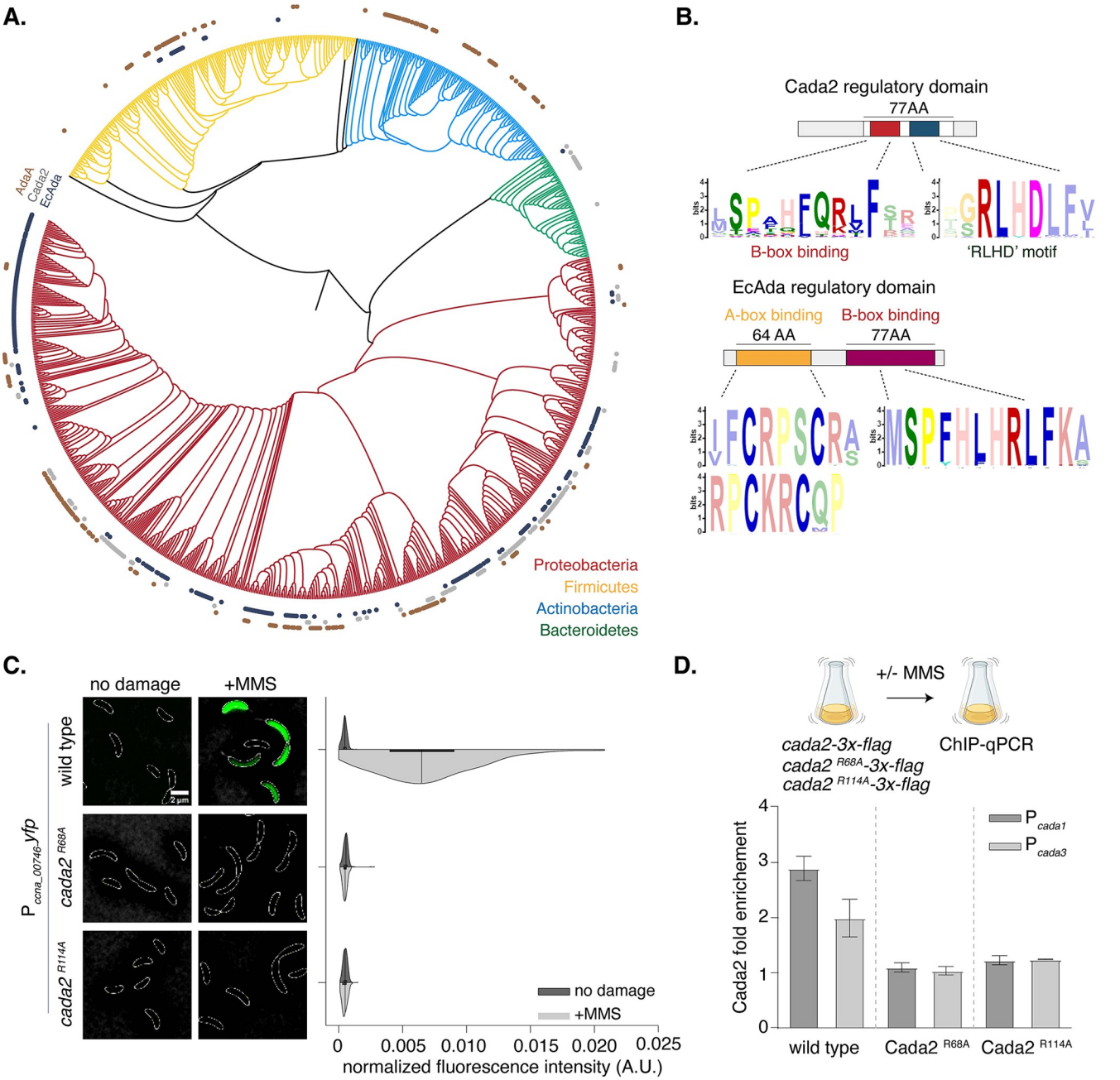

Fig 4. Cada2-like proteins are widespread and encode a novel DNA-binding domain. (A) Phylogenetic distribution at the genus level of EcAda-like (black), Cada2-like (gray), and AdaA-like (brown) proteins across the genomes of 4 bacterial phyla: proteobacteria (red), firmicutes (yellow), actinomycetes (blue), and bacteriodetes (green). Presence/absence is shown on a 16S rRNA-based phylogenetic tree of bacterial genomes. **(B)** [Top] MEME motif derived from the regulatory domain of 1,000 EcAda-like proteins. [Bottom] MEME motif derived from 1,000 Cada2-like proteins reveal conserved residues in the regulatory domain. **(C)** [Left] Representative images showing P_{ccna_00746}-*yfp* reporter induction in wild type and *cada2* mutant (*cada2*^{R68A} and *cada2*^{R114A}) background under 1.5 mM MMS damage. [Right] Half-violin plots show fluorescence intensity distribution normalized to cell area from single cells in the presence (dark) and absence of damage (light) (n = 300, from 3 biological replicates). Wild-type data are represented again from Fig 1B. The underlying data are available in S1 Data. **(D)** ChIP-qPCR comparing wild-type Cada2, Cada2^{R68A}, and Cada2^{R114A} enrichment at *cada1* and *cada3* promoters following exposure to 1.5 mM MMS. In all cases, flag-tagged version of protein is expressed from a xylose-inducible promoter on a high-copy replicating vector (in a *Δcada2* background). Bar graphs represent mean fold enrichment in comparison to wild-type *cada2* (n = 3 independent repeats, error bars represent standard error). The underlying data are available in S1 Data.

https://doi.org/10.1371/journal.pbio.3002540.g004

residue (B-box *cada2*^{R68A} and RLHD-motif *cada2*^{R114A}) abrogated P$_{ccna_00746}$-*yfp* reporter activity under methylation damage (Fig 4C). The 2 mutants were not compromised in expression and phenocopied the *cada2* deletion when assessed for survival under methylation damage (S3B and S3C Fig). Promoter binding was also compromised in both mutants as seen from ChIP-qPCR experiments for Cada2 association at promoter regions of genes belonging to the Cada2 regulon (Fig 4D). We conclude that the B-box and the novel and conserved "RLHD" domains on the Cada2 protein enable it to associate with its promoters in a sequence-specific manner.

Cada2 is a methylation-responsive transcription factor

We noticed in the ChIP profiles that Cada2 appeared to be modestly enriched at its own promoter even in no damage conditions (Figs 3A, S2C, and S2D). This was in contrast to RNA polymerase that was observed to bind to the *cada2* promoter region only in the present of methylation damage (Figs 3A, S2C, and S2D). Indeed, *Caulobacter* induced the adaptive response only upon exposure to DNA methylation damage (Fig 1B), and overexpression of *cada2* from a xylose-inducible promoter was insufficient to induce expression from the P$_{ccna_00746}$-*yfp* reporter in the absence of damage (Fig 5A).

We asked whether RNA polymerase recruitment to the promoter regions by Cada2 was mediated by physical interaction between Cada2 and the RNA polymerase holoenzyme, and if this step could be methylation dependent. We thus carried out bacterial-two-hybrid interaction analysis of Cada2 against various RNA polymerase holoenzyme subunits (RpoA, RpoB, RpoC, RpoD, and RpoZ). As a positive control, we used the helicase-nuclease protein complex components AddA and AddB [58]. In the presence of methylation damage, we observed interaction signal between Cada2 and RNA polymerase subunit A (S4A Fig).

How does methylation damage activate Cada2 function? Under damage, EcAda is posttranslationally methylated at conserved cysteine residues in its A-box binding and methyltransferase ("PCHR") domains, respectively [59,60]. However, it is the methylation of Cys$_{38}$ in the A-box binding domain that is required for its activation as a transcription factor via modulation of sequence-specific DNA binding affinity of EcAda [35]. While Cada2 lacks the A-box binding domain, it does possess the methyltransferase domain (in its C-terminus) (Fig 5B). Hence, we tested whether Cada2 is methylated post exposure to methylation DNA damage.

Using mass spectrometry, we identified a methylation modification on a cysteine residue (Cys$_{267}$, part of the "P**C**HR" methyltransferase domain) of Cada2 in cells treated with methylation damage, with no detectable methylation in the absence of damage (Fig 5B). We mutated the Cys$_{267}$ residue to an alanine to disrupt the methylation modification. This mutant, *cada2*^{C267A}, was unable to drive the induction of the P$_{ccna_00746}$-*yfp* reporter and phenocopied a *cada2* deletion under streptozotocin treatment (S4B and S4C Fig). Significantly, mutations in the promoter-binding domains of Cada2 (B-box *cada2*^{R68A} and RLHD-motif *cada2*^{R114A}) did not affect the ability of Cada2 to act as a methyltransferase, suggesting that this activity occurred in a DNA sequence-independent manner (S4D Fig), and that methylation likely precedes the transcriptional response regulated by methylated Cada2.

To test if methylation of Cada2 was sufficient for its activation, we mutated the Cys$_{267}$ residue to a glycine (*cada2*^{C267G}). Such a mutation in the A-box of EcAda results in its constitutive activation as a transcription factor [35]. In contrast to the alanine mutant which abrogated *cada2* activity, we found that the *cada2*^{C267G} exhibited constitutive expression of the complete Cada2 regulon (S4E Fig). Consistent with this, the P$_{ccna_00746}$-*yfp* reporter showed robust expression even in the absence of methylation damage (Fig 5C). Furthermore, *cada2*^{C267G} did

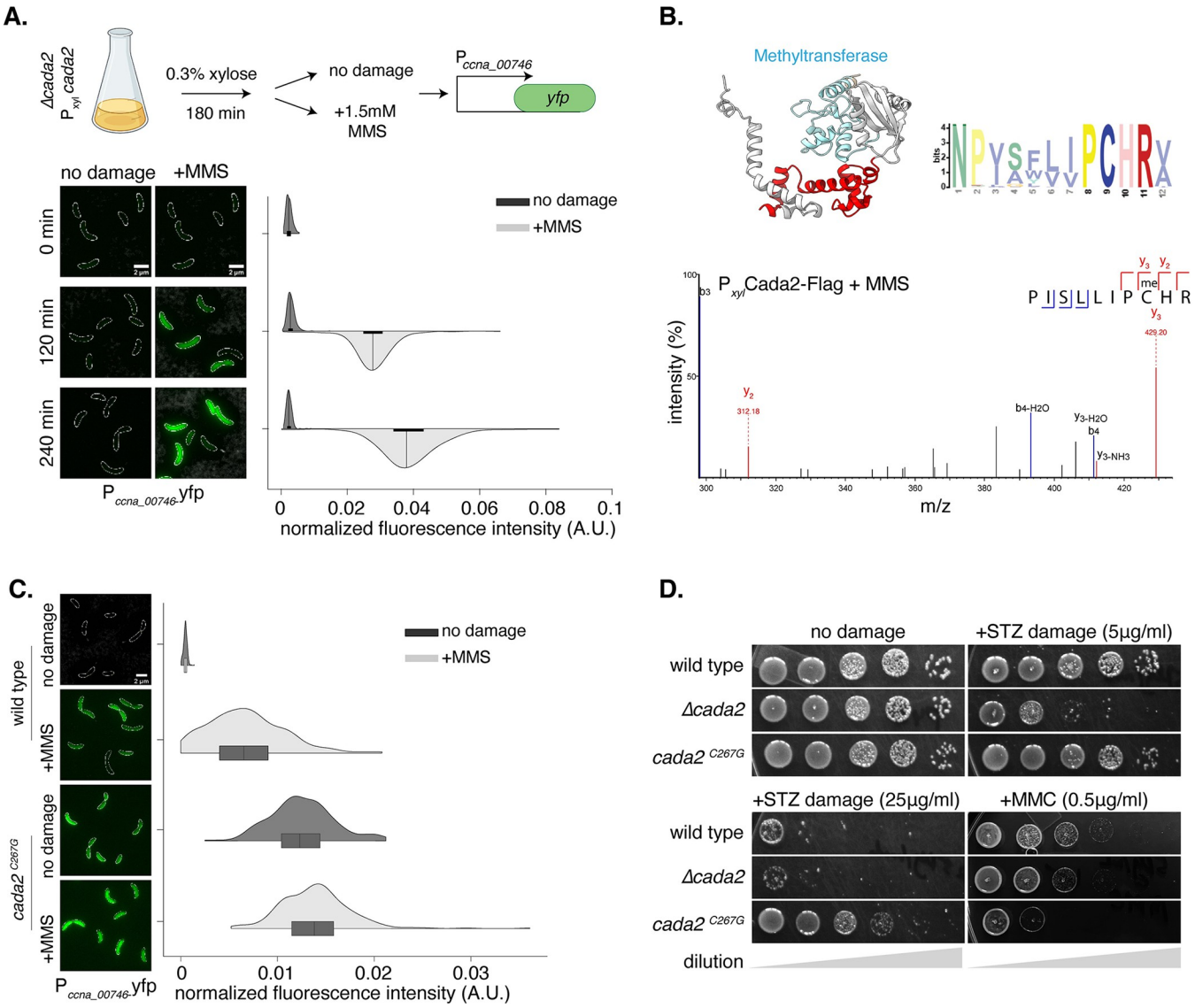

Fig 5. Cada2 is a methylation-dependent transcription factor. (A) [Top] Schematic representing experimental protocol for *cada2* overexpression analysis (see main text for details). [Bottom] Representative images and normalized fluorescence intensity for P_{ccna_00746}-*yfp* expression where *cada2* is overexpressed from a xylose-inducible promoter in a *Δcada2* background in presence/absence of 1.5 mM MMS damage ($n = 300$, from 3 biological replicates). The underlying data are available in S1 Data. **(B)** (Top left) Alphafold structure of Cada2 highlighting the presence of the methyltransferase domain is schematized. (Top right) MEME generated from 1,000 bacterial Cada2 proteins indicate conservation of the "PCHR" motif across methyltransferase domains. (Bottom) Zoomed in mass spectrometry fragmentation spectrum of a methylated peptide detected from a *cada2*-flag strain under 1.5 mM MMS exposure. The peptide (inset) includes the conserved "PCHR" motif of the Cada2 methyltransferase domain. Representative spectrum from 8 fragments across 2 biological replicates is shown. Methylation modification on Cys267 was detected in 6 out of 8 fragments. **(C)** [Left] Representative cells showing P_{ccna_00746}-*yfp* reporter induction in wild type (from Fig 1B) and $cada2^{C267G}$ mutant under 1.5 mM MMS damage. [Right] Violin plots show fluorescence intensity distribution normalized to cell area from single cells ($n = 300$, from 3 biological replicates). The underlying data are available in S1 Data. (D) Survival assay of $cada2^{C267G}$ mutant in the presence and absence of streptozotocin (5 and 25 μg/ml) and mitomycin C (0.5 μg/ml) damage.

https://doi.org/10.1371/journal.pbio.3002540.g005

not display growth defects associated with the *cada2* deletion or $cada2^{C267A}$ under methylation damage (Figs 5D and S4C). Thus, it is likely that Cada2 is activated as a regulator of the *Caulobacter* adaptive response via methylation of the conserved cysteine in its methyltransferase domain (PF01035).

Given that the adaptive response was constitutively ON in *cada2*^{C267G} cells, we asked whether such cells would perform better that wild type under methylation damage (as they would resemble cells "primed/adapted" to methylation damage). Indeed, in comparison to wild-type cells, the *cada2*^{C267G} mutant had a significant growth advantage under methylation damage and appeared to have no growth defect in untreated conditions (Fig 5D). However, despite this growth advantage, sequence analysis of the "PCHR" domain in Cada2-like proteins showed that a constitutive "ON" version of this protein is not found in nature (Fig 5B, 1,000 nonredundant genomes analyzed). We thus wondered whether there could be growth conditions where such a response must remain repressed. To test this, we subjected wild-type and *cada2*^{C267G} mutant cells to MMC-treatment, an unrelated DNA damage condition where Cada2 activity is not essential and would ordinarily be in an "OFF" state. In contrast to the growth advantage under methylation damage, we observed that *cada2*^{C267G} cells were considerably compromised in growth in comparison to wild type when exposed to MMC (that induces mono-adducts and intra-strand crosslinks) (Fig 5D). While the molecular mechanism underlying Cada2 activation under methylation damage as well as the crosstalk in case of the *cada2*^{C267G} mutant requires further investigation, these observations support the possibility that methylation-specific regulation of Cada2 is important, due to antagonistic effects associated with this response being mis-regulated in other stress conditions.

Regulatory features of Cada2 are conserved across bacteria and distinct from the EcAda paradigm

Taken together, the following regulatory features constitute the *Caulobacter* adaptive response: (a) Promoter regions of genes under this regulon carry a B-box motif (GCAA) and an X-box motif (CGG). (b) Cada2 protein N-terminus encodes B-box binding and RLHD domains that enable it to interact with promoter regions in a sequence-specific manner. (c) Cada2 protein C-terminus encodes a methyltransferase domain. Methylation of the conserved cysteine in this domain activates Cada2 as a transcription factor, likely via enabling interaction with RNA polymerase.

We asked how extrapolatable these observations would be across bacterial species. For this, we collated the presence or absence of the features listed above for a set of bacteria (including closely related and unrelated) that carry either an EcAda-like or Cada2-like protein (Fig 6A). To our surprise, we found that every organism that encoded an EcAda-like protein carried all the associated regulatory features for an *E. coli*-like adaptive response (Fig 6A), while organisms with Cada2-like protein had regulatory features observed in case of the *Caulobacter* adaptive response (Fig 6A).

The conservation of the protein and promoter sequence across Cada2-like proteins motivated us to test whether these proteins are functionally inter-changeable. For this, we expressed a Cada2-like protein from *Myxococcus xanthus* (Cada2^{myxo}) in *Caulobacter* lacking its own *cada2*. Cada2^{myxo} has overall low identity (44%) to the *Caulobacter* Cada2 protein; however, it shares high similarity in terms of conserved regulatory features (Fig 6A). As a control, we expressed EcAda in the same background to assess whether EcAda could complement the absence of Cada2. We found that Cada2^{myxo} was able to fully complement the absence of *Caulobacter cada2* under methylation damage, as assessed by cell survival as well as promoter activity of P_{ccna_00746}-*yfp* (Figs 6B and S5A). In contrast, EcAda was not able to complement a *cada2* deletion phenotype (Fig 6B). This protein, expressed in *Caulobacter*, was fully proficient in driving expression of *yfp* from a promoter carrying the EcAda DNA binding sequence in a methylation-dependent manner (Fig 6B). However, the Cada2^{myxo} showed Cada2-specific activity and could not induce *yfp* expression from the EcAda reporter (Fig 6B). Together, these

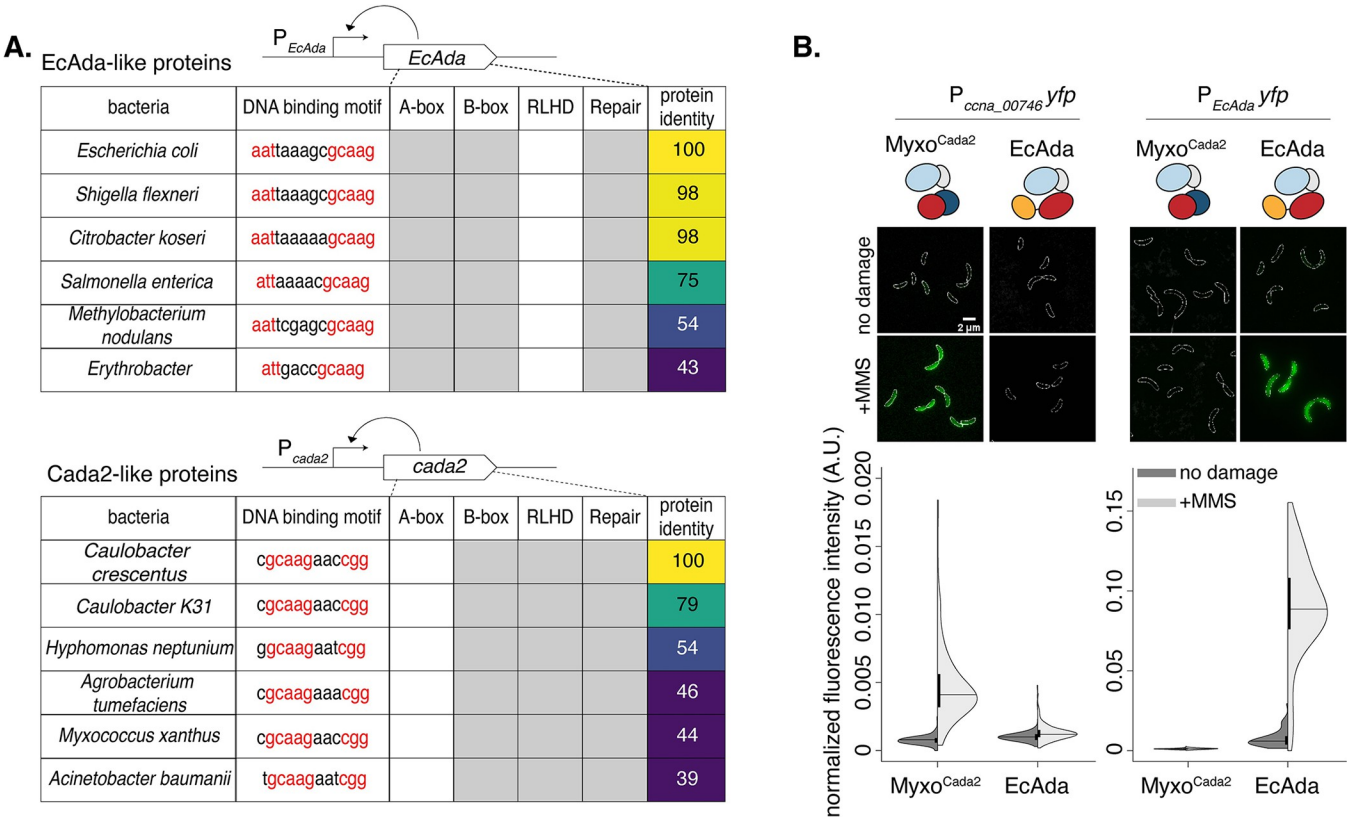

Fig 6. Regulatory features of Cada2 are conserved across bacteria and distinct from the EcAda paradigm. (A) Regulatory Ada methyltransferases such as EcAda and Cada2 are auto-regulatory; as represented in schematics, these proteins regulate activity of their own promoters (in a methylation-dependent manner). This allowed us to predict presence/absence of regulatory features in their respective promoters. The table represents conservation of regulatory features in the protein (A, B, or RLHD amino acid sequences) and in the cognate promoter sequences (A, B, or X Box DNA motif) of identified EcAda-like and Cada2-like proteins. Presence (gray) or absence of features (blank) is indicated along with % protein sequence. **(B)** Cada2 from *Myxococcus* (Myxo^{Cada2}) can drive P_{ccna_00746}-*yfp* induction in *Δcada2* in response to MMS damage, but cannot drive *yfp* expression from the *EcAda* promoter. Conversely, Ada from *E. coli* can activate expression of *yfp* from the *EcAda* promoter, but not from the P_{ccna_00746} promoter. The underlying data are available in S1 Data.

https://doi.org/10.1371/journal.pbio.3002540.g006

data illustrate the specificity, sufficiency, and necessity of the identified regulatory features in the control of the adaptive responses across bacteria.

Discussion

In this study, we identify Cada2, a novel methylation damage-specific transcription factor conserved across all major phyla of bacteria. We further implicate a critical and conserved role for PTM-based activation of this class of bacterial transcription factors, albeit with varying mechanisms of action (Cada2 versus EcAda). We hypothesize that the organizational diversity observed across regulators of the adaptive response pathways are likely an adaptation to the physiological features of bacteria. The downstream adaptive response is, however, robust to this variability.

Diverse mechanisms, unifying function

The mechanistic differences between Cada2 and EcAda would suggest that the response dynamics would also be dissimilar. Cada2 and EcAda use distinct DNA-binding domains to associate with their cognate promoters (B-box and "RLHD" domains versus A and B-box binding domains, respectively [35]). The 2 proteins are also activated as transcription

factors in non-overlapping ways: In case of EcAda, methylation of the cysteine residue within the sequence-specific DNA-binding A-box domain activates it as a transcription factor [59,61]. In contrast, Cada2 activation occurs via PTM of the cysteine in its methyltransferase domain that is distally located from the sequence-specific DNA-bindings domain. Furthermore, unlike EcAda, unmethylated Cada2 appears to form sequence-specific interactions with its own promoter (Figs 3A and S2A). Presence of electrostatic repulsion (as in EcAda) would hinder this interaction. This would suggest that Cada2 acts via an alternate, non-electrostatic mechanism.

Yet despite these contrasts, the response kinetics are invariant between the 2 organisms. The adaptive response of *Caulobacter* is tightly regulated, showing low levels of expression (Ada OFF) in the absence of methylation DNA damage. Upon exposure to methylating agents, *Caulobacter* cells also elicit a delayed adaptive response, subsequent to the SOS response. Similar to the *E. coli* Ada response [31], the *Caulobacter* response exhibits bi-stability with a subpopulation of cells resembling expression of Ada OFF cells even in the presence of DNA methylation damage. Interestingly, the OFF population is abolished in a strain overexpressing Cada2, mirroring observations made in case of EcAda [31] (Figs 1B and 5A). The factors influencing the differences in the mechanisms of action of the response regulators, while retaining the defining features of the response itself are an important and exciting avenue for future investigations.

Requirement for a methylation damage-specific response in bacteria

How pervasive is the adaptive response to methylation damage? Our computational analysis reveals that the regulatory Ada methyltransferase is widely conserved, but it appears to have diverse domain organizations (S3D and S3E Fig). Given the apparent modularity of these domains (S3D and S3E Fig), it is tempting to speculate that such diverse organizations arise as a result of domain-shuffling from a common ancestor protein [62–64]. Despite their prevalence, multiple regulators rarely occur on the same genome. A case in point is the comparison between EcAda and Cada2. We were able to detect only approximately 7% of genomes possessing both proteins. The genomes that encoded both EcAda and Cada2 displayed varied and unique regulatory circuits, including autoregulation as well as cross-regulation (S3F Fig).

What drives the observed diversity? Comparison of EcAda and Cada2 might provide a compelling hypothesis. Both EcAda and Cada2 share the B-box motif in their cognate promoters. However, instead of the AT-rich A-box motif observed in the promoters of EcAda regulon genes, Cada2 binds to a GC-rich X-box motif. This correlates well with the GC-content of the genomes that encode these proteins. The GC-content of the *E. coli* genome is approximately 50% (low), while the GC-content of the *Caulobacter* genome is approximately 67% (high). This hypothesis is consistent with previous studies that have also suggested the interdependence of genome GC-content and presence/absence of specific DNA repair pathways [65–68].

Thus, it is likely that distinguishing physiological characteristics of bacteria warrant organizational adaptation and domain reorganizations of regulatory Ada methyltransferases. However, regardless of the diversity observed at a regulator level, the adaptive response-like pathways seem to retain certain characteristics as emphasized previously. In this context, we highlight the PTM-based transcriptional switch required for activation and regulation of the response. It is possible that costs (e.g., Fig 5D) or fitness advantages associated with this response serve as constraining forces that render this system robust to variations in regulator organization, and yet drive the retention of key regulatory features. Together, the conservation of bacterial adaptive response mechanisms, albeit in diverse regulatory forms, underscores the fundamental requirement for a dedicated methylation-specific DNA damage response.

Materials and methods

Bacterial strains and growth conditions

All strains, plasmids, and oligos used in this study are listed in S1–S3 Tables, respectively. *Caulobacter crescentus* cells were grown at 30˚C in peptone yeast extract (PYE) media (0.2% peptone, 0.1% yeast extract, and 0.06% $MgSO_4$) and supplemented with antibiotics or inducers, as required, at suitable concentrations. For induction of *cada2* expression, 0.3% xylose was introduced into the cultures, unless otherwise stated.

Fluorescence microscopy and image analysis

Imaging was performed on an epifluorescence, wide-field microscope (Eclipse Ti-2E, Nikon) with a 60×/1.4 NA oil immersion objective and a motorized stage. pE-4000 (CoolLED) was used as the LED excitation source. Exposure times for all *yfp* samples (λ = 490nm) was 500 ms and exposure was maintained at 50% of LED power. Images were captured using a Hamamatsu Orca Flash 4.0 camera. Focus was ensured via an infrared-based Perfect Focusing System (Nikon). Samples were prepared as detailed in [69]. For time course microscopy, 1 ml of cultures were aliquoted at various time points, pelleted and resuspended in appropriate volume of growth media, and 2 μl of the resuspended cells were spotted on 1% agarose (Invitrogen ultrapure) pads and imaged. For time lapse microscopy, 2 μl of culture were spotted on 1.5% low-melting GTG agarose pads supplemented with PYE media and 1.5 mM MMS damage. Throughout the duration of the time lapse, samples were grown in an OkoLab incubation chamber maintained at 30˚C and imaged at regular intervals. Cell segmentation and fluorescence intensity analysis was carried out via Oufti software [70].

RNA-sequencing

Sample collection. Overnight cultures of *Caulobacter* cells were back-diluted to 0.025 OD_{600} and grown at 30˚C for 3 h. When OD_{600} of culture was approximately 0.1, DNA damage (MMS (1.5 mM)/MMC (0.25 μg/ml)/norfloxacin (8 μg/ml)) was introduced. Samples from 4 biological replicates were collected for no damage conditions and 2 biological replicates for samples in the presence of damage, and 2 ml of *Caulobacter* cells were harvested and spun at 10,000 g for 5 min. Supernatant was then discarded and the cell pellets were snap frozen with liquid nitrogen and stored at –80˚C until the samples were further processed. Cells were harvested at 0, 20, and 40 min post DNA damage exposure.

RNA-extraction. RNA was extracted from the cell pellets according to a previous study [58]. Cells were lysed using 400 μl of pre-heated Trizol in a thermomixer at 65˚C, 2000 rpm. The lysed cells were transferred to –80˚C for a minimum of 30 min and were then centrifuged at 4˚C. The supernatant was carefully transferred to 100% ethanol of equivalent volume. RNA was then extracted via the protocol mentioned in the Direct-zol RNA MiniPrep (Zymo, Cat. no. R2052) kit. The extracted RNA was then subjected to DNAse treatment in order to remove any genomic DNA from the sample. Total RNA from the samples was purified using the RNA Clean & Concentrator-25 (Zymo, Cat. no. R1018) kit. Integrity of the RNA was tested using Bioanalyzer instrument. mRNA was isolated from the total sample using the RiboMinus kit (Thermo, Cat. no. K155004) and submitted for RNA-seq at the NCBS next generation sequencing facility. Details regarding kits used for RNA library preparation and the sequencing platform used for the RNA-seq experiment are summarized in S4 Table.

RNA-sequencing analysis

Raw reads for the sequencing results were obtained as fastq files. The reference genome sequence (.fna) and annotation (.gff) files for the *Caulobacter crescentus NA1000* strain (accession number: NC_011916.1) were downloaded from the NCBI file transfer protocol (ftp) website ("ftp.ncbi.nlm.nih.gov"). The raw read quality was checked using the FastQC software (version 0.11.5). Burrows–Wheeler Aligner (BWA) (version 0.7.17-r1188) was used to index the reference genome. Reads with raw read quality > = 20 were aligned to the indexed genome using the BWA aln -q option. Samtools (version 0.1.7) was used to filter out multi-mapped reads. Bedtools (version 2.26.0) was used to calculate the reads count per gene using the annotation file (.bed) [71]. The normalization and differential gene expression analysis for the samples with replicates were carried out using edgeR [72]. Samples were grouped together with their cognate biological replicates. To estimate differential gene expression, a quasi-likelihood F-test was applied comparing normalized read values from no damage samples to samples exposed to DNA damage. For $cada2^{C267G}$ data, samples were normalized by calculating Reads per Kilobase per Million mapped reads (RPKM) along with a corresponding wild-type control. Fold change for $cada2^{C267G}$ samples was calculated by comparing RPKM values with the wild-type control (in the absence of damage). A gene was defined as up-regulated based on 2 thresholds; in a certain condition if the $\log_2$ fold change (from control) of the gene is greater than or equal to 1 (i.e., the gene content has doubled in this condition) and if the false discovery rate across replicates was less than 0.05. Similarly, a gene was down-regulated in that condition if the gene's $\log_2$ fold change from control was less than −1 and the false discovery rate across replicates is less than 0.05. DGE analysis was done using Google Colab with R version 4.2.3 (2023-03-15).

Western blotting

Western blotting was performed as described in a previous study [73]. Briefly, overnight cultures in PYE media was back-diluted to 0.025 OD_{600} and grown at 30˚C for 3 h. At approximately 0.1OD_{600}, 1.5 mM MMS was added to the culture; 0, 2, and 4 h post introduction of DNA damage, cells corresponding to 0.6 OD_{600} were harvested. The culture was pelleted following which supernatant was discarded. The pelleted cells were resuspended in 200 μl of lysis buffer [500 mM Tris-HCL (pH 6.8), 8% SDS, 40% glycerol, 2-Mercaptoethanol, and sufficient amount of bromophenol blue] and heated at 95˚C for 3 min. This was followed by 2 rounds of 30 s vortex spins, each followed by short spins. The resuspended cells were again heated to 95˚C post which samples were loaded on to a 10% SDS-PAGE gel and subjected to electrophoresis. The resolved protein bands were then transferred on to a PVDF membrane and probed with anti-flag antibodies for estimation of Cada2 levels. The same samples were also probed with anti-RpoA antibodies in order to verify uniform loading of samples. The blots were developed using SuperSignal West PICO PLUS chemiluminescent substrate.

Mass spectrometry

Sample collection. Cells overexpressing flag-tagged *wild-type cada2* or its mutants ($cada2_{R68A}$, $cada2_{R114A}$) from a xylose-inducible promoter in a *Δcada2* background were used. Overnight cultures grown in PYE media containing gentamycin was back-diluted to 0.025 OD_{600} and grown at 30˚C for 3 h. At approximately 0.1OD_{600}, 1.5 mM MMS and 0.3% xylose were added to the culture and culture volume corresponding to 0.6 OD_{600} was collected after 4 h of treatment. The culture was pelleted following which supernatant was discarded. The pelleted cells were resuspended in 200 μl of lysis buffer [500 mM Tris-HCL (pH 6.8), 8% SDS, 40% glycerol, 2-Mercaptoethanol, and sufficient amount of bromophenol blue] and heated at

95˚C for 3 min. This was followed by 30 s vortex spin, followed by a short spin. The above 2 steps were repeated twice. The resuspended cells were again heated to 95˚C post which samples were loaded on to a 10% SDS-PAGE gel and subjected to electrophoresis. The gel was appropriately resolved and stained with Coomassie dye. Bands corresponding to the size of Cada2-3x-Flag (approximately 37 kDa) were excised from the gel, resuspended in milliQ water, and submitted to the NCBS mass spectrometry facility for further processing.

Sample preparation and mass spectrometry. The excised gel fragments were first washed thrice with LC-MS water post which the supernatant water was discarded. Next, the gel pieces were chopped followed by destaining with 1:1 100 mM Triethyl ammonium bicarbonate (TEAB) with 100% acetonitrile (ACN). Post destaining, the supernatant was discarded. Approximately 200 μl of 100% ACN was then added and the sample was kept at room temperature till the gel pieces shrunk, following which the supernatant was discarded. The previous step of ACN addition was repeated for a second time. The gel pieces were allowed to dry for a few minutes at room temperature. Next, the gel was suspended in a minimum of 500 ng of Trypsin with 100 to 200 μl of 100 mM TEAB. The samples were then kept for overnight trypsin digestion at 37˚C, and 100 μl of 100% ACN with 0.1% formic acid was then introduced to the digested sample and the digested sample were sonicated for 5 min. The supernatant was then collected into a fresh tube. The post-digestion step was repeated for a second time. The collected supernatant was then dried using a SpeedVac vacuum concentrator and reconstituted in 0.1% formic acid. This was followed by desalting post which samples were injected into Orbitrap Fusion Tribrid Mass Spectrometer coupled to a Thermo EASY nanoLC 1200 chromatographic system for mass spectrometry. The results were analyzed using the PeakStudio 8.0 for identifying peptide fragments and their associated PTMs (if any).

Chromatin immunoprecipitation sequencing (ChIP-Seq)

Fifty ml PYE was sub-inoculated with overnight *Caulobacter crescentus* cultures to OD_{600} of 0.025, cultures were grown at 30˚C for another 3 h with shaking at 250 rpm until reaching OD_{600} of 0.1. MMS (final concentration of 1.5 mM) was then added, and the cultures were outgrowth for another 2 or 4 h before fixation with formaldehyde. Controls of no MMS treatment were also included. Cells were fixed with formaldehyde (final concentration of 1%) at room temperature for 30 min, then quenched with 0.125 M glycine for another 15 min at room temperature. Cells were washed 3 times with 1× PBS (pH 7.4) and resuspended in 1 ml of lysis buffer [20 mM K-HEPES (pH 7.9), 50 mM KCl, 10% glycerol, and Roche EDTA-free protease inhibitors]. ChIP-seq was performed as described previously in Tran and colleagues [74]. Briefly, the cell suspension was sonicated on ice using a Soniprep 150 probe-type sonicator (11 cycles, 15s on, 15s off, at setting 8) to shear the chromatin to below 1 kb, and the cell debris was cleared by centrifugation (20 min at 13,000 rpm at 4˚C). The supernatant was then transferred to a new 2 ml tube and the buffer conditions were adjusted to 10 mM Tris-HCl (pH 8), 150 mM NaCl, and 0.1% NP-40. Fifty microliters of the supernatant were transferred to a separate tube for control (the input fraction) and stored at −20˚C. Meanwhile, antibodies-coupled beads were washed off storage buffers before being added to the above supernatant. α-M2 FLAG antibodies coupled to sepharose beads (Merck, United Kingdom) were employed for ChIP-seq of Cada2-FLAG and RpoC-FLAG. Briefly, 100 μl α-FLAG beads were washed off the storage buffer by repeated centrifugation and resuspension in IPP150 buffer [10 mM Tris-HCl (pH 8), 150 mM NaCl, and 0.1% NP-40]. Beads were then introduced to the cleared supernatant and incubated with gentle shaking at 4˚C overnight. Beads were then washed 5 times at 4˚C for 2 min each with 1 ml of IPP150 buffer, then twice at 4˚C for 2 min each in 1× TE buffer [10 mM Tris-HCl (pH 8) and 1 mM EDTA]. Protein-DNA complexes were then eluted twice

from the beads by incubating the beads first with 150 μl of the elution buffer [50 mM Tris-HCl (pH 8), 10 mM EDTA, and 1% SDS] at 65˚C for 15 min, then with 100 μl of 1X TE buffer + 1% SDS for another 15 min at 65˚C. The supernatant (the ChIP fraction) was then separated from the beads and further incubated at 65˚C overnight to completely reverse the crosslink. The input fraction was also de-crosslinked by incubation with 200 μl of 1X TE buffer + 1% SDS at 65˚C overnight. DNA from the ChIP and input fraction were then purified using the PCR purification kit (Qiagen) according to the manufacturer's instruction, then eluted out in 40 μl water. Purified DNA was then made into libraries suitable for Illumina sequencing using the NEXT Ultra II library preparation kit (NEB, UK). ChIP libraries were sequenced on the Illumina Hiseq 2500 at the Tufts University Genomics facility.

ChIP-seq analysis

Raw reads for the ChIP-seq results were obtained as fastq files. BWA was used in order to align the reads to the *Caulobacter* genome similar to analysis of RNA-seq data. The aligned reads in. bam format were converted into a .bed file format using the bamtobed option of Bedtools. Coverage at each nucleotide position was calculated using the Bedtools genomecov command. Subsequently, ChIP-seq profiles from these coverage files were plotted using a custom python script. Pearson correlation between ChIP-seq profiles for *cada2 flag* and *rpoC flag* was calculated using Numpy [75]. Peakzilla was used in order to identify bona fide enrichment peaks [76], and.bed files from *cada2 3xflag* or rpoC 3xflag ChIP-seq under MMS exposure were compared to untagged *wild type* cells treated under identical conditions for peak-calling.

ChIP and quantitative PCR analysis

For quantitative-PCR (q-PCR) analysis of *cada1* and *cada3* promoter enrichment in ChIP experiments, purified DNA from both input and ChIP fractions was diluted 1:4 in water and 1 μl was used for qPCR using a SYBR Green JumpStart *Taq* ReadyMix (CAT S4438, Merck, UK) and a BioRad CFX96 instrument. Fold enrichments were calculated using the comparative Ct method (ΔΔCt) and represent the relative abundance of *cada1* and *cada3* promoter DNA compared to *rpoD* DNA as a negative control. All fold enrichment values represent the average of 3 biological replicates.

The following oligos were used in qPCR reactions:
cada1 (fw, TAGGACGCGACTGCTGA, rv, TCCTTTCGTGAGGAGACCA)
cada3 (fw, ATCGCCCGCATGGAATAC, rv, AGAAGGAAGCTACTACCGGAT)
rpoD (fw, TCAGGCCAAGAAGGAAATGG, rv, GCCTTCATCAGGCCGATATT).

Survival assay

Overnight cultures of *Caulobacter* strains were back-diluted to OD_{600} 0.1. The cultures were incubated at 30˚C for 3 h. All cultures were then normalized to 0.3 OD_{600} and serially diluted in 10-fold increments (10^{-1}–10^{-8}), and 6 μl of each dilution was spotted on PYE plates containing appropriate chemicals (DNA damaging agents/antibiotics/inducer). The plates were incubated at 30˚C for 48 h post which survival of the respective strains was determined by counting the number of spots.

Computational predictions of protein structures via AlphaFold

Three-dimensional structural predictions for EcAda, Cada1, Cada2, and Cada3 were made by ColabFold (AlphaFold2 coupled with MMSeq2 hosted on Google Colaboratory) [56,57]. Primary sequence of each protein was used as query sequence. Default parameters were used for

other settings. Structural models shown in this paper are models with the highest pLDDT score.

Computational analysis of protein conservation

All "complete" and "latest" (assembly_summary.txt; as of January 2017) genome information files for approximately 6,000 bacteria were downloaded from the NCBI ftp website using in-house scripts for whole-genome sequences (.fna), protein coding nucleotide sequences (.fna), RNA sequences (.fna), and protein sequences (.faa). All the organisms were assigned respective phylum based on the KEGG classification (https://www.genome.jp/kegg/genome.html; as of May 2018).

For identification of Ada domains and Ada variants, initial blastp was run using each of the 4 protein domain sequences of EcAda from *Escherichia coli MG1655* as the query sequence against the UniprotKB database with an E-value cutoff of 0.0001. The 4 domains of Ada include A-box, B-box, RNAse-like, and repair. The top 1,000 full-length sequence hits were downloaded from UniProt for all the 4 domains. A domain multiple sequence alignment (MSA) was made using phmmer–A option with the top 1,000 hits as the sequence database and *E. coli* domain sequences as the query. An hmm profile was built using the hmmbuild command for the MSA obtained in the previous step. To find domain homologs, hmmsearch command with an E-value cutoff of 0.0001 was used with the hmm profile as the query against a database of 5,973 bacterial genome sequences. These homolog searches were done using HMMER package v3.3. Different Ada variants were identified based on the assignment of the different combinations of Ada domains to the same protein.

A phylogenetic species tree was constructed using 16S rRNA sequences. One sequence per genome was extracted from a multi-fasta file. An MSA was built using muscle v3.8.31 with default options, followed by alignment trimming using BMGE v1.12. Using IQTREE v1.6.5, a maximum likelihood-based phylogeny was built with the best model chosen using ModelFinder (-m MF option) against 285 other models. Branch supports were assessed using both 1,000 ultrafast bootstrap approximations (-bb 1,000 –bnni option) and SH-like approximate likelihood ratio test (-alrt 1,000 option). Final tree for visualization was pruned to contain 4 major phyla—Proteobacteria, Actinobacteria, Firmicutes, and Bacteroidetes. Online tool Itol [77] was used for visualization and EcAda and Cada2 presence/absence was overlaid on the phylogeny.

For identification of essential regulatory motifs of Cada2 and EcAda, protein blast was run on the coding sequences of EcAda or Cada2 against the UniprotKB database with an E-value cutoff of 10. The top 1,000 full-length sequence hits were downloaded from UniProt for both proteins and were submitted to the MEME discovery program in order to identify conserved features.

Supporting information

S1 Fig. A methylation-specific DNA damage response in Caulobacter. (A) Heat map of $\log_2$FC values for genes up-regulated in wild-type cells following MMS, MMC, or Norfloxacin treatment. **(B)** [Left] Representative cells showing P_{ccna_00746}-*yfp* reporter induction upon exposure to STZ, MMC, norfloxacin, and HU. [Right] Violin plots show fluorescence intensity distribution normalized to cell area from single cells (n = 300, from 3 biological replicates). The underlying data are available in S1 Data. **(C)** [Left] Representative cells showing P_{ccna_00746}-*yfp* reporter induction in *ΔdriD* background under 1.5 mM MMS damage. [Right] Violin plots showing fluorescence intensity distribution normalized to cell area from single cells (n = 300, from 3 biological replicates). The underlying data are available in S1 Data. **(D)**

[Top] Schematic of the experimental protocol for testing the adaptive property of the *Caulobacter* methylation-specific damage response. Cultures of P_{ccna_00746}-*yfp* cells were exposed to a sublethal dose of 0.5 mM either MMS (adapted) or no MMS (non-adapted). The cells were subsequently exposed to a higher dose of 1.5 mM MMS on agarose pads supplemented with PYE medium. [Bottom] Normalized fluorescence intensity kinetics was measured via time lapse microscopy over 3 h of lethal MMS exposure. Dotted lines (in dark) indicate mean time trace of induction kinetics while the shaded region (in light) indicates the standard deviation of all time traces for the respective conditions (here and for all other time lapse data) ($n = 25$). The underlying data are available in S1 Data. **(E)** Survival assay of individual deletions of *cada* genes with and without MMS exposure (1.5 mM).
(TIF)

S2 Fig. Cada2 associates with adaptive response promoters in a sequence-specific manner. **(A)** Survival assay of *cada2-3x-flag* strain in the presence or absence of STZ damage (5 μg/ml). **(B)** Western blot showing Cada2-3x-flag levels at 0, 2, and 4 h after 1.5 mM MMS exposure. As a loading control, RpoA is probed. **(C)** ChIP-seq profiles for wild type (control), Cada2-3x-Flag, or RpoC-3x-Flag ±2.5 kb around *cada2* CDS before (no damage) and after (+damage) exposure to 1.5 mM MMS. **(D)** Zoomed-in ChIP-seq profiles for Cada2-3x-Flag and RpoC-3x-Flag ±2.5 kb around *cada2* CDS before and after exposure to 1.5 mM MMS.
(TIF)

S3 Fig. Cada2-like proteins are widespread and encode a novel DNA-binding domain. **(A)** (Top) Domain organization and predicted Alphafold structure of Cada2 and EcAda is indicated. (Bottom) Closeup of the EcAda and Cada2 regulatory domain reveals that the Cada2 B-box binding domain (possessing the "SPFHQR" amino acid sequence) and the newly identified sequence-specific binding domain of Cada2 (possessing the "RLHD" amino acid sequence) are part of a helix-turn-helix domain similar to the B-box binding domain of EcAda. The position of these conserved motifs are highlighted and labeled as a ball-and-stick in the overall ribbon representation of the models. **(B)** Western blot of flag-tagged Cada2 mutants ($Cada2^{R68A}$ and $Cada2^{R114A}$) overexpressed in a *Δcada2* strain from a xylose-inducible promoter treated with MMS damage. **(C)** Survival assay of $cada2^{R68A}$ and $cada2^{R114A}$ with or without STZ damage. **(D)** [Left] Prevalence of EcAda-like, Cada2-like, and AdaA-like proteins estimated from a curated, nonredundant database of bacterial genomes is analyzed at the genus level. Numbers represent the presence of these proteins for the major bacterial clades. [Right] domain organization of the respective adaptive response regulatory proteins analyzed here. **(E)** [Left] As (D) for a curated, nonredundant database of bacterial genomes at the species level. [Right] Prevalence of other adaptive response methyltransferases estimated from a curated, nonredundant database of bacterial genomes is analyzed at the species level. **(F)** Table represents presence and absence of EcAda-like and Cada2-like proteins and their co-occurrence. In the instances of Cada2-EcAda co-occurrence, potential regulatory circuits predicted via identifying Cada2 and EcAda binding motif in their cognate promoters are shown.
(TIF)

S4 Fig. Cada2 is a methylation-dependent transcription factor. **(A)** Bacterial two-hybrid assay to test interaction between Cada2 and RNA polymerase subunits. T18-Cada2 was tested for interaction with T25-RNA polymerase holoenzyme subunits with and without 1.5 mM MMS. The presence of red colonies indicated positive interaction. As a positive control AddA (T18) and AddB (T25) are used, and empty vectors (T18 and T25) are used as negative control. Representative images from 2 independent repeats are shown. **(B)** [Left] Representative cells showing P_{ccna_00746}-*yfp* reporter induction in wild type and $cada2^{267A}$ mutant background

under 1.5 mM MMS damage. [Right] Violin plots showing fluorescence intensity distribution normalized to cell area from single cells in the presence (dark) and absence of damage (light) (n = 200, from 2 biological replicates). Wild-type data are represented again from Fig 1B. The underlying data are available in S1 Data. **(C)** Survival assay of the *cada2 mutants* (*cada2*^{C267A} or *cada2*^{C267G}) with and without STZ damage (5 μg/ml). (D) [Above] Table representing peptides corresponding to the Cada2 methyltransferase domain bearing the "PCHR" motif as identified via mass spectrometry. In the absence of MMS, peptides corresponding to wild-type Cada2-flag are unmethylated. Upon exposure to MMS damage, peptides corresponding to wild type as well as mutant Cada2-3x-Flag exhibit methylation at the Cys^{267} residue. [Below] Representative mass spectrometry fragmentation patterns for peptides in the above table. In case of wild type (no damage) representative spectrum from 3 fragments across 2 biological replicates is shown. No methylation modification on Cys^{267} was detected. In case of Cada2^{R68A} representative spectrum from 4 fragments across 2 biological replicates is shown. Methylation modification on Cys^{267} was detected in 3 out of 4 fragments. In case of Cada2^{R114A} representative spectrum from 7 fragments across 2 biological replicates is shown. Methylation modification on Cys^{267} was detected in 5 out of 7 fragments. **(E)** Heat maps represent $\log_2$FC values for Cada2 regulon genes in wild type and *cada2*^{C267G} cells under MMS exposure.
(TIF)

S5 Fig. Regulatory features of Cada2 are conserved across bacteria and distinct from the EcAda paradigm. (A) Survival assay of *Δcada2* strain overexpressing *cada2*^{caulo} or *cada2*^{myxo} (from xylose-inducible promoter) under methylation damage (5 μg/ml STZ). Survival of these strains was compared to a control strain comprising of an empty vector in a *Δcada2* background.
(TIF)

S1 Table. Strains used in present study.
(DOCX)

S2 Table. Plasmids used in present study.
(DOCX)

S3 Table. Oligos used in present study.
(DOCX)

S4 Table. Kits used for preparing RNA libraries and sequencing platform used for the RNA-seq experiment.
(DOCX)

S1 Data. Numerical data underlying graphs and plots shown in main and supplementary figures.
(XLSX)

S1 Raw Images. Uncropped images for western blot in S2B and S3B Figs.
(PDF)

Acknowledgments

We thank Dr. Asha Joseph for contribution to RNA-sequencing experiments and Kaustav Mitra for carrying out initial bacterial-two-hybrid experiments. We are grateful to members of the AB lab for helpful comments and discussions. We acknowledge support from NCBS Mass Spectrometry facility and the Next Generations Sequencing facility for key experiments.

Author Contributions

Conceptualization: Aditya Kamat, Tung B. K. Le, Anjana Badrinarayanan.

Data curation: Aditya Kamat, Tung B. K. Le, Anjana Badrinarayanan.

Formal analysis: Aditya Kamat, Ngat T. Tran, Mohak Sharda, Neha Sontakke, Tung B. K. Le.

Funding acquisition: Tung B. K. Le, Anjana Badrinarayanan.

Investigation: Aditya Kamat, Ngat T. Tran, Tung B. K. Le, Anjana Badrinarayanan.

Methodology: Aditya Kamat, Mohak Sharda, Neha Sontakke.

Project administration: Anjana Badrinarayanan.

Resources: Aditya Kamat, Neha Sontakke.

Software: Aditya Kamat.

Supervision: Anjana Badrinarayanan.

Validation: Aditya Kamat, Tung B. K. Le.

Visualization: Aditya Kamat, Tung B. K. Le, Anjana Badrinarayanan.

Writing – original draft: Aditya Kamat, Anjana Badrinarayanan.

Writing – review & editing: Aditya Kamat, Ngat T. Tran, Mohak Sharda, Neha Sontakke, Tung B. K. Le, Anjana Badrinarayanan.

References

1. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. Neuropsychopharmacol. 2013; 38:23–38. https://doi.org/10.1038/npp.2012.112 PMID: 22781841
2. Seong HJ, Han S-W, Sul WJ. Prokaryotic DNA methylation and its functional roles. J Microbiol. 2021; 59:242–248. https://doi.org/10.1007/s12275-021-0674-y PMID: 33624263
3. Robertson KD, Jones PA. DNA methylation: past, present and future directions. Carcinogenesis. 2000; 21:461–467. https://doi.org/10.1093/carcin/21.3.461 PMID: 10688866
4. Mattei AL, Bailly N, Meissner A. DNA methylation: a historical perspective. Trends Genet. 2022; 38:676–707. https://doi.org/10.1016/j.tig.2022.03.010 PMID: 35504755
5. Sánchez-Romero MA, Cota I, Casadesús J. DNA methylation in bacteria: from the methyl group to the methylome. Curr Opin Microbiol. 2015; 25:9–16. https://doi.org/10.1016/j.mib.2015.03.004 PMID: 25818841
6. Bignami M O'Driscoll M, Aquilina G, Karran P. Unmasking a killer: DNA O6-methylguanine and the cytotoxicity of methylating agents. Mutat Res/Rev Mutat Res. 2000; 462:71–82.
7. Boiteux S, Laval J. Mutagenesis by alkylating agents: Coding properties for DNA polymerase of poly (dC) template containing 3-methylcytosine. Biochimie. 1982; 64:637–641. https://doi.org/10.1016/s0300-9084(82)80103-x PMID: 6814512
8. Larson K, Sahm J, Shenkar R, Strauss B. Methylation-induced blocks to in vitro DNA replication. Mutat Res/Fund Mol Mech Mutagenesis. 1985; 150:77–84. https://doi.org/10.1016/0027-5107(85)90103-4 PMID: 4000169
9. Sedgwick B. Repairing DNA-methylation damage. Nat Rev Mol Cell Biol. 2004; 5:148–157. https://doi.org/10.1038/nrm1312 PMID: 15040447
10. Little JW, Gellert M. The SOS regulatory system: Control of its state by the level of RecA protease. J Mol Biol. 1983; 167:791–808. https://doi.org/10.1016/s0022-2836(83)80111-9 PMID: 6410076
11. Giese KC, Michalowski CB, Little JW. RecA-Dependent Cleavage of LexA Dimers. J Mol Biol. 2008; 377:148–161. https://doi.org/10.1016/j.jmb.2007.12.025 PMID: 18234215
12. Kovačič L, Paulič N, Leonardi A, Hodnik V, Anderluh G, Podlesek Z, et al. Structural insight into LexA-RecA* interaction. Nucleic Acids Res. 2013; 41:9901–9910. https://doi.org/10.1093/nar/gkt744 PMID: 23965307

13. Little JW, Edmiston SH, Pacelli LZ, Mount DW. Cleavage of the *Escherichia coli* lexA protein by the recA protease. Proc Natl Acad Sci U S A. 1980; 77:3225–3229.

14. Simmons LA, Foti JJ, Cohen SE, Walker GC. The SOS Regulatory Network. EcoSal Plus. 2008; 3: https://doi.org/10.1128/ecosalplus.5.4.3 PMID: 25325076

15. Jaramillo-Riveri S, Broughton J, McVey A, Pilizota T, Scott M, El Karoui M. Growth-dependent heterogeneity in the DNA damage response in *Escherichia coli*. Mol Syst Biol. 2022; 18:e10441.

16. Jones EC, Uphoff S. Single-molecule imaging of LexA degradation in *Escherichia coli* elucidates regulatory mechanisms and heterogeneity of the SOS response. Nat Microbiol. 2021; 6:981–990. https://doi.org/10.1038/s41564-021-00930-y PMID: 34183814

17. Fuchs RP, Fujii S. Translesion DNA synthesis and mutagenesis in prokaryotes. Cold Spring Harb Perspect Biol. 2013; 5:a012682. https://doi.org/10.1101/cshperspect.a012682 PMID: 24296168

18. Kamat A, Badrinarayanan A. SOS-independent bacterial DNA damage responses: diverse mechanisms, unifying function. Curr Opin Microbiol. 2023; 73:102323.

19. Olivencia BF, Müller AU, Roschitzki B, Burger S, Weber-Ban E, Imkamp F. *Mycobacterium smegmatis* PafBC is involved in regulation of DNA damage response. Sci Rep. 2017; 7:13987.

20. Müller AU, Imkamp F, Weber-Ban E. The Mycobacterial LexA/RecA-Independent DNA Damage Response Is Controlled by PafBC and the Pup-Proteasome System. Cell Rep. 2018; 23:3551–3564. https://doi.org/10.1016/j.celrep.2018.05.073 PMID: 29924998

21. Müller AU, Kummer E, Schilling CM, Ban N, Weber-Ban E. Transcriptional control of mycobacterial DNA damage response by sigma adaptation. Sci Adv. 2021; 7:eabl4064. https://doi.org/10.1126/sciadv.abl4064 PMID: 34851662

22. Gozzi K, Salinas R, Nguyen VD, Laub MT. Schumacher MA. ssDNA is an allosteric regulator of the *C. crescentus* SOS-independent DNA damage response transcription activator, DriD. Genes Dev. 2022; 36:618–633.

23. Modell JW, Kambara TK, Perchuk BS, Laub MT. A DNA Damage-Induced, SOS-Independent Checkpoint Regulates Cell Division in *Caulobacter crescentus*. PLoS Biol. 2014; 12:e1001977.

24. Gozzi K, Tran NT, Modell JW, Le TBK, Laub MT. Prophage-like gene transfer agents promote *Caulobacter crescentus* survival and DNA repair during stationary phase. PLoS Biol. 2022; 20:e3001790.

25. Keller LM, Weber-Ban E. An emerging class of nucleic acid-sensing regulators in bacteria: WYL domain-containing proteins. Curr Opin Microbiol. 2023; 74:102296. https://doi.org/10.1016/j.mib.2023.102296 PMID: 37027901

26. Adefisayo OO, Dupuy P, Nautiyal A, Bean JM, Glickman MS. Division of labor between SOS and PafBC in mycobacterial DNA repair and mutagenesis. Nucleic Acids Res. 2021; 49:12805–12819. https://doi.org/10.1093/nar/gkab1169 PMID: 34871411

27. Samson L, Cairns J. A new pathway for DNA repair in *Escherichia coli*. Nature. 1977; 267:281–283.

28. Mielecki D, Grzesiuk E. Ada response—a strategy for repair of alkylated DNA in bacteria. FEMS Microbiol Lett. 2014; 355:1–11. https://doi.org/10.1111/1574-6968.12462 PMID: 24810496

29. Uphoff S. Real-time dynamics of mutagenesis reveal the chronology of DNA repair and damage tolerance responses in single cells. Proc Natl Acad Sci U S A. 2018; 115:E6516–E6525. https://doi.org/10.1073/pnas.1801101115 PMID: 29941584

30. Jeggo P, Defais M, Samson L, Schendel P. An adaptive response of *E. coli* to low levels of alkylating agent: Comparison with previously characterised DNA repair pathways. Molec Gen Genet. 1977; 157:1–9.

31. Uphoff S, Lord ND, Okumus B, Potvin-Trottier L, Sherratt DJ, Paulsson J. Stochastic activation of a DNA damage response causes cell-to-cell mutation rate variation. Science. 2016; 351:1094–1097. https://doi.org/10.1126/science.aac9786 PMID: 26941321

32. Friedberg EC, Walker GC, Siede W, Wood RD, Schultz RA, Ellenberger T. DNA Repair and Mutagenesis. ASM Press; 2005.

33. Nakabeppu Y, Sekiguchi M. Regulatory mechanisms for induction of synthesis of repair enzymes in response to alkylating agents: ada protein acts as a transcriptional regulator. Proc Natl Acad Sci U S A. 1986; 83:6297–6301. https://doi.org/10.1073/pnas.83.17.6297 PMID: 3529081

34. Sakumi K, Igarashi K, Sekiguchi M, Ishihama A. The Ada protein is a class I transcription factor of *Escherichia coli*. J Bacteriol. 1993; 175:2455–2457.

35. He C, Hus J-C, Sun LJ, Zhou P, Norman DPG, Dötsch V, et al. A Methylation-Dependent Electrostatic Switch Controls DNA Repair and Transcriptional Activation by *E. coli* Ada. Mol Cell. 2005; 20:117–129.

36. Abdollahi M, Hosseini A. Streptozotocin. In *Encyclopedia of Toxicology* (Third Edition). Wexler P, editor. Academic Press; 2014. p. 402–404.

37. Ng TL, Rohac R, Mitchell A, Boal AK, Balskus EP. An N-nitrosating metalloenzyme constructs the pharmacophore of streptozotocin. Nature. 2019; 566:94. https://doi.org/10.1038/s41586-019-0894-z PMID: 30728519

38. Poncin K, Roba A, Jimmidi R, Potemberg G, Fioravanti A, Francis N, et al. Occurrence and repair of alkylating stress in the intracellular pathogen *Brucella abortus*. Nat Commun. 2019; 10:4847. https://doi.org/10.1038/s41467-019-12516-8 PMID: 31649248

39. Rošić S, Amouroux R, Requena CE, Gomes A, Emperle M, Beltran T, et al. Evolutionary analysis indicates that DNA alkylation damage is a byproduct of cytosine DNA methyltransferase activity. Nat Genet. 2018; 50:452–459. https://doi.org/10.1038/s41588-018-0061-8 PMID: 29459678

40. Witkin EM. The radiation sensitivity of *Escherichia coli* B: a hypothesis relating filament formation and prophage induction. Proc Natl Acad Sci U S A. 1967; 57:1275–1279.

41. Mielecki D, Wrzesiński M, Grzesiuk E. Inducible repair of alkylated DNA in microorganisms. Mutat Res/Rev Mutat Res. 2015; 763:294–305. https://doi.org/10.1016/j.mrrev.2014.12.001 PMID: 25795127

42. Morohoshi F, Hayashi K, Munakata N. *Bacillus subtilis ada* operon encodes two DNA alkyltransferases. Nucleic Acids Res. 1990; 18:5473–5480.

43. Mielecki D, Saumaa S, Wrzesiński M, Maciejewska AM, Żuchniewicz K, Sikora A, et al. *Pseudomonas putida* AlkA and AlkB Proteins Comprise Different Defense Systems for the Repair of Alkylation Damage to DNA–In Vivo, In Vitro, and In Silico Studies. PLoS ONE. 2013; 8:e76198. https://doi.org/10.1371/journal.pone.0076198 PMID: 24098441

44. Colombi D, Gomes SL. An *alkB* gene homolog is differentially transcribed during the *Caulobacter crescentus* cell cycle. J Bacteriol. 1997; 179:3139–3145.

45. Sedgwick B, Vaughan P. Widespread adaptive response against environmental methylating agents in microorganisms. Mutat Res. 1991; 250:211–221. https://doi.org/10.1016/0027-5107(91)90178-q PMID: 1944338

46. Beranek DT. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. Mutat Res. 1990; 231:11–30. https://doi.org/10.1016/0027-5107(90)90173-2 PMID: 2195323

47. Bargonetti J, Champeil E, Tomasz M. Differential Toxicity of DNA Adducts of Mitomycin C. J Nucleic Acids. 2010; 2010:698960. https://doi.org/10.4061/2010/698960 PMID: 20798760

48. Pohlhaus JR, Kreuzer KN. Norfloxacin-induced DNA gyrase cleavage complexes block *Escherichia coli* replication forks, causing double-stranded breaks in vivo. Mol Microbiol. 2005; 56:1416–1429.

49. Modell JW, Hopkins AC, Laub MT. A DNA damage checkpoint in *Caulobacter crescentus* inhibits cell division through a direct interaction with FtsW. Genes Dev. 2011; 25:1328–1343.

50. Joseph AM, Nahar K, Daw S, Hasan MM, Lo R, Le TBK, et al. Mechanistic insight into the repair of C8-linked pyrrolobenzodiazepine monomer-mediated DNA damage. RSC Med Chem. 2022; 13:1621–1633. https://doi.org/10.1039/d2md00194b PMID: 36561066

51. Lagage V, Chen V, Uphoff S. Adaptation delay causes a burst of mutations in bacteria responding to oxidative stress. EMBO Rep. 2023; 24:e55640. https://doi.org/10.15252/embr.202255640 PMID: 36397732

52. Nakabeppu Y, Kondo H, Kawabata S, Iwanaga S, Sekiguchi M. Purification and structure of the intact Ada regulatory protein of *Escherichia coli K12*, O6-methylguanine-DNA methyltransferase. J Biol Chem. 1985; 260:7281–7288.

53. Singer B, Grunberger D. Molecular Biology of Mutagens and Carcinogens. Springer Science & Business Media; 2012.

54. Lemotte PK, Walker GC. Induction and autoregulation of *ada*, a positively acting element regulating the response of *Escherichia coli K-12* to methylating agents. J Bacteriol. 1985; 161:888–895.

55. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011; 7:e1002195. https://doi.org/10.1371/journal.pcbi.1002195 PMID: 22039361

56. </>Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596:583–589.

57. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022; 19:679–682. https://doi.org/10.1038/s41592-022-01488-1 PMID: 35637307

58. Badrinarayanan A, Le TBK, Spille J-H, Cisse II, Laub MT. Global analysis of double-strand break processing reveals in vivo properties of the helicase-nuclease complex AddAB. PLoS Genetics. 2017; 13:e1006783. https://doi.org/10.1371/journal.pgen.1006783 PMID: 28489851

59. McCarthy TV, Lindahl T. Methyl phosphotriesters in alkylated DNA are repaired by the Ada regulatory protein of *E. coli*. Nucleic Acids Res. 1985; 13:2683–2698.

60. Olsson M, Lindahl T. Repair of alkylated DNA in *Escherichia coli*. Methyl group transfer from O6-methylguanine to a protein cysteine residue. J Biol Chem. 1980; 255:10569–10571.

61. Takano K, Nakabeppu Y, Sekiguchi M. Functional sites of the Ada regulatory protein of *Escherichia coli*: Analysis by amino acid substitutions. J Mol Biol. 1988; 201:261–271.

62. Iyer LM, Koonin EV, Aravind L. Extensive domain shuffling in transcription regulators of DNA viruses and implications for the origin of fungal APSES transcription factors. Genome Biol. 2002; 3: research0012.1. https://doi.org/10.1186/gb-2002-3-3-research0012 PMID: 11897024

63. Kupferschmied P, Péchy-Tarr M, Imperiali N, Maurhofer M, Keel C. Domain Shuffling in a Sensor Protein Contributed to the Evolution of Insect Pathogenicity in Plant-Beneficial Pseudomonas protegens. PLoS Pathog. 2014; 10:e1003964. https://doi.org/10.1371/journal.ppat.1003964 PMID: 24586167

64. Cheatle Jarvela AM, Hinman VF. Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. EvoDevo. 2015; 6:3. https://doi.org/10.1186/2041-9139-6-3 PMID: 25685316

65. Joseph AM, Badrinarayanan A. Visualizing mutagenic repair: novel insights into bacterial translesion synthesis. FEMS Microbiol Rev. 2020; 44:572–582. https://doi.org/10.1093/femsre/fuaa023 PMID: 32556198

66. Sharda M, Badrinarayanan A, Seshasayee ASN. Evolutionary and Comparative Analysis of Bacterial Nonhomologous End Joining Repair. Genome Biol Evol. 2020; 12:2450–2466. https://doi.org/10.1093/gbe/evaa223 PMID: 33078828

67. Weissman JL, Fagan WF, Johnson PLF. Linking high GC content to the repair of double strand breaks in prokaryotic genomes. PLoS Genetics. 2019; 15:e1008493. https://doi.org/10.1371/journal.pgen.1008493 PMID: 31703064

68. Wu H, Zhang Z, Hu S, Yu J. On the molecular mechanism of GC content variation among eubacterial genomes. Biol Direct. 2012; 7:2. https://doi.org/10.1186/1745-6150-7-2 PMID: 22230424

69. Chimthanawala A, Badrinarayanan A. Live-Cell Fluorescence Imaging of RecN in Caulobacter crescentus Under DNA Damage. In SMC Complexes: Methods and Protocols. Badrinarayanan A, editor. Springer; 2019. p. 239–250.

70. Paintdakhi A, Parry B, Campos M, Irnov I, Elf J, Surovtsev I, et al. Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. Mol Microbiol. 2016; 99:767–777. https://doi.org/10.1111/mmi.13264 PMID: 26538279

71. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. https://doi.org/10.1093/bioinformatics/btq033 PMID: 20110278

72. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010; 26:139–140. https://doi.org/10.1093/bioinformatics/btp616 PMID: 19910308

73. Chimthanawala A, Parmar JJ, Kumar S, Iyer KS, Rao M, Badrinarayanan A. SMC protein RecN drives RecA filament translocation and remodelling for in vivo homology search. 2022. https://doi.org/10.1101/2021.08.16.456443

74. Tran NT, Stevenson CE, Som NF, Thanapipatsiri A, Jalal ASB, Le TBK. Permissive zones for the centromere-binding protein ParB on the *Caulobacter crescentus* chromosome. Nucleic Acids Res. 2018; 46:1196–1209.

75. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020; 585:357–362. https://doi.org/10.1038/s41586-020-2649-2 PMID: 32939066

76. Bardet AF, Steinmann J, Bafna S, Knoblich JA, Zeitlinger J, Stark A. Identification of transcription factor binding sites from ChIP-seq data at high resolution. Bioinformatics. 2013; 29:2705–2713. https://doi.org/10.1093/bioinformatics/btt470 PMID: 23980024

77. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2007; 23:127–128. https://doi.org/10.1093/bioinformatics/btl529 PMID: 17050570