



## Mutations in the Global Transcription Factor CRP/CAP: Insights from Experimental Evolution and Deep Sequencing

Frendorf, Pernille Ott; Lauritsen, Ida; Sekowska, Agnieszka; Danchin, Antoine; Nørholm, Morten H.H.

*Published in:*  
Computational and Structural Biotechnology Journal

*Link to article, DOI:*  
[10.1016/j.csbj.2019.05.009](https://doi.org/10.1016/j.csbj.2019.05.009)

*Publication date:*  
2019

*Document Version*  
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

*Citation (APA):*  
Frendorf, P. O., Lauritsen, I., Sekowska, A., Danchin, A., & Nørholm, M. H. H. (2019). Mutations in the Global Transcription Factor CRP/CAP: Insights from Experimental Evolution and Deep Sequencing. *Computational and Structural Biotechnology Journal*, 17, 730-736. <https://doi.org/10.1016/j.csbj.2019.05.009>

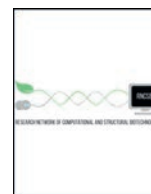
---

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



## Mini Review

# Mutations in the Global Transcription Factor CRP/CAP: Insights from Experimental Evolution and Deep Sequencing

Pernille Ott Frendorf<sup>a,1</sup>, Ida Lauritsen<sup>a,1</sup>, Agnieszka Sekowska<sup>b</sup>, Antoine Danchin<sup>c</sup>, Morten H.H. Nørholm<sup>a,\*</sup>

<sup>a</sup> Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet B220, DK-2800 Kgs. Lyngby, Denmark

<sup>b</sup> Institut de Cardiométabolisme et Nutrition, CHU Pitié-Salpêtrière, 47 boulevard de l'Hôpital, 75013 Paris, France

<sup>c</sup> Institut Cochin, INSERM U1016, CNRS UMR8104, Université Paris Descartes, 24 rue du Faubourg Saint-Jacques, 75014 Paris, France

## ARTICLE INFO

## Article history:

Received 10 April 2019

Received in revised form 22 May 2019

Accepted 26 May 2019

Available online 18 June 2019

## Keywords:

cAMP receptor protein

Carbon catabolite repression

Global transcriptional regulation

Experimental evolution

## ABSTRACT

The *Escherichia coli* cyclic AMP receptor protein (CRP or catabolite activator protein, CAP) provides a textbook example of bacterial transcriptional regulation and is one of the best studied transcription factors in biology. For almost five decades a large number of mutants, evolved *in vivo* or engineered *in vitro*, have shed light on the molecular structure and mechanism of CRP. Here, we review previous work, providing an overview of studies describing the isolation of CRP mutants. Furthermore, we present new data on deep sequencing of different bacterial populations that have evolved under selective pressure that strongly favors mutations in the *crp* locus. Our new approach identifies more than 100 new CRP mutations and paves the way for a deeper understanding of this fascinating bacterial master regulator.

© 2019 The Authors. Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Contents

1. Introduction . . . . .	730
2. Evolved and Engineered Mutations in <i>crp</i> . . . . .	732
3. Deep Sequencing of CRP in Aging Bacterial Populations . . . . .	733
Acknowledgements . . . . .	734
Appendix A. Supplementary data. . . . .	734
References. . . . .	734

## 1. Introduction

Studies of CRP date back to the earliest days of molecular biology, shortly after the model for negative *lac* gene regulation was presented [25]. In the decade following Jacob and Monod's groundbreaking discoveries, several reports on positive gene regulation were published including catabolite activation by CRP [14,15,69]. CRP is mostly known for its global regulatory role in carbon catabolism in the model bacterium *Escherichia coli* (*E. coli*): In the absence of readily metabolized carbon sources such as glucose, the enzyme adenylate cyclase is activated, producing cyclic AMP (cAMP) from ATP. cAMP binds and activates CRP, increasing the affinity for DNA, which in many cases activate operons

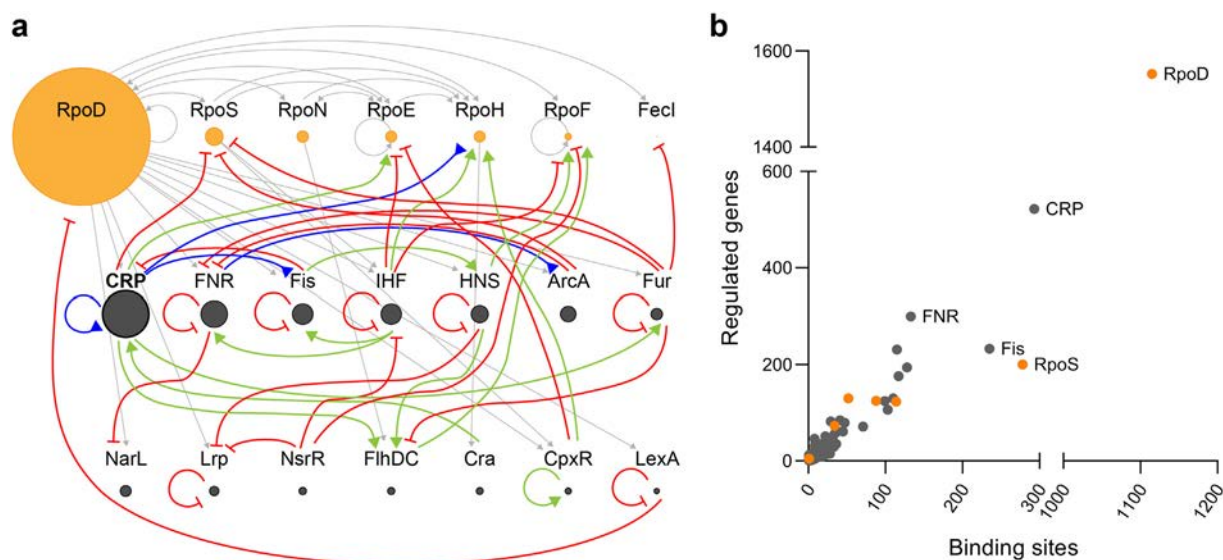
involved in the utilization of alternative carbon sources such as lactose and maltose. However, CRP can also repress gene expression and has been shown to regulate hundreds of genes in the *E. coli* genome, earning it the status of “global” or “master” regulator (Fig. 1) [31]. In fact, beyond the many specific binding sites experimentally validated in *E. coli*, CRP exhibits unspecific DNA binding affinity and together these observations points towards a role more akin to that of a nucleoid-associated protein involved in the organisation of the bacterial chromosome [55].

The first experimentally solved three-dimensional structure of CRP bound to cAMP was published in 1981 [32], but for many years the absence of an experimentally determined structure of apo-CRP hindered the understanding of the conformational changes that occur upon cAMP binding [18,27]. Presently, only one crystal structure [62] and one NMR structure [39] of wildtype apoCRP have been published. Both structures indicate that large structural rearrangements take place for DNA binding to occur. These include reorientation of the

\* Corresponding author.

E-mail address: [mormo@biosustain.dtu.dk](mailto:mormo@biosustain.dtu.dk) (M.H.H. Nørholm).

<sup>1</sup> These authors contributed equally.



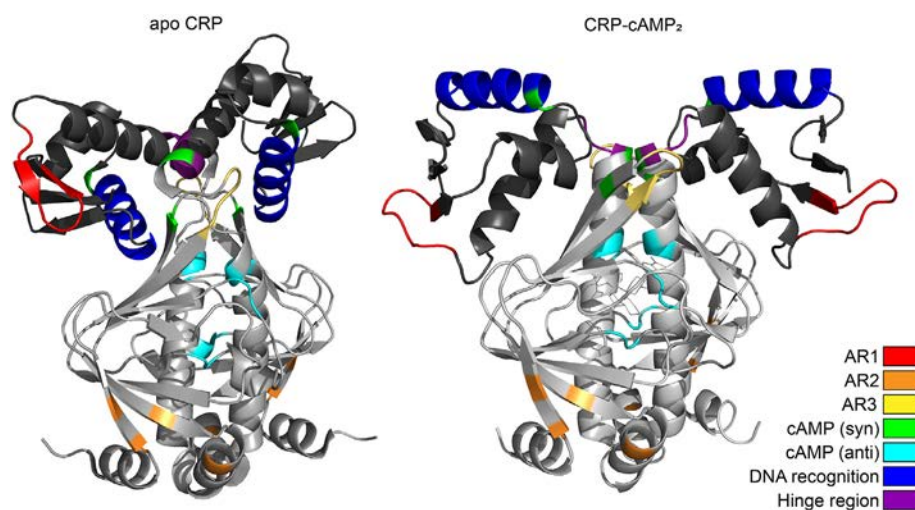
**Fig. 1.** Network of global transcription factors and sigma factors in *Escherichia coli*. Illustrated is a) The sigma factors (orange circles) and the 14 biggest regulators by regulon size (grey circles). The size of the circles is directly proportional to the sigulon or regulon size by number of genes directly affected (RegulonDB 04-03-2019) [19]. Arrows designate regulation of regulator expression; sigma factor transcription (grey arrows), activation (green, arrowheads), repression (red, perpendicular line ends) or dual regulation (blue, reverse arrowheads) [26]. b) The sigma factors (orange circles) and 207 regulators (grey circles) of *E. coli* plotted by the number of promoters or binding sites recognized, respectively, and the number of genes directly affected (RegulonDB 04-03-2019) [19].

DNA-binding domain and stabilization of the backbone helix, but the two different structures do not agree on the orientation of the C-terminal domain. This observation, and the limited number of published apoCRP structures, suggests that apoCRP may be unstable due to flexibility of the C-terminal domain [44].

CRP is a 45 kDa homodimer (Fig. 2), with each monomer consisting of 209 amino acids in two separate domains. The larger N-terminal domain (residues 1–138) binds the allosteric effector cAMP in the *anti*-conformation (residues 71, 72, 82, 83, 127 and 128) with reported binding constants in the range of 1–28  $\mu\text{M}$  [2,13,14,22,30,56]. The C-terminal domain (residues 139–209) houses the DNA-recognition helix (residues 181–193) as part of a helix-turn-helix (HTH) motif [6]. The N- and C-terminal domains are connected by a hinge region (residues 135–138). In the absence of ligand, CRP exists in a closed conformation (Fig. 2, left), where the HTH motif is secluded inside the C-terminal domain. When cAMP binds the main binding site, allosteric change stabilizes an open complex (Fig. 2,

right), resulting in the HTH motif protruding from the surface of the protein, thereby enabling DNA binding. A secondary effector binding site binds cAMP in its *syn*-conformation (residues 58, 135, 180) with binding constants in the millimolar range likely to be of limited physiological relevance [30,38]. In the presence of excess cAMP, both cAMP binding sites in one monomer are occupied, resulting in a lower DNA-binding affinity [38].

When CRP binds, the DNA is bent 90°, which likely significantly affects protein-protein and protein-DNA interactions in promoter regions [17,28,47,61] and several studies suggest that CRP is involved both in recruiting the RNA polymerase and in post-recruitment regulation [7,29,35,41]. In the case of transcriptional activation, regulation is mediated by direct interactions with the RNA polymerase holoenzyme through so-called activating regions (ARs). AR1 (residues 156–164) interacts with the  $\alpha\text{CTD}$  domain [36,46,63–66], AR2 (residues 19, 21, 96 and 101) with the  $\alpha\text{NTD}$  domain [35,58] and AR3 (residues 52–55 and 58) with the sigma factor [4,41,42,58–60].



**Fig. 2.** Illustration of functional domains in a closed and an open conformation of CRP. The locations of functionally important CRP domains in the tertiary structure of apo-CRP (left) and CRP-cAMP<sub>2</sub> (right). The protein structures are from Protein Data Bank entries 3FWE (apo-CRP) and 1ZRC (CRP-cAMP<sub>2</sub>) and were modified using the PyMOL Molecular Graphics System, Version 2.3 Schrödinger, LLC.

Complementary to structural and biochemical studies, evolved and rationally engineered mutations are imperative when elucidating protein structure and function. In this respect, CRP is probably one of the best studied transcription factors. For several decades, mutant generation has been applied to study the physiological role and molecular mechanism of CRP [37,45]. A comprehensive list of mutations identified in CRP was previously assembled [37], but to our knowledge the different approaches towards obtaining these CRP mutants have not been reviewed previously. In light of recent high throughput sequencing (HTS) approaches applied to *crp* mutants [48], we here aim at presenting a short overview on the more than 40 years of exploration of the CRP mutational space and in addition provide complementary new data on deep sequencing of experimentally evolved *crp*, obtained in aging bacterial colonies. An overview of all CRP mutants reviewed here is shown in Tables S1 and S2.

## 2. Evolved and Engineered Mutations in *crp*

The outcome of an evolution experiment is dependent on the mutational space available under the given experimental conditions, the genotype of the organism and the selection pressure applied. Evolution of CRP has been pursued by inducing mutagenesis using e.g. UV radiation or chemical reagents [33,45] or by directly targeting the *crp* gene with error-prone PCR [67], but several studies have also relied on spontaneously arising mutants [9,43,48,50]. In most studies, the CRP variants were generated in a cAMP-deficient production strain ( $\Delta$ *cya*) and screened for mutations enabling fermentation of a carbon source such as lactose or maltose. Such cAMP suppressor mutations were termed *csm* [33] and CRP variants called CRP\* [45] or CRP<sup>i</sup> [5].

In the 1970's, early after the discovery of CRP, researchers began isolating *csm*, CRP\* [5,12,45,51,54] and defective *crp*-mutants [3] mostly under conditions of induced mutagenesis. However, the exact molecular nature of these mutants was unknown for some time due to the lack of DNA sequencing and amplification technologies.

In 1985, Aiba and colleagues published a paper where they had exposed a plasmid-borne *crp* gene to UV-radiation and selected for lactose utilization in a  $\Delta$ *cya* strain. The obtained *crp* mutations in the isolated strains caused amino acid substitutions in positions 53, 62, 141, 142 and 148 of the CRP protein [1]. It was noted that positions 53 and 62 were in vicinity of the cAMP binding site, but that the phenotypes differed in that only CRP D53H was activated by the alternative cyclic nucleotide cGMP. The role of position 53 being located in the AR3 region, possible interacting with sigma factors, was not discussed in this work as AR3 was unknown at this time. Amino acids 141, 142 and 148 are part of the D- $\alpha$ -helix and it was speculated that they were critical in the allosteric transition, from the N-terminal domain to the DNA binding C-terminal domain, normally caused by binding of cAMP.

Around the same time, Garges and Adhya used *crp*-carrying phages for infection and growth in a mutator *E. coli* strain for CRP mutant generation [21]. Phages carrying mutagenized *crp* variants were isolated as positive lactose utilizing plaques in an  $\Delta$ *cya* background. The detected mutations were in positions 72, engaged in cAMP binding, and again in D- $\alpha$ -helix residues 141, 142 and 144.

The following year, Harman and colleagues sequenced and characterized three CRP\* mutants that previously were selected by different methods [23,33,40,45]. The CRP mutation A144T in the D- $\alpha$ -helix was again identified – in this case from a  $\Delta$ *cya* strain selected on xylose as carbon source. A T127I mutation in the cAMP binding site was identified in combination with Q170K from a strain that complemented a CRP binding site mutation designated L8 in the *lac* promoter [45]. This double mutant showed a CRP\* phenotype and was activated by cGMP, although the physiological relevance of the latter was questioned by Harman and colleagues. The individual effects of the 127 and 170 mutations were not explored further in this work. Finally, the mutation L195R was evolved in the  $\Delta$ *cya*, *crp* T127I, Q170K mutant background

and the extra mutation enabled growth on arabinose in the absence of cAMP [40]. The authors suggested that the increased positive charge of this L195R mutation in the DNA binding domain caused an increase in the affinity for DNA.

In a follow-up study by Garges and Adhya, CRP\* suppressor mutants, causing a loss of the G141S and A144T CRP\* phenotypes on lactose, were identified as T127A and R169C/E171G, respectively [20]. In case of the G141S CRP\* mutant, it is perhaps not surprising that a mutation near the cAMP binding (T127A) can neutralize cAMP independence. Similarly, it was noted that amino acids near the two mutations in positions 169 and 171, identified in the CRP A144T mutant background, were previously suggested to interact with the amino acid Y63 near the cAMP binding site, but also could be in direct interaction with the DNA [57].

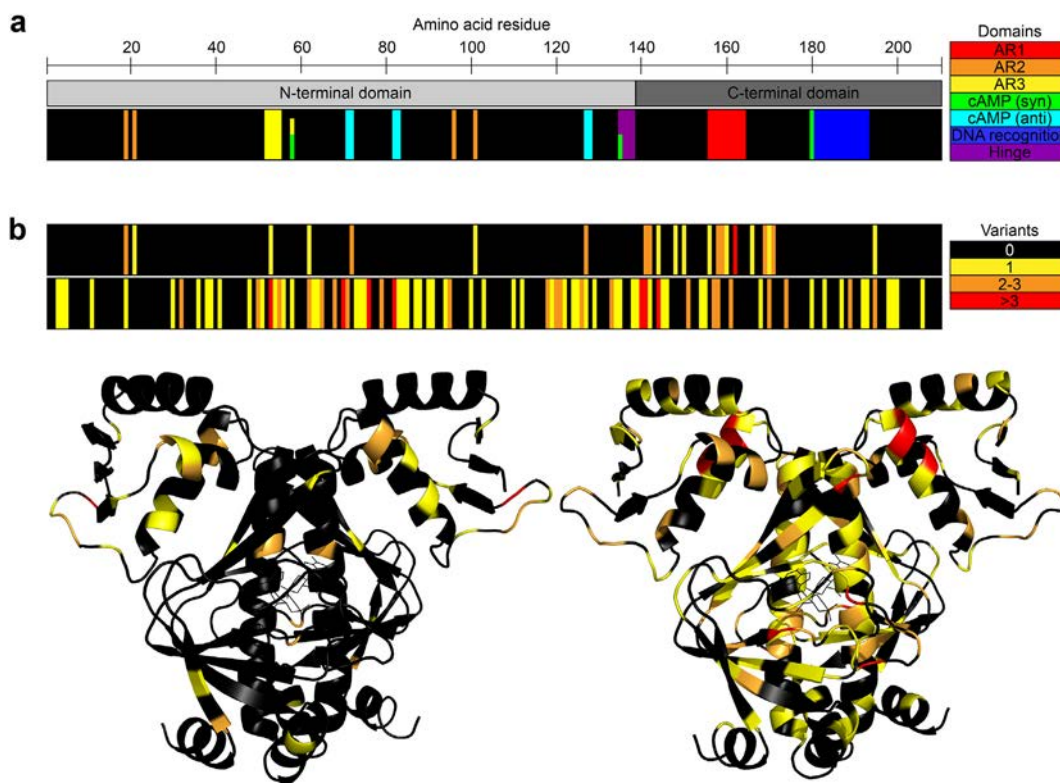
Mutations in positions 141, 141 and 144 again occurred in a study that described the selection of CRP\* mutants based on growth on lactose in a  $\Delta$ *cya* mutant background [53]. In this case, two different mutations in position 144, A144T and A144E, were found in combination with T28K. These two combinations were found to be toxic when expressed on a multicopy plasmid, whereas two other D-helix mutations T140K and G141D were tolerated in high copy.

Another broad category of CRP mutations, more generally termed *positive control mutants* (*pcm*), were selected for their inability to induce transcription while retaining binding to specific CRP DNA binding sites. The first attempt at creating CRP *pc* mutants introduced the mutations E171Q, E171K and Q170K based on similarity to the lambda repressor, but these caused different effects at different promoters [4,24]. A more clear *pcm* phenotype was observed with the mutation CRP H159L in AR1 and second site revertants was identified as K52N and K52Q in AR3 [4]. Eschenlauer and Reznikoff screened for CRP mutants that repressed the *gal* promoter but had lost their ability to induce the *lac* promoter [16]. This way, they identified mutations in cAMP binding position 72, and in position 162 in the AR1 region. In a similar study by Zhou and co-workers, plasmid-harbored *crp* genes were mutagenized by error-prone PCR and screened in an engineered  $\Delta$ *cya* strain for defective ribose fermentation, while retaining the ability to repress a modified *lac* promoter [66]. The identified mutations were in positions 156, 158, 159, and 162 that are all part of AR1. Finally, Niu and co-workers identified mutations in the AR1 region that could not activate transcription of Class I and II CRP-dependent promoters as well as the mutations H19L, H19Y, H21L and K101E in the AR2 region that were only defective in Class II promoter activation [35].

Experimental evolution of CRP has also addressed the complex interplay between CRP and the CytR transcription factor. A CytR-repressed *tsx* promoter construct was screened in combination with a mutagenized *crp* plasmid library for CRP variants that were dominantly activating the promoter. This approach identified mutations in positions 17, 18, 108 and 110 in CRP – all in the vicinity of AR2 [52].

Recent work has again explored CRP mutants that evolve spontaneously under different selection regimes. Sievert and co-workers observed that by growing an *E. coli* strain in high levels of the (CRP-dependent) carbon source xylose, the CRP G141D mutation again was found to evolve, promoting increased xylose utilization and growth rate [50]. In adaptive laboratory evolution for improved fitness in minimal medium supplied with lactate, the CRP mutations L150Q and I165T evolved [9]. These are both located near AR1 and presumably cause a changed interaction with the RNA polymerase.

In summary, adaptive mutations identified in the previous four decades of CRP studies (Fig. 3) occur predominantly in the cAMP binding site, the D- $\alpha$ -helix, and in the RNA polymerase activating domains AR1 and AR2. These three categories are intuitively easy to understand as they likely either directly affect ligand binding, ligand-induced allosteric transitions, or the productive interaction with the core RNA polymerase, respectively. Mutations around



**Fig. 3.** Illustration of adaptive mutations identified in CRP a) The locations of functionally important CRP domains in the primary amino acid sequence. b) Locations of mutations identified in the primary and tertiary structure of CRP in previous studies (upper black bar and structure to the left) and in our laboratories by deep sequencing (lower black bar and structure to the right). The number of variants identified per residue is colour coded (black: 0 variants, yellow: 1 variant, orange: 2–3 variants, and red: more than three variants).

position 170 are also frequently observed but are more difficult to interpret and have been discussed both to directly affect interactions with cAMP and the DNA.

To supplement *crp* mutants that evolve under selective pressure, the advent of PCR enabled hypothesis-driven site-directed mutagenesis exploring the role of specific amino acid residues in the different domains of the CRP protein. As summarized in Table S2, these are often studied in areas where mutations evolve naturally (the cAMP-*anti* binding pocket, the D- $\alpha$ -helix and the activating regions AR1 and AR2), but have also explored mutations in e.g. the hinge region that connects the C- and D-helices, AR3, and the DNA binding domain in the C-terminus.

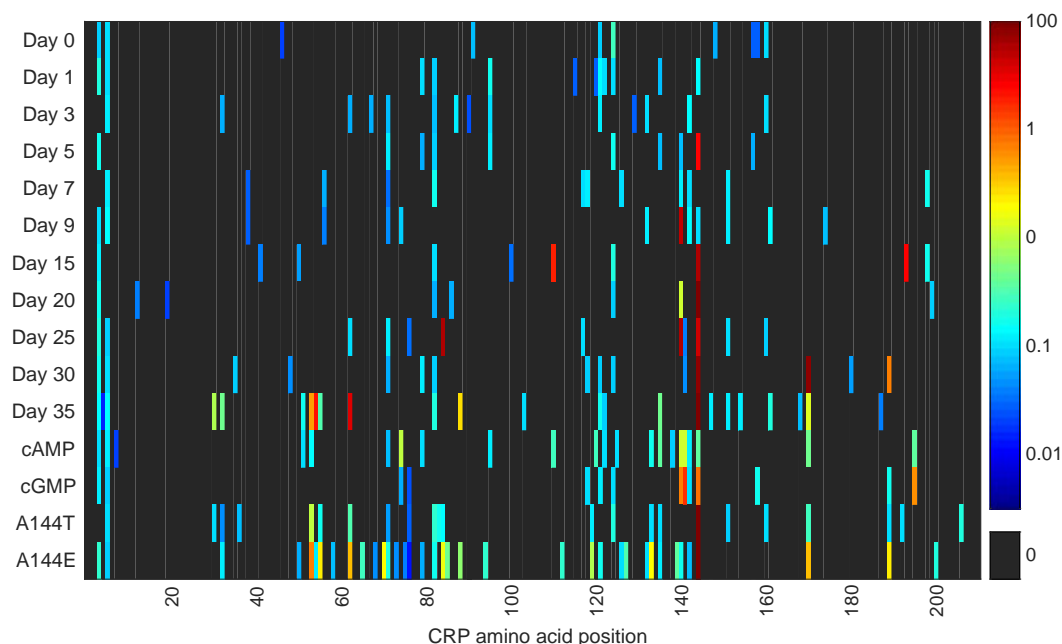
In an experiment designed to study adaptation when bacteria age and starve over two months, our recent work identified a large number of CRP mutants. By plating a  $\Delta$ *cya* *E. coli* strain on MacConkey agar supplied with the CRP-dependent carbon source maltose, mutant red “*papillae*” - or “secondary colonies” - appeared that were able to use the extra carbon source. 96 mutants were selected for genome sequencing based on their different CRP\* phenotypes and their temporal appearance. In addition, an additional approximately 500 mutants had their *crp* loci sequenced by PCR amplification and Sanger sequencing [48]. 35 different missense mutations were identified in *crp*. Seven of the identified mutations, S62F, T127I, G141D, G141S, A144T, A144E and L195R were previously observed in adaptive evolution studies and have all been classified as CRP\*. Other *crp* mutations have not been identified elsewhere, including P110Q, L134M, T140P/R/K, A144K, G162S and M189K. The CRP mutant S62Y has not been identified before. However, the substitution of serine to phenylalanine (S62F) has been observed by induced mutagenesis and screening on lactose previously [1]. The Q170K mutation was always observed in combination with an additional *crp* mutation such as T140R, A144T/E or M198K [48]. Interestingly, a similar trend was observed in the study by Harman and co-workers in which the Q170K mutation was paired with T127I or T127I and L195R [23].

### 3. Deep Sequencing of CRP in Aging Bacterial Populations

We reasoned that we could obtain deeper insights into the mutational space of CRP by growing a large number of different bacterial colonies, followed by deep sequencing of the *crp* locus. To this end, we followed the same workflow as previously described [48], but now isolated DNA from more than 500 colonies per plate at different time points, followed by PCR amplification of *crp* and HTS.

The new data presented here identifies more than 100 new *crp* missense mutations (Figs. 3–4, Table S3), although we only observed a significant increase in different mutations towards the end of the 35 days. The HTS approach provides a previously unmatched look at the mutational landscape of CRP, revealing novel insights into the evolutionary response of a strain during a selective event. However, a tradeoff for the sequencing depth is its inability to distinguish between amplification of a single mutation and multiple occurrences of the same mutation. Thus, mutations detected by this HTS approach can only be roughly categorized into those appearing in higher frequencies (likely due to a clear fitness advantage leading to a dominating population) or those that appear in low frequencies. Similar, we cannot distinguish between mutations that occur alone and those that only occur in combination with other mutations such as described above for the mutation Q170K.

The data presented in Figs. 3 and 4 is generally well aligned with previous observations. Several high-frequency mutations are observed in the residues D53, S62, T140, G141, A144, Q170, and relatively few mutations are seen in the C-terminus of CRP. In addition, DNA was isolated from aging bacteria in the presence of two CRP ligands: 0.1 mM cAMP and 0.5 mM cGMP, and with parental strains that already contained the frequently observed CRP\* mutants A144T and A144E (Fig. 4). From plates supplemented with the two cyclic nucleotides, binding site mutations again dominate together with mutations in the D- $\alpha$ -helix. Interestingly, although the A144T/E mutations observed by



**Fig. 4.** Heatmap of the mutation frequencies (%) of each CRP residue in an *E. coli*  $\Delta cyaA$  pTIG strain. Strain and evolution experiment set-up were as previously described [48]. Biomass samples of the populations were taken from preculture (day 0) or from one of 20 identical plates during the evolution experiment (days 1–35). Furthermore, samples were taken at day 7 of the  $\Delta cyaA$  strain supplemented with 0.1 mM cAMP or 0.5 mM cGMP, and of *E. coli*  $\Delta cyaA$  strains with CRP A144T or CRP A144E mutations. The *crp* gene was sequenced by amplicon next-generation sequencing and the data processed by CLC Genomics Workbench (Qiagen, Aarhus, Denmark).

Sekowska *et al.* occur in the same residue, they have distinct effects on the mutational landscape of CRP. The A144T mutation causes almost uniformly distributed low frequency mutations, while the A144E mutation causes increased mutagenesis generally concentrated around residues 50–90 (AR3 and the cAMP binding site) and 110–150 (the cAMP binding site and inter-domain stability).

To our knowledge, the data presented here is the most comprehensive overview of the natural mutational landscape of CRP to date. Our approach provides a detailed map of CRP mutations for future in-depth characterization. With the increasingly affordable deep sequencing methodologies, the approach could be generalized to study similar evolutionary tracks in transcriptional regulation. How can we try and extend our knowledge of structural and regulatory features of the CRP protein?

Besides exploring changes in physical growth conditions (altering temperature or osmolarity) coupling with other master regulatory systems should be rewarding. Global regulators such as CRP manage coordination of gene expression in a variety of conditions, that are also coordinated by other regulatory molecules, in particular the alarmone ppGpp. This “magic spot” has been discovered half a century ago as involved in monitoring amino acid availability [8]. Yet its role is far from fully understood and still a matter of considerable research. It is now known that altered levels of this regulatory molecule in *relA spoT* mutants – coding for enzymes controlling the synthesis and turnover of the molecule – resulted in non-optimal resource allocation in *E. coli* [68]. Interestingly, this happened under conditions where it is expected that CRP is involved in the management of ppGpp-mediated effects [49] and CRP-mediated contribution to *relA* expression has been demonstrated [34]. This regulation must match the coupling between the cAMP-CRP regulation and amino acid biosynthesis.

As a case in point indeed, it has long been known that there is an explicit link between ppGpp synthesis and serine/one carbon metabolism. Serine excess resulted in growth inhibition of *relA* mutants, while *relA cya*-defective or *crp*-defective mutants became resistant to excess serine. In *relA cya* strains, sensitivity to serine was restored when the growth medium was supplemented with cAMP, substantiating the serine-mediated interference in the cAMP-CRP control of gene

expression [10]. To be sure, this effect was reverted in a *crp\** background. It was therefore interesting to isolate secondary mutants that would again be resistant to excess serine in order to better understand how CRP was involved in this regulation. A new class of CRP mutants was identified in *E. coli cya relA crp\** strains. These mutants were mapped in the *crp* gene, and their physiological features differed from both the wild type *crp* and the *crp\** allele [11]. However, they could not be studied more in-depth at the time. Exploring this selection procedure with the “omics” techniques that are now familiar should allow us to enter a new evolution landscape of the protein.

Similar approaches could be developed to study other global regulators. In general, letting genes that are expressed under stationary conditions evolve should bring about new observations in the uncharted territory of adaptive mutations.

## Acknowledgements

We would like to thank the three anonymous reviewers for their positive, insightful and constructive comments. This work was supported by the Novo Nordisk Foundation.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.csbj.2019.05.009>.

## References

- [1] Aiba H, Nakamura T, Mitani H, Mori H. Mutations that alter the allosteric nature of cAMP receptor protein of *Escherichia coli*. *EMBO J* 1985;4:3329–32. <https://doi.org/10.1002/j.1460-2075.1985.tb04084.x>.
- [2] Anderson WB, Schneider AB, Emmer M, Perlman RL, Pastan IRA. Purification of and properties of the cyclic adenosine monophosphate receptor protein which mediates cyclic gene transcription in *Escherichia coli*. *J Biol Chem* 1971;246:5929–37.
- [3] Artman M, Werthamer S. Effects of cyclic guanosine 3',5'-monophosphate on the synthesis of enzymes sensitive to catabolite repression in intact cells of *Escherichia coli*. *J Bacteriol* 1974;120:980–3.
- [4] Bell A, Gaston K, Williams R, Chapman K, Kolb A, Buc H, et al. Mutations that alter the ability of the *Escherichia coli* cyclic AMP receptor protein to activate transcription. *Nucleic Acids Res* 1990;18:7243–50. <https://doi.org/10.1093/nar/18.24.7243>.

- [5] Botsford JL, Drexler M. The cyclic 3',5'-adenosine monophosphate receptor protein and regulation of cyclic 3',5'-adenosine monophosphate synthesis in *Escherichia coli*. MGG Mol Gen Genet 1978;165:47–56. <https://doi.org/10.1007/BF00270375>.
- [6] Brennan RG. Interactions of the helix-turn-helix binding domain. Curr Opin Struct Biol 1991. [https://doi.org/10.1016/0959-440x\(91\)90015-1](https://doi.org/10.1016/0959-440x(91)90015-1).
- [7] Busby S, Ebright RH. Transcription activation by catabolite activator protein (CAP). J Mol Biol 1999;293:199–213. <https://doi.org/10.1006/jmbi.1999.3161>.
- [8] Cashel M, Gallant J. Two compounds implicated in the function of RC gene. Nature 1969;221:838–41.
- [9] Conrad TM, Joyce AR, Applebee MK, Barrett CL, Xie B, Gao Y, et al. Whole-genome resequencing of *Escherichia coli* K-12 MG1655 undergoing short-term laboratory evolution in lactate minimal media reveals flexible selection of adaptive mutations. Genome Biol 2009;10:1–12. <https://doi.org/10.1186/gb-2009-10-10-r118>.
- [10] Daniel J, Danchin A. Involvement of cyclic AMP and its receptor protein in the sensitivity of *Escherichia coli* K 12 toward serine. Mol Gen Genet MGG 1979;176:343–50. <https://doi.org/10.1007/BF00333096>.
- [11] Daniel J, Joseph E, Ullmann A, Danchin A. CrpX mutants of *Escherichia coli* K-12: selection and physiological properties. FEMS Microbiol Lett 1981;10:389–93. <https://doi.org/10.1111/j.1574-6968.1981.tb06277.x>.
- [12] Dessein A, Schwartz M, Ullmann A. Catabolite repression in *Escherichia coli* mutants lacking cyclic AMP. MGG Mol Gen Genet 1978;162:83–7. <https://doi.org/10.1007/BF00333853>.
- [13] Donoso-Pardo JL, Turner PC, King RW. Cyclic nucleotide binding to cAMP receptor protein from *Escherichia coli* optical and ligand-binding studies. Eur J Biochem 1987. <https://doi.org/10.1111/j.1432-1033.1987.tb13470.x>.
- [14] Emmer M, deCrombrugge B, Pastan I, Perlman R. Cyclic AMP receptor protein of *E. coli*: its role in the synthesis of inducible enzymes. Proc Natl Acad Sci 1970;66:480–7. <https://doi.org/10.1073/pnas.66.2.480>.
- [15] Englesberg E, Irr J, Power J, Lee N. Positive control of enzyme synthesis by gene C in the L-arabinose system. J Bacteriol 1965;90:946–57.
- [16] Eschenlauer AC, Reznikoff WS. *Escherichia coli* catabolite gene activator protein mutants defective in positive control of lac operon transcription. J Bacteriol 1991;173:5024–9. <https://doi.org/10.1128/jb.173.16.5024-5029.1991>.
- [17] Fried MG, Crothers DM. CAP and RNA polymerase interactions with the lac promoter: binding stoichiometry and long range effects. Nucleic Acids Res 1983. <https://doi.org/10.1093/nar/11.1.141>.
- [18] Gallagher DT, Smith N, Kim SK, Robinson H, Reddy PT. Profound asymmetry in the structure of the cAMP-free cAMP receptor protein (CRP) from *Mycobacterium tuberculosis*. J Biol Chem 2009. <https://doi.org/10.1074/jbc.C800215200>.
- [19] Gama-Castro S, Salgado H, Santos-Zavaleta A, Ledezma-Tejeda M, Muñiz-Rascado L, García-Sotelo JS, et al. RegulonDB version 9.0: high-level integration of gene regulation, coexpression, motif clustering and beyond. Nucleic Acids Res 2016. <https://doi.org/10.1093/nar/gkv1156>.
- [20] Garges S, Adhya S. Cyclic AMP-induced conformational change of cyclic AMP receptor protein (CRP): intragenic suppressors of cyclic AMP-independent CRP mutations. J Bacteriol 1988;170:1417–22. <https://doi.org/10.1128/jb.170.4.1417-1422.1988>.
- [21] Garges S, Adhya S. Sites of allosteric shift in the structure of the cyclic AMP receptor protein. Cell 1985;41:745–51. [https://doi.org/10.1016/S0092-8674\(85\)80055-6](https://doi.org/10.1016/S0092-8674(85)80055-6).
- [22] Green J, Stapleton MR, Smith LJ, Artymiuk PJ, Kahramanoglu C, Hunt DM, et al. Cyclic-AMP and bacterial cyclic-AMP receptor proteins revisited: adaptation for different ecological niches. Curr Opin Microbiol 2014;18:1–7. <https://doi.org/10.1016/j.mib.2014.01.003>.
- [23] Harman JG, McKenney K, Peterkofsky A. Structure-function analysis of three cAMP-independent forms of the cAMP receptor protein. J Biol Chem 1986;261:16332–9 [doi:3023348].
- [24] Irwin N, Ptashne M. Mutants of the catabolite activator protein of *Escherichia coli* that are specifically deficient in the gene-activation function. Proc Natl Acad Sci 1987;84:8315–9. <https://doi.org/10.1073/pnas.84.23.8315>.
- [25] Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 1961;3:318–56. [https://doi.org/10.1016/S0022-2836\(61\)80072-7](https://doi.org/10.1016/S0022-2836(61)80072-7).
- [26] Keseler IM, Mackie A, Santos-Zavaleta A, Billington R, Bonavides-Martínez C, Caspi R, et al. The EcoCyc database: reflecting new knowledge about *Escherichia coli* K-12. Nucleic Acids Res 2017. <https://doi.org/10.1093/nar/gkw1003>.
- [27] Kumar P, Joshi DC, Akif M, Akhter Y, Hasnain SE, Mande SC. Mapping conformational transitions in cyclic AMP receptor protein: crystal structure and normal-mode analysis of *Mycobacterium tuberculosis* apo-cAMP receptor protein. Biophys J 2010. <https://doi.org/10.1016/j.bpj.2009.10.016>.
- [28] Liu-Johnson HN, Gartenberg MR, Crothers DM. The DNA binding domain and bending angle of *E. coli* CAP protein. Cell 1986;47:995–1005. [https://doi.org/10.1016/0092-8674\(86\)90814-7](https://doi.org/10.1016/0092-8674(86)90814-7).
- [29] Liu B, Hong C, Huang RK, Yu Z. Structural basis of bacterial transcription activation. Science 2017;951:947–51 (80- ).
- [30] Malecki J, Polit A, Wasylewski Z. Kinetic studies of cAMP-induced allosteric changes in cyclic AMP receptor protein from *Escherichia coli*. J Biol Chem 2000;275:8480–6. <https://doi.org/10.1074/jbc.275.12.8480>.
- [31] Martínez-Antonio A, Collado-Vides J. Identifying global regulators in transcriptional regulatory networks in bacteria. Curr Opin Microbiol 2003;6:482–9. <https://doi.org/10.1016/j.mib.2003.09.002>.
- [32] McKay DB, Steitz TA. Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. Nature 1981;290:744–9. <https://doi.org/10.1038/290744a0>.
- [33] Melton T, Snow LL, Freitag CS, Dobrogosz WJ. Isolation and characterization of cAMP suppressor mutants of *Escherichia coli* K12. Mol Gen Genet 1981;182:480–9.
- [34] Nakagawa A, Oshima T, Mori H. Identification and characterization of a second, inducible promoter of *relA* in *Escherichia coli*. Genes Genet Syst 2006;81:299–310. <https://doi.org/10.1266/ggs.81.299>.
- [35] Niu W, Kim Y, Tau G, Heyduk T, Ebright RH. Transcription activation at class II CAP-dependent promoters: two interactions between CAP and RNA polymerase. Cell 1996;87:1123–34. [https://doi.org/10.1016/S0092-8674\(00\)81806-1](https://doi.org/10.1016/S0092-8674(00)81806-1).
- [36] Niu W, Zhou Y, Dong Q, Ebright YW, Ebright RH. Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP) I. saturation and alanine-scanning mutagenesis. J Mol Biol 1994. [https://doi.org/10.1016/0022-2836\(94\)90034-5](https://doi.org/10.1016/0022-2836(94)90034-5).
- [37] Passner JM, Schultz SC, Steitz TA. Modeling the cAMP-induced allosteric transition using the crystal structure of CAP-cAMP at 2.1 Å resolution. J Mol Biol 2000;304:847–59. <https://doi.org/10.1006/jmbi.2000.4231>.
- [38] Passner JM, Steitz TA. The structure of a CAP-DNA complex having two cAMP molecules bound to each monomer. Proc Natl Acad Sci U S A 1997;94:2843–7. <https://doi.org/10.1073/pnas.94.7.2843>.
- [39] Popovych N, Tzeng S-R, Tonelli M, Ebright RH, Kalodimos CG. Structural basis for cAMP-mediated allosteric control of the catabolite activator protein. Proc Natl Acad Sci 2009. <https://doi.org/10.1073/pnas.0900595106>.
- [40] Puskas R, Fredd N, Gazdar C, Peterkofsky A. Methylglyoxal-mediated growth inhibition in an *Escherichia coli* cAMP receptor protein mutant. Arch Biochem Biophys 1983;223:503–13.
- [41] Rhodius VA, Busby SJW. Interactions between activating region 3 of the *Escherichia coli* cyclic AMP receptor protein and region 4 of the RNA polymerase  $\sigma 70$  subunit: application of suppression genetics. J Mol Biol 2000;311–24. <https://doi.org/10.1006/jmbi.2000.3737>.
- [42] Rhodius VA, Busby SJW. Transcription activation by the *Escherichia coli* cyclic AMP receptor protein: determinants within activating region 3. J Mol Biol 2000;299:295–310. <https://doi.org/10.1006/jmbi.2000.3736>.
- [43] Sabourin D, Beckwith J. Deletion of the *Escherichia coli* crp gene. J Bacteriol 1975;122:338–40.
- [44] Saha A, Mukhopadhyay J, Datta AB, Parrack P. Revisiting the mechanism of activation of cyclic AMP receptor protein (CRP) by cAMP in *Escherichia coli*: lessons from a subunit-crosslinked form of CRP. FEBS Lett 2015. <https://doi.org/10.1016/j.febslet.2014.12.021>.
- [45] Sanders R, McGeoch D. A mutant transcription factor that is activated by 3':5'-cyclic guanosine monophosphate. Proc Natl Acad Sci 1973;70:1017–21.
- [46] Savery NJ, Lloyd GS, Busby SJW, Thomas MS, Ebright RH, Gourse RL. Determinants of the C-terminal domain of the *Escherichia coli* RNA polymerase  $\alpha$  subunit important for transcription at class I cyclic AMP receptor protein-dependent promoters. J Bacteriol 2002. <https://doi.org/10.1128/JB.184.8.2273-2280.2002>.
- [47] Schultz SC, Shields GC, Steitz TA. Crystal-structure of a cap-DNA complex - the DNA is bent by 90-degrees. Science 1991;253:1001–7 (80- ).
- [48] Sekowska A, Wendel S, Fischer EC, Nørholm MHH, Danchin A. Generation of mutation hotspots in ageing bacterial colonies. Sci Rep 2016;6:2. <https://doi.org/10.1038/s41598-016-0005-4>.
- [49] Shimizu K. Metabolic regulation and coordination of the metabolism in bacteria in response to a variety of growth conditions. In: Ye Q, Bao J, Zhong J-J, editors. Bioreactor engineering research and industrial applications I: cell factories. Berlin, Heidelberg: Springer Berlin Heidelberg; 2016. p. 1–54. [https://doi.org/10.1007/10\\_2015\\_320](https://doi.org/10.1007/10_2015_320).
- [50] Sievert C, Nieves LM, Panyon LA, Loeffler T, Morris C, Cartwright RA, et al. Experimental evolution reveals an effective avenue to release catabolite repression via mutations in XylR. Proc Natl Acad Sci 2017;114:7349–54. <https://doi.org/10.1073/pnas.1700345114>.
- [51] Silverstone AE, Goman M, Scaife JG. ALT: a new factor involved in the synthesis of RNA by *Escherichia coli*. MGG Mol Gen Genet 1972;118:223–34. <https://doi.org/10.1007/BF00333459>.
- [52] Søgaard-Andersen L, Mironov AS, Pedersen H, Sukhedelets VV, Valentin-Hansen P. Single amino acid substitutions in the cAMP receptor protein specifically abolish regulation by the CytR repressor in *Escherichia coli*. Proc Natl Acad Sci 1991;88:4921–5. <https://doi.org/10.1073/pnas.88.11.4921>.
- [53] Tagami H, Inada T, Kunimura T, Aiba H. Glucose lowers CRP levels resulting in repression of the lac operon in cells lacking cAMP. Mol Microbiol 1995;17:251–8. <https://doi.org/10.1111/j.1365-2958.1995.mmi.17020251.x>.
- [54] Takebe Y, Shibuya M, Science M. A new Extragenic suppressor of cya mutation. J Biochem 1978;1623:1615–23.
- [55] Visweswariah SS, Busby SJW. Evolution of bacterial transcription factors: how proteins take on new tasks, but do not always stop doing the old ones. Trends Microbiol 2015;23:463–7. <https://doi.org/10.1016/j.tim.2015.04.009>.
- [56] Weber IT, Steitz TA. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. J Mol Biol 1987;198:311–26. [https://doi.org/10.1016/0022-2836\(87\)90315-9](https://doi.org/10.1016/0022-2836(87)90315-9).
- [57] Weber IT, Steitz TA. Model of specific complex between catabolite gene activator protein and B-DNA suggested by electrostatic complementarity. Proc Natl Acad Sci 1984;81:3973–7. <https://doi.org/10.1073/pnas.81.13.3973>.
- [58] West D, Williams R, Rhodius V, Bell A, Sharma N, Zou C, et al. Interactions between the *Escherichia coli* cyclic AMP receptor protein and RNA polymerase at class II promoters. Mol Microbiol 1993;10:789–97. <https://doi.org/10.1111/j.1365-2958.1993.tb00949.x>.
- [59] Williams R, Bell A, Sims G, Busby S. The role of two surface exposed loops in transcription activation by the *Escherichia coli* CRP and FNR proteins. Nucleic Acids Res 1991. <https://doi.org/10.1093/nar/19.24.6705>.
- [60] Williams RM, Rhodius VA, Bell AI, Kolb A, Busby SJW. Orientation of functional activating regions in the *Escherichia coli* CRP protein during transcription activation at class II promoters. Nucleic Acids Res 1996;24:1112–8. <https://doi.org/10.1093/nar/24.6.1112>.
- [61] Wu HM, Crothers DM. The locus of sequence-directed and protein-induced DNA bending. Nature 1984. <https://doi.org/10.1038/308590a0>.

- [62] Yu S, Steitz TA, Sharma H, Wang J, Kong J. Structure of apo-CAP reveals that large conformational changes are necessary for DNA binding. *Proc Natl Acad Sci U S A* 2009;106:16604–9. <https://doi.org/10.1073/pnas.0908380106>.
- [63] Zhou Y. The functional subunit of a dimeric transcription activator protein depends on promoter architecture. *Trends Genet* 1995. [https://doi.org/10.1016/0168-9525\(95\)90593-6](https://doi.org/10.1016/0168-9525(95)90593-6).
- [64] Zhou Y, Busby S, Ebright R. Identification of the functional subunit of a dimeric transcription activator protein by use of oriented heterodimers. *Cell* 1993;73:375–9. [https://doi.org/10.1016/0092-8674\(93\)90236-j](https://doi.org/10.1016/0092-8674(93)90236-j).
- [65] Zhou Y, Merkel TJ, Ebright RH. Characterization of the activating region of *Escherichia coli* catabolite gene activator protein (CAP) II. Role at class I and class II CAP-dependent promoters. *J Mol Biol* 1994. [https://doi.org/10.1016/0022-2836\(94\)90035-3](https://doi.org/10.1016/0022-2836(94)90035-3).
- [66] Zhou Y, Zhang X, Ebright RH. Identification of the activating region of catabolite gene activator protein (CAP): isolation and characterization of mutants of CAP specifically defective in transcription activation. *Proc Natl Acad Sci U S A* 1993;90:6081–5.
- [67] Zhou Y, Zhang X, Ebright RH. Random mutagenesis of gene-sized DNA molecules by use of CRP with Taq DNA polymerase. *Nucleic Acids Res* 1991;19:6052. <https://doi.org/10.1093/nar/19.21.6052>.
- [68] Zhu M, Dai X. Growth suppression by altered (p)ppGpp levels results from non-optimal resource allocation in *Escherichia coli*. *Nucleic Acids Res* 2019;47:4684–93. <https://doi.org/10.1093/nar/gkz211>.
- [69] Zubay G, Schwartz D, Beckwith J. Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proc Natl Acad Sci* 1970;66:104–10. <https://doi.org/10.1101/sqb.1970.035.01.056>.