Epistatic Interactions in the Arabinose *Cis*-Regulatory Element

Mato Lagator,*[1] Claudia Igler,[1] Anaísa B. Moreno,[1] Călin C. Guet,[1] and Jonathan P. Bollback[1]
[1]IST Austria, Klosterneuburg, Austria
***Corresponding author:** E-mail: mlagator@ist.ac.at.
Associate editor: Patricia Wittkopp

Abstract

Changes in gene expression are an important mode of evolution; however, the proximate mechanism of these changes is poorly understood. In particular, little is known about the effects of mutations within *cis* binding sites for transcription factors, or the nature of epistatic interactions between these mutations. Here, we tested the effects of single and double mutants in two *cis* binding sites involved in the transcriptional regulation of the *Escherichia coli araBAD* operon, a component of arabinose metabolism, using a synthetic system. This system decouples transcriptional control from any posttranslational effects on fitness, allowing a precise estimate of the effect of single and double mutations, and hence epistasis, on gene expression. We found that epistatic interactions between mutations in the *araBAD cis*-regulatory element are common, and that the predominant form of epistasis is negative. The magnitude of the interactions depended on whether the mutations are located in the same or in different operator sites. Importantly, these epistatic interactions were dependent on the presence of arabinose, a native inducer of the *araBAD* operon in vivo, with some interactions changing in sign (e.g., from negative to positive) in its presence. This study thus reveals that mutations in even relatively simple *cis*-regulatory elements interact in complex ways such that selection on the level of gene expression in one environment might perturb regulation in the other environment in an unpredictable and uncorrelated manner.

Key words: gene expression; gene regulation; epistasis; *cis*-regulatory elements.

Introduction

Changes in the *cis* regulation of gene expression have been proposed as a major source of evolutionary innovation (King and Wilson 1975; Wittkopp and Kalay 2012). For example, across insect species there has been increasing evidence for the essential role that *cis* regulatory changes have in shaping body plan formation (Carroll 2000, 2008; Wittkopp and Kalay 2012). Changes in the regulation of gene expression can occur through mutations in the transcription factor coding sequence (transregulatory elements) and/or in *cis*-regulatory elements (CREs), which contain the transcription factor and the RNA polymerase (RNAP) binding sites (Jacob and Monod 1961). Mutations in CREs may be important targets of selection (Stern and Orgogozo 2008), as it is hypothesized that, compared with *trans* elements, mutations in CREs have a wider range of effects, giving rise to a greater diversity of phenotypes that could be selected upon (Wray 2007).

Previous studies have described distributions of mutational effects for several prokaryotic CREs both in vitro (Maerkl and Quake 2007, 2009; Geertz et al. 2012) and in vivo (Patwardhan et al. 2009; Kinney et al. 2010; Brewster et al. 2012; Sharon et al. 2012; Kosuri et al. 2013). As these studies predominantly focused on characterizing general relationships between *trans* factors (TFs and RNAP) and their *cis* binding sites, the analysis of interactions between individual *cis* mutations has been limited (Kwasnieski et al. 2012), partly due to restrictions in the techniques used (Melnikov et al. 2012; Patwardhan et al. 2012). Understanding the dependence of the effect of a mutation on the genetic background in which it appears, a phenomenon termed epistasis (Fisher 1918; Phillips 2008) is

critical to understanding adaptation and the engineering of synthetic promoters with specific properties (Kinkhabwala and Guet 2008). That is, the phenotypes (gene expression levels, in the case of CREs) of individuals containing multiple mutations are expected to correspond to those of the underlying single mutant phenotypes. Instead, however, the phenotypes might deviate from this expectation, resulting in positive epistasis if the mean double mutant expression is greater than the expression level predicted from single mutants. If the double mutant expression is lower than predicted, then mutations are considered to be in negative epistasis. If the individual mutation causes an increase in the expression but the double mutant containing that mutation leads to its reduction (or vice versa), the mutation is deemed to be in sign epistasis (Phillips 2008). Epistatic interactions between mutations in CREs define the robustness as well as the evolvability of regulatory elements—not only how transcription levels can be modulated but also how new functional CREs could evolve (Payne and Wagner 2014).

The functional effects of epistatic interactions are complex and therefore poorly understood, as epistasis may be influenced by many factors (Lehner 2011). Epistasis can depend on environmental factors, as has been demonstrated for mutations in bacterial and bacteriophage proteins (You and Yin 2002; Hayden and Wagner 2012; Lalic and Elena 2012; Wang, Diaz Arenas, et al. 2012; de Vos et al. 2013; Flynn et al. 2013; Caudle et al. 2014). As cellular responses to environmental changes can be complex, understanding their effect on epistasis is inherently difficult. Some prokaryotic regulatory pathways, however, offer a tractable system for understanding the

Open Access

Article

environmental response, with predictable responses to single stimuli modulating gene expression (Browning and Busby 2004). One such regulatory pathway is that of the well-studied *Escherichia coli* arabinose operon *araBAD* (Helling and Weinberg 1963; Schleif 2000) (fig. 1). This operon is activated in the presence of its substrate, the sugar arabinose, which interacts with the transcription factor AraC. By synthetically associating the *araBAD* CRE with a fluorescence marker instead of the native *araBAD* operon, we have decoupled the effects of mutations in CRE on expression from any potential downstream effects that might affect global expression (fig. 1c). This allows us to directly study the effects of mutations on CRE and their epistatic interactions, in the presence and absence of a single, well-understood, and controlled environmental variable, arabinose.

Epistatic interactions can also depend on the physical location of mutations. For example, the type and magnitude of interactions can be different for pairs of mutations found within a gene and those in different genes (Szendro et al. 2013), or for mutations found in genes that interact and those that do not (Lalic and Elena 2012). We investigated if the differences in epistasis in CRE might depend on whether those mutations are found in the same or in different operators. Epistasis in a CRE could arise from the effects of mutations on transcription factor binding to its operator site. But it could also arise from an interaction between the effects of mutations and the constraints imposed by protein–protein interactions that stabilize transcription factor DNA binding, as could be true for both AraC and the AraC–arabinose complex, which bind as dimers (Schleif 2003). Epistasis could also depend on the relative importance of an operator site for the control of transcription, which could be the case as both AraC and AraC–arabinose complex preferentially binding to *araI₁* (Schleif 2010).

Our knowledge of epistasis comes almost exclusively from studying the effects of mutations on protein-coding sequences. With the exception of a few studies that looked at pairwise interactions (Kwasnieski et al. 2012), statistical epistasis (Otwinowski and Nemenman 2013), and broad patterns of interactions (Patwardhan et al. 2009, 2012; Sharon et al. 2012; Jolma et al. 2013), assessments of epistatic interactions between mutations in CREs have largely been absent. Given the potentially large role played by changes in *cis* regulatory regions on shaping evolutionary outcomes, we explored how epistatic interactions in a CRE depend on both the environment and the location of mutations. We introduced single and double point mutations into *araI₁* and *araI₂*, the two proximal AraC operator sites of the *araBAD* CRE, and studied the effects of these mutations on expression in two different environments, defined by the presence or absence of the natural inducer arabinose. We found that both the environment and the location of the mutations in the CRE affect the nature of epistatic interactions.

Results

Mutational Effects

Most mutations, both single and double, significantly altered expression relative to the wild type (figs. 2 and 3 and supplementary tables S1 and S2, Supplementary Material online), which is not surprising as all mutated sites are fully conserved within the Enterobacteriaceae family (supplementary fig. S2c, Supplementary Material online). Surprisingly, three mutations increased expression in the presence of arabinose (fig. 2a), in apparent contradiction to previous reports that all possible mutations in tested sites decreased AraC binding by at least 10-fold (Niland et al. 1996). We tested if this disparity arose from the fact that the construct was on a low copy number plasmid rather than on the chromosome, but found no difference in the response of the two systems to arabinose (supplementary fig. S1, Supplementary Material online). The observed results, however, are in accordance with the position weight matrix for AraC binding (supplementary fig. S2, Supplementary Material online) obtained from RegulonDB (Salgado et al. 2013). As these sites are fully conserved, our result suggests that the optimal level of *araBAD* expression, even in the presence of arabinose, is lower than the maximum possible expression level. In the absence of arabinose, all tested mutations either significantly increased expression from *araBAD* CRE, leading to less tight repression, or left it unaffected (figs. 2b and 3b and supplementary table S2, Supplementary Material online).

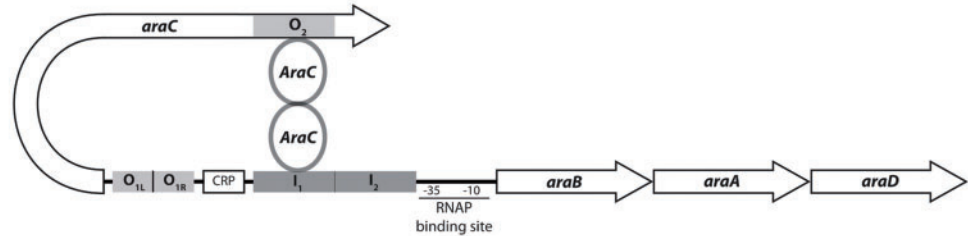### Epistatic Interactions Are Environment Dependent

We identified significant epistatic interactions, measured as the deviation of the double mutant expression from the multiplicative expectation of expression based on the corresponding single mutant effects, for exactly half (10 out of 20) of the double mutants in both environments (fig. 4 and supplementary tables S3 and S4, Supplementary Material online). Negative epistasis, where the expression of a double mutant is less than expected, was predominant in both environments (fig. 4): Only one double mutant in the presence of arabinose and two in its absence exhibited positive epistasis (fig. 4). In the presence of arabinose, sign epistasis was observed in 6 of the 10 epistatically interacting double mutants, as one of the component single mutations individually had a positive effect on expression while the double mutant negatively affected transcription (fig. 4b). The presence of epistasis did not depend on the magnitude of double mutant effects in either environment (supplementary fig. S4, Supplementary Material online).

Interestingly, these epistatic interactions were environment dependent, as evidenced by a significant $G \times G \times E$ interaction ($F_{19,120} = 21.51$, $P < 0.0001$). This interaction was not dependent on the three double mutants that changed the sign of interaction between two environments. The identities of most epistatically interacting double mutants differed depending on the presence of arabinose, sometimes even changing the sign of interaction between the two environments (fig. 4).
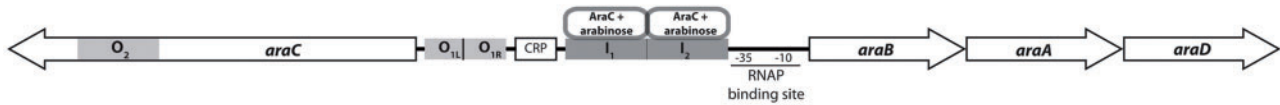
Epistatic Interactions Depend on the Physical Location of Mutations

We asked whether epistatic interactions differ depending on the location of the mutations—whether the magnitude and sign of epistasis differ if the mutations are in the *araI₁*, *araI₂*, or
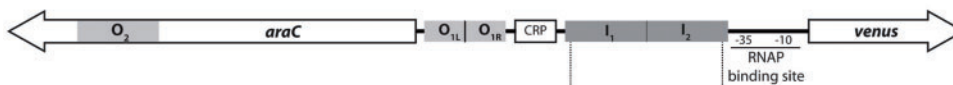
Fig. 1. Structure of *araBAD* operon and regulatory function of AraC. The *cis*-regulatory region of the *araBAD* operon consists of two proximal AraC operators, *araI₁* and *araI₂*, two intermediate operators, *araO₁L* and *araO₁R*, and the distal operator, *araO₂*. The *cis*-regulatory element also contains a CRP binding site. (a) In the absence of arabinose, AraC preferentially binds to operators *araI₁* and *araO₂*, forming a DNA loop and repressing transcription. (b) In the presence of arabinose, the AraC–arabinose complex binds to *araI₁* and *araI₂* operator sites, actively recruiting RNA polymerase and promoting transcription. (c) The structure of the experimental synthetic construct in which the *araBAD* operon has been replaced by a fluorescence reporter gene, *venus-yfp*. (d) We introduced point mutations in all base pairs in *araI₁* and *araI₂* operators for which it was previously determined through gel shift assays that any substitution decreased AraC binding by at least 10-fold (Niland et al. 1996).

with one mutation in each of the operators. We found a significant effect of which operator the mutations were in, both in the presence ($F_{2,17} = 25.083$, $P < 0.0001$) and absence of arabinose ($F_{2,17} = 39.089$, $P < 0.0001$) (fig. 5). We conducted pair-wise tests in order to analyze the differences in epistasis between mutations in the same or in different operators. In the presence of arabinose, we found greater negative epistasis when both mutations were in *araI₁* than in double mutants in *araI₂* ($t_8 = -3.257$, $P < 0.05$), and those with a mutation in each of the two operators ($t_{13} = -4.304$, $P < 0.001$) (fig. 5a). In the absence of arabinose, greater negative epistasis was observed in double mutants with mutations in different operators, when compared with those with both mutations in either operator *araI₁* ($t_{13} = 4.366$, $P < 0.001$) or operator *araI₂* ($t_{13} = 2.165$, $P < 0.05$).

Discussion

In this study, we evaluated the epistatic interactions between mutations in a *cis*-regulatory region of the *araBAD* operon.

We used a synthetic system that decouples the control of transcription from any posttranslational effects on fitness, thus allowing us an estimate of the effects of epistatic interactions on gene expression. We demonstrated that epistatic interactions are a common feature of the *araBAD* CRE, as we observed them in exactly half of the tested double mutants in both environments (fig. 4). Furthermore, we showed that interactions between mutations are strongly environmentally dependent (fig. 4).

The dependence of epistasis on the environment has been previously reported (You and Yin 2002; Hayden and Wagner 2012; Lalic and Elena 2012; Wang, Diaz Arenas, et al. 2012; de Vos et al. 2013; Flynn et al. 2013; Caudle et al. 2014), but only in proteins and under conditions when the environment elicits a complex cellular response. Here, the sole difference between the two environments is the presence or absence of arabinose, which binds to AraC to form the AraC–arabinose complex (Schleif 2010), which in turn affects the binding affinity of AraC for its operator sites (Lobell and Schleif 2005).
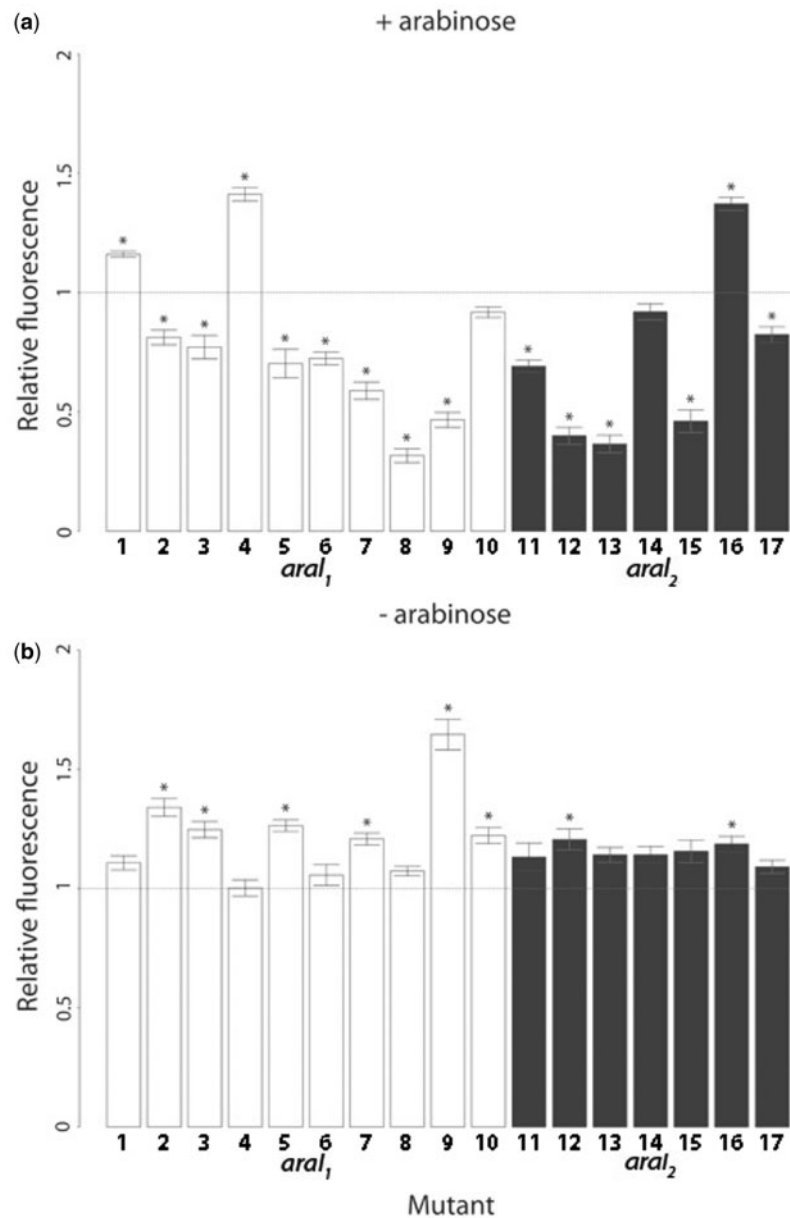
Fig. 2. Relative fluorescence of single mutants in two environments. Bars are mean fluorescence relative to wild type. White bars are single mutants in the operator $araI_1$, dark bars are single mutants in $araI_2$. Stars indicate mutants that significantly differ from the wild type. The dotted line represents wild-type fluorescence, normalized to 1. Error bars are standard errors of the mean. Measurements were taken in the (a) presence and (b) absence of arabinose.

We demonstrated that such a simple and controllable environmental change is sufficient to drastically alter epistatic interactions between mutations in operator sites.

The relationship between epistasis and environment plays a crucial role in determining how a population responds to environmental change (de Vos et al. 2013). When epistatic interactions between the same mutations are environment dependent, beneficial genotypes do not necessarily correlate between environments (Coyne et al. 1997, 2000; Goodnight and Wade 2000). In terms of transcriptional regulation, selection toward one optimum in one environment might perturb regulation in the other environment in an unpredictable and uncorrelated manner. To illustrate this point, consider a scenario with strong selection to lower expression when the

system is in its activated state (i.e., when arabinose is present). In order to most drastically decrease expression, selection might favor negatively epistatic mutations, for example, double mutant (4,10) in our experiment. When arabinose becomes depleted from the environment, these same mutations would be in positive epistasis, so that repression is less tight and efficient. Such effects would be particularly emphasized when mutations exhibit sign epistasis in one environment but not in the other, and therefore a more rugged fitness landscape is translated into a smoother one. Given the frequency and relative importance of changes in CREs for evolution (Wray 2007), understanding how fitness landscapes change between environments is of critical importance.
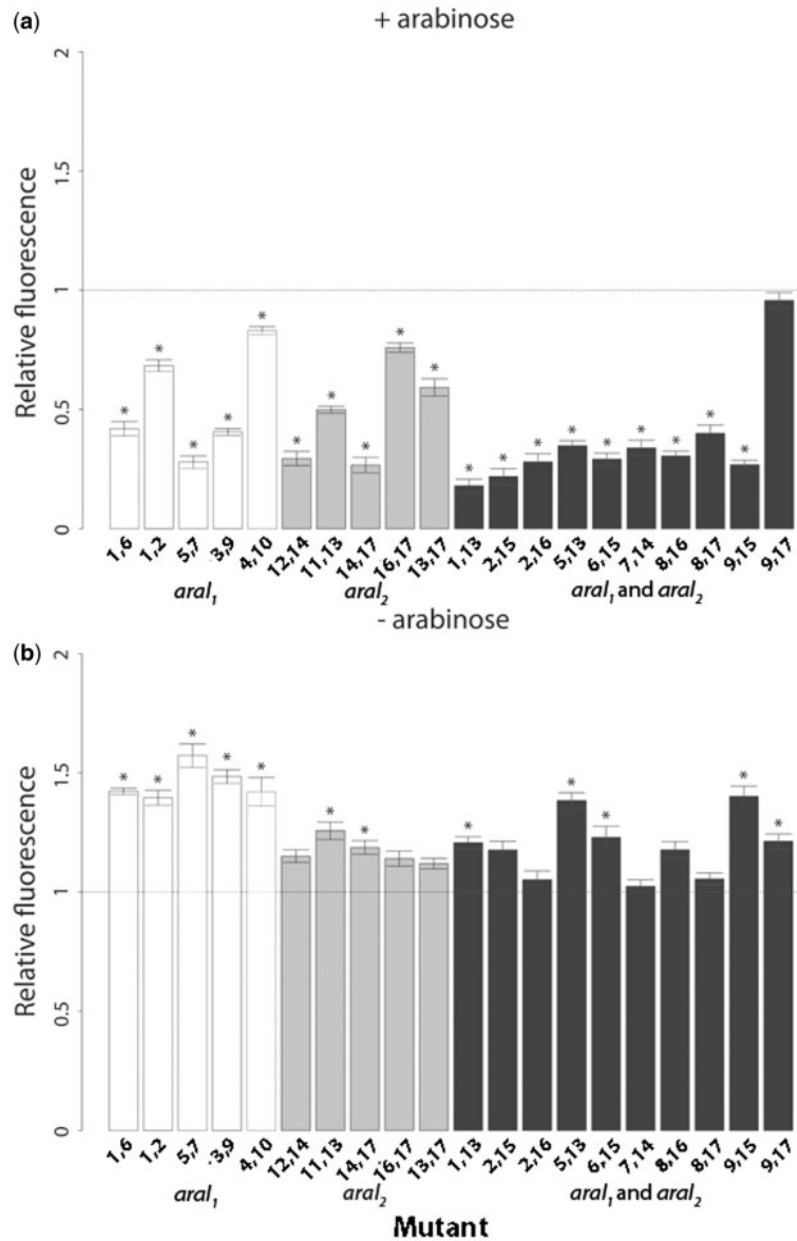
Fig. 3. Relative fluorescence of double mutants in two environments. Bars are mean fluorescence relative to wild type. White bars are double mutants with both mutations in the operator *araI₁*; light gray bars with both mutations in *araI₂*; and dark bars have one mutation in each of the regions (*araI₁* and *araI₂*). Stars indicate mutants that significantly differ from the wild type. The dotted line represents the wild-type relative fluorescence of 1. Error bars are standard errors of the mean. Measurements were taken in the (*a*) presence and (*b*) absence of arabinose.

We also demonstrated that epistasis depends on the physical location of mutations, and in particular whether mutations are located in the same or in different operators (fig. 5). This phenomenon is conceptually similar to the prediction that epistasis is stronger when mutations are found in the same genes (Szendro et al. 2013). In the presence of arabinose, when the system is in its activated state, the mutations in operator *araI₁* show stronger negative epistasis than those in operator *araI₂* (fig. 5*a*). This might reflect the stronger binding affinity of the AraC–arabinose dimer for *araI₁* than for *araI₂* (Niland et al. 1996). In the absence of arabinose, the *araI₁* operator plays a dominant role in repression, as the AraC dimer is bound to it and to operator *araO₂* to form a DNA loop (Lobell and Schleif 2005). Given this model of AraC regulation, the observation of any effects of single mutations in *araI₂* on expression as well as the occurrence of epistatic interactions between them in the absence of arabinose is surprising, as *araI₂* should not be bound to AraC (Schleif 2010). It is possible that mutations in operator *araI₂* affect the accessibility of *araI₁* by modifying the local tertiary structure of the DNA, as has been shown for mutations at sites flanking the transcription factor binding site (Levo and Segal 2014). Such effects depend on the local DNA context (Gordân et al. 2013), so that changes in the stability of the DNA loop could alter mutational effects on expression. Therefore, a mutation in operator *araI₂* that increases expression in the absence of
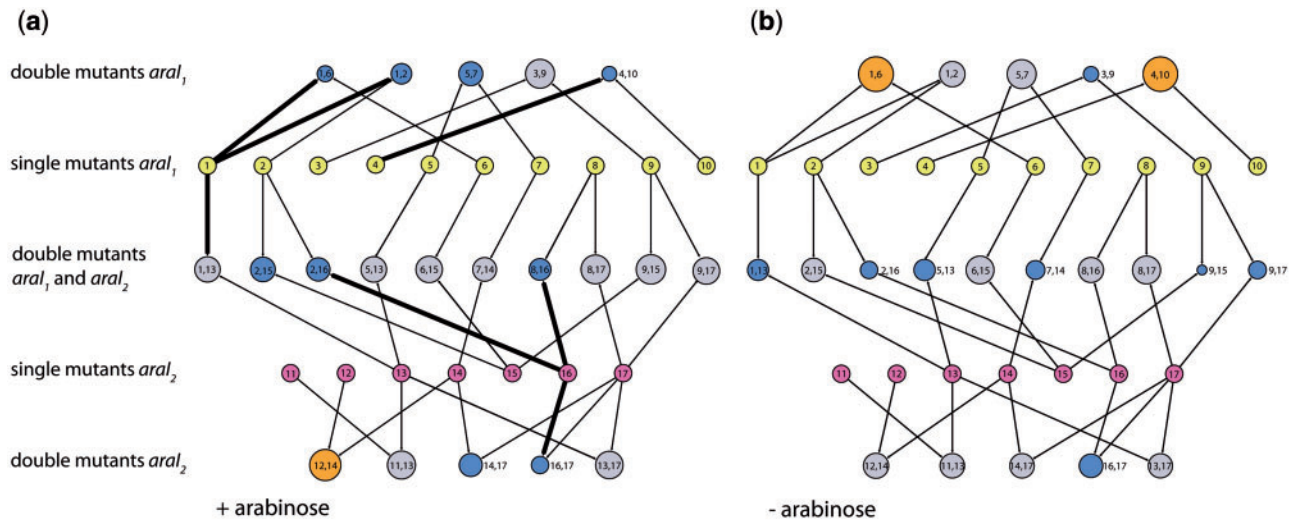
Fig. 4. Epistasis in the mutation network. Single mutations are shown in yellow and purple, and are connected to the corresponding double mutant(s) that contain them. For double mutants, the size of the circle is proportional to the magnitude of epistasis, with more negative epistasis values corresponding to smaller circles. Significant negative epistasis is shown in blue, and significant positive interactions in orange, while noninteracting mutations are in gray. Thick connecting lines indicate sign epistasis, with the effect of a single mutation having an opposite sign to that of a double mutant. Interaction network was measured in (a) the presence of arabinose and (b) the absence of arabinose.
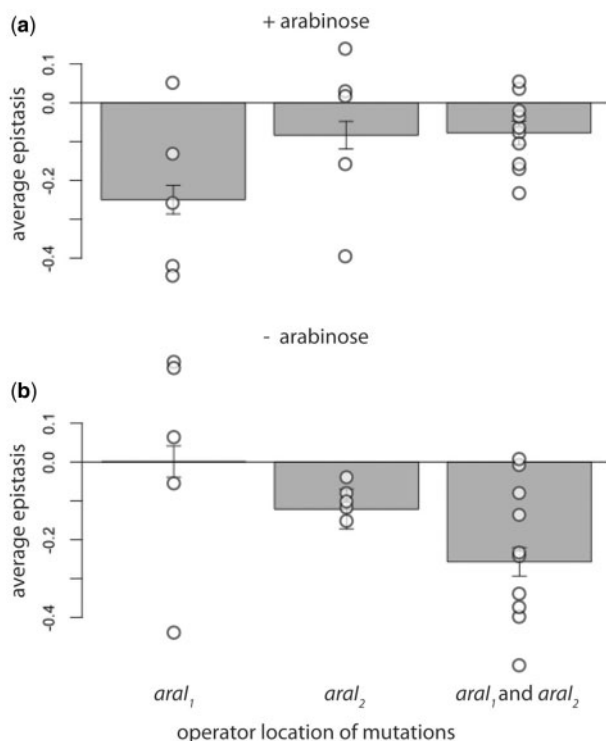

Fig. 5. Epistatic interactions depend on the location of mutations. Bars are mean epistasis values for double mutants in $aral_1$, in $aral_2$, and those with one mutation in each of the regions ($aral_1$ and $aral_2$), respectively. Circles are epistasis values for individual double mutants measured in the (a) presence of arabinose and (b) absence of arabinose. Error bars are standard deviations.

arabinose by destabilizing the looping structure could be doing so by modifying the local DNA structure and hence decreasing AraC binding at operator $aral_1$. Such an effect could also explain the observed negative epistasis when the

two mutations are in different operators, as they are not predicted to directly interact (fig. 5b).

In this study, we explored the effects of mutations only on the direct phenotype, that is, gene expression, allowing measurements of epistasis present in the genotype–phenotype map. Epistatic interactions can also arise from nonlinearity in the phenotype-fitness mapping (de Visser et al. 2011). By creating an isolated, synthetic system we studied only the nature of the genotype–phenotype mapping and its environment dependence, without considering any downstream effects on fitness. As fitness effects of mutations on CREs depend on the particular properties of the regulated gene, and cannot be a priori inferred (Wittkopp and Kalay 2012), constructing a full genotype–phenotype fitness map was beyond the scope of this work.

Our observation of the dependency of epistatic interactions on the environment and the location of mutations have implications not only for the evolution of CREs but also for engineering regulatory elements with desired expression profiles, a critical task for the optimal design of functional synthetic systems (Purnick and Weiss 2009; Wang, Ma, et al. 2012; Levo and Segal 2014). Epistatic interactions might also impact the overall function of a synthetic construct, potentially leading to unpredictable network properties (Guet et al. 2002). We have shown that large modifications to expression levels are more likely to be achieved by modifying the operator site with highest affinity for the transcription factor, as the epistatic interactions in that operator might be stronger. The fine tuning of expression, on the other hand, is best achieved by introducing mutations in different operators. Thus, the epistatic landscape of CREs may have arisen from the difficulty in finding solutions to two opposing forces—optimal expression in the presence of the inducer and the ability to repress the operon when not required.

Materials and Methods

The transcription factor, AraC, has five operator binding sites in the CRE of the *araBAD* operon—two proximal sites ($araI_2$ and $araI_1$), two intermediate sites ($araO_{1L}$ and $araO_{1R}$), and a distal site ($araO_2$) (Schleif 2010). In the absence of the natural inducer arabinose, AraC exists predominantly as a dimer bound to the $araI_1$ and $araO_2$ operator sites, forming a DNA loop that prevents transcription (Schleif 2003) (fig. 1a). The sugar arabinose acts as an inducer, by binding to AraC and introducing a conformational change that prevents DNA looping by the preferential binding of the AraC–arabinose complex to the $araI_1$ and $araI_2$ sites (fig. 1b). In addition, binding at the $araI_2$ site directly recruits RNAP, activating transcription of the operon (Schleif 2010).

Construction of the Plasmid System

The experimental construct consisted of the native *E. coli* K-12 *araBAD* operon regulatory region containing the promoter (P_{BAD}), the upstream regulatory region consisting of five AraC operators involved in *araBAD* operon regulation ($araI_1$, $araI_2$, $araO_{1L}$, $araO_{1R}$, and $araO_2$), and the *araC* gene followed by a terminator sequence (fig. 1c). The *araBAD* operon was replaced by the fluorescent protein *venus-YFP* (Nagai et al. 2002), followed by the *E. coli* alpha operon tL17 terminator. This construct allows the effects of mutations in the *araBAD* operators to be measured in terms of their effect on expression. The whole construct was cloned into a low copy number plasmid pZS* with a kanamycin resistance marker (Lutz and Bujard 1997).

Mutant Library Construction

A library of single and double mutants in the $araI_1$ and $araI_2$ operators involved in the regulation of the *araBAD* promoter was created using Quick-Change II™ site-directed mutagenesis protocol (Agilent Technologies). Following the mutagenesis protocol, plasmids were cloned into BW25113 strain (CGSC# 7636) in which the *araBAD* operon has been deleted (Datsenko and Wanner 2000), and were then plated on Luria-Bertani (LB) plates and 50 µg/ml kanamycin.

We introduced point mutations in all base pairs in $araI_1$ and $araI_2$ operators for which it was previously reported that any substitution decreased AraC binding by at least 10-fold (Niland et al. 1996). Based on a search of public sequence databases, the $araI_1$ and $araI_2$ operator sites are highly conserved within the Enterobacteriaceae family (supplementary fig. S2c, Supplementary Material online). For each mutation, we randomly selected the base to be introduced, the only constraint being that the ratio between transitions and transversions was approximately 1:2. We tested a total of 17 single mutants, 10 in $araI_1$ and 7 in $araI_2$ (fig. 1), and 20 double mutants. The double mutations consisted of random combinations of single mutants. For five of the double mutants, both mutations were in the $araI_1$ operator; five were both in the $araI_2$ operator; and ten had one mutation in each of the two operators (supplementary fig. S4, Supplementary Material online).

Expression Assays

Single colonies with the desired point mutations (each confirmed by Sanger sequencing) were grown overnight at 37 °C on LB containing 50 µg/ml kanamycin. These cultures were used to inoculate four replicate populations in arabinose and four in the absence of arabinose. These were grown overnight in M9 media, supplemented with 0.01% casamino acids, 50 µg/ml kanamycin, 0.2% glycerol, and either containing 0.1% arabinose or not. The populations were then serially diluted, grown for 4 h, and then used to inoculate 1.2 ml of the corresponding media with 0.1% of the grown culture, to ensure that measurements are taken during the exponential growth. When the cultures reached an OD_{600} of approximately 0.05, 150 µl of the culture was sampled and fluorescence and OD_{600} measurements taken using Biotek H1 plate reader. Fluorescence measurements were normalized by the OD_{600} measurements to account for the variation in the size of the initial inoculum. With this design, the expression of each mutant was measured four times in each environment, with measurements performed on independent cultures.

Chromosome Insertion

The observed disparity between the effects of mutations measured by Niland et al. (1996) and in our study could have been a result of a difference in how the system behaves when on a low copy number plasmid or in the chromosome. To test if this were true, we inserted the wild-type construct with the kanamycin resistance gene into the *araBAD–araC* locus on the BW25113 chromosome using lambda-red recombineering. We tested how the two systems respond to arabinose by measuring fluorescence in the manner previously described, on a variety of arabinose concentrations (0.2%, 0.1%, 0.05%, 0.025%, 0.0125%, 0.00625%, 0.003125%, and 0%). We used six replicates at each concentration.

Data Analysis

Expression across the Mutant Library

For each mutant, we measured the fluorescence and normalized it by the fluorescence for the wild-type plasmid in the same environment. Because there is variation in the fluorescence of the wild-type strain, we used error propagation when calculating standard deviations of the mean normalized fluorescence (Ku 1966). We tested for an effect of each mutant on expression using ANOVA (aov function in R statistical software version 3.1.1.; R Core Team 2015), with relative fluorescence as the response variable, mutation as the fixed factor (37 levels), and replicate (4 levels) as a random factor, and used post hoc FDR-corrected *t*-tests to compare the mean fluorescence of each mutant to that of the wild type.

Epistatic Interactions

To estimate the interaction between two mutations, we used a multiplicative epistasis model, as the studied mutations were not expected to be independent of each other (Cordell 2002). In this model, epistasis is calculated as $\varepsilon = \omega_{m12} - \omega_{m1} \times \omega_{m2}$, where ω_{m12} is the relative fluorescence of a double mutant, and ω_{m1} and ω_{m2} the relative

fluorescence of the two corresponding single mutants, respectively. It is worth noting that we calculated epistasis based on expression levels of a reporter gene, rather than based on the strength of binding of the transcription factor to the operator. The relationship between the epistasis based on expression to that based on binding depends on the role of the transcription factor and the context of its binding. Broadly speaking, if two mutations have lower expression than expected based on the single mutant phenotypes (i.e., show negative epistasis on expression), this might be due to negative epistasis on activator binding, or positive epistasis on repressor binding.

To test whether the estimated epistasis was significantly different from zero, we conducted a series of FDR-corrected t-tests. The errors were calculated based on four replicates, using error propagation to account for the inherent variance of each replicate that was due to normalization by the wild type. To test for a relationship between epistasis and the magnitude of the corresponding double mutant effects, we used a linear regression model weighted by the cumulative error. To test for the effects of the environment on epistasis, we tested for a genotype-by-genotype-by-environment ($G \times G \times E$) interaction using ANOVA, with replicate as a random factor. The difference in the magnitude and sign of epistasis was compared between the double mutants depending on the region that the mutations were in (both mutations in operator $araI_1$, both in operator $araI_2$, or one mutation in each of the two operators). FDR-corrected pairwise t-tests were carried out between the three regions, with mean epistasis as a response variable. We tested whether magnitude of epistasis depended on the proximity of mutations within an operon by carrying out ANOVA with distance between introduced mutations in base pairs as the fixed factor. Possibly due to small sample size, we did not find a significant effect of the distance between mutations in either environment (in the absence of arabinose: $F_{1,8} = 0.206$, $P = 0.662$; in the presence of arabinose: $F_{1,8} = 0.163$, $P = 0.697$).

Supplementary Material

Supplementary figures S1–S4 and tables S1–S4 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

References

Brewster RC, Jones DL, Phillips R. 2012. Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. *PLoS Comput Biol.* 8:e1002811.

Browning DF, Busby SJW. 2004. The regulation of bacterial transcription initiation. *Nat Rev Microbiol.* 2:57–65.

Carroll SB. 2000. Endless forms: the evolution of gene regulation and morphological diversity. *Cell* 101:577–580.

Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134:25–36.

Caudle SB, Miller CR, Rokyta DR. 2014. Environment determines epistatic patterns for a ssDNA virus. *Genetics* 196:267–279.

Cordell HJ. 2002. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet.* 11:2463–2468.

Coyne JA, Barton NH, Turelli M. 1997. Perspective: a critique of Sewall Wright's shifting balance theory of evolution. *Evolution* 51:643–671.

Coyne JA, Barton NH, Turelli M. 2000. Is Wright's shifting balance process important in evolution? *Evolution* 54:306–317.

Datsenko KA, Wanner BR. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* 97:6640–6645.

de Visser JAGM, Cooper TF, Elena SF. 2011. The causes of epistasis. *Proc R Soc B Biol Sci.* 278:3617–3624.

de Vos M, Poelwijk FJ, Battich N, Ndika JDT, Tans SJ. 2013. Environmental dependence of genetic constraint. *PLoS Genet.* 9:e1003580.

Fisher RA. 1918. The correlations between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb.* 52:399–433.

Flynn KM, Cooper TF, Moore FB-G, Cooper VS. 2013. The environment affects epistatic interactions to alter the topology of an empirical fitness landscape. *PLoS Genet.* 9:e1003426.

Geertz M, Shore D, Maerkl SJ. 2012. Massively parallel measurements of molecular interaction kinetics on a microfluidic platform. *Proc Natl Acad Sci U S A.* 109:16540–16545.

Goodnight CJ, Wade MJ. 2000. The ongoing synthesis: a reply to Coyne, Barton, and Turelli. *Evolution* 54:317–324.

Gordân R, Shen N, Dror I, Zhou T, Horton J, Rohs R, Bulyk ML. 2013. Genomic regions flanking E-box binding sites influence DNA binding specificity of bHLH transcription factors through DNA shape. *Cell Rep.* 3:1093–1104.

Guet CC, Elowitz MB, Hsing W, Leibler S. 2002. Combinatorial synthesis of genetic networks. *Science* 296:1466–1470.

Hayden EJ, Wagner A. 2012. Environmental change exposes beneficial epistatic interactions in a catalytic RNA. *Proc R Soc B Biol Sci.* 279:3418–3425.

Helling RB, Weinberg R. 1963. Complementation studies of arabinose genes in *Escherichia coli*. *Genetics* 48:1397–1410.

Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol.* 3:318–356.

Jolma A, Yan J, Whitington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G, et al. 2013. DNA-binding specificities of human transcription factors. *Cell* 152:327–339.

King MC, Wilson AC. 1975. Evolution at two levels in humans and chimpanzees. *Science* 188:107–116.

Kinkhabwala A, Guet CC. 2008. Uncovering *cis* regulatory codes using synthetic promoter shuffling. *PLoS One* 3:e2030.

Kinney JB, Murugan A, Callan CGJ, Cox EC. 2010. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc Natl Acad Sci U S A.* 107:9158–9163.

Kosuri S, Goodman DB, Cambray G, Mutalik VK, Gao Y, Arkin AP, Endy D, Church GM. 2013. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 110:14024–14029.

Ku HH. 1966. Notes on the use of propagation of error formulas. *J Res Natl Bureau Stand C Eng Instrum.* 70C:263–273.

Kwasnieski JC, Mogno I, Myers CA, Corbo JC, Cohen BA. 2012. Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc Natl Acad Sci U S A.* 109:19498–19503.

Lalic J, Elena SF. 2012. Epistasis between mutations is host-dependent for an RNA virus. *Biol Lett.* 9:20120396.

Lehner B. 2011. Molecular mechanisms of epistasis within and between genes. *Trends Genet.* 27:323–331.

Levo M, Segal E. 2014. In pursuit of design principles of regulatory sequences. *Nat Rev Genet.* 15:453–468.

Lobell RB, Schleif R. 2005. DNA looping and unlooping by AraC protein. *Science* 250:528–532.

Lutz R, Bujard H. 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res.* 25:1203–1210.

Maerkl SJ, Quake SR. 2007. A systems approach to measuring the binding energy landscapes of transcription factors. *Science* 315:233–237.

Maerkl SJ, Quake SR. 2009. Experimental determination of the evolvability of a transcription factor. *Proc Natl Acad Sci U S A.* 106:18650–18655.

Melnikov A, Murugan A, Zhang X, Tesileanu T, Wang L, Rogov P, Feizi S, Gnirke A, Callan CG, Kinney JB, et al. 2012. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat Biotechnol.* 30:271–277.

Nagai T, Ibata K, Park ES, Kubota M, Mikoshiba K, Miyawaki A. 2002. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat Biotechnol.* 20:87–90.

Niland P, Hühne R, Müller-Hill B. 1996. How AraC interacts specifically with its target DNAs. *J Mol Biol.* 264:667–674.

Otwinowski J, Nemenman I. 2013. Genotype to phenotype mapping and the fitness landscape of the *E. coli lac* promoter. *PLoS One* 8:e61570.

Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, Lee C, Andrie JM, Lee SI, Cooper GM, et al. 2012. Massively parallel functional dissection of mammalian enhancers *in vivo. Nat Biotechnol.* 30:265–270.

Patwardhan RP, Lee C, Litvin O, Young DL, Pe'er D, Shendure J. 2009. High-resolution analysis of DNA regulatory elements by synthetic saturation mutagenesis. *Nat Biotechnol.* 27:1173–1175.

Payne JL, Wagner A. 2014. The robustness and evolvability of transcription factor binding sites. *Science* 343:875–877.

Phillips PC. 2008. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet.* 9:855–867.

Purnick PEM, Weiss R. 2009. The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol.* 10:410–422.

R Core Team. 2015. A language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Salgado H, Peralta-Gil M, Gama-Castro S, Santos-Zavaleta A, Muñiz-Rascado L, García-Sotelo JS, Weiss V, Solano-Lira H, Martínez-Flores I, Medina-Rivera A, et al. 2013. RegulonDB v8.0: omics data sets, evolutionary conservation, regulatory phrases, cross-validated gold standards and more. *Nucleic Acids Res.* 41:D203–D213.

Schleif R. 2000. Regulation of the L-arabinose operon of *Escherichia coli. Trends Genet.* 16:559–565.

Schleif R. 2003. AraC protein: a love-hate relationship. *Bioessays* 25:274–282.

Schleif R. 2010. AraC protein, regulation of the L-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. *FEMS Microbiol Rev.* 34:779–796.

Sharon E, Kalma Y, Sharp A, Raveh-Sadka T, Levo M, Zeevi D, Keren L, Yakhini Z, Weinberger A, Segal E. 2012. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat Biotechnol.* 30:521–530.

Stern DL, Orgogozo V. 2008. The loci of evolution: how predictable is genetic evolution. *Evolution* 62:2155–2177.

Szendro IG, Schenk MF, Franke J, Krug J, de Visser JAGM. 2013. Quantitative analyses of empirical fitness landscapes. *J Stat Mech.* 2013.01:P01005.

Wang T, Ma X, Du G, Chen J. 2012. Overview of regulatory strategies and molecular elements in metabolic engineering of bacteria. *Mol Biotechnol.* 52:300–308.

Wang Y, Diaz Arenas C, Stoebel DM, Cooper TF. 2012. Genetic background affects epistatic interactions between two beneficial mutations. *Biol Lett.* 9:20120328.

Wittkopp PJ, Kalay G. 2012. *Cis*-Regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat Rev Genet.* 13:59–69.

Wray GA. 2007. The evolutionary significance of *cis*-regulatory mutations. *Nat Rev Genet.* 8:206–216.

You L, Yin J. 2002. Dependence of epistasis on environment and mutation severity as revealed by in silico mutagenesis of phage T7. *Genetics* 160:1273–1281.