J A C S
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Subscriber access provided by ISTANBUL TEKNIK UNIV

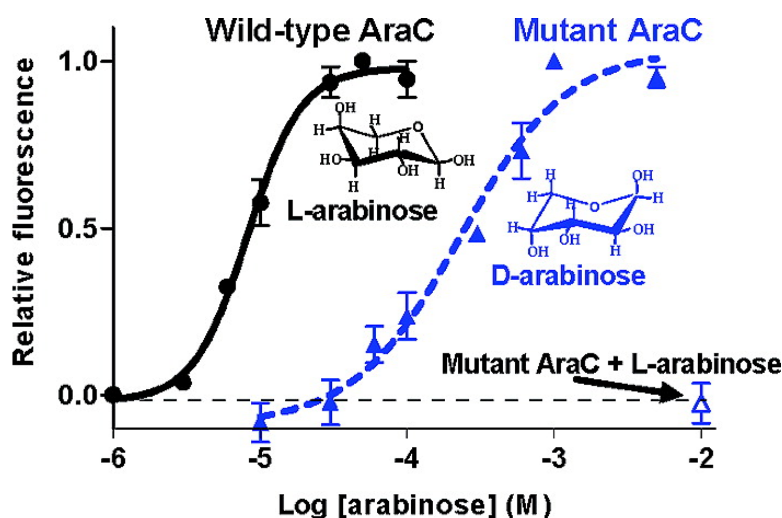
Article

AraC Regulatory Protein Mutants with Altered Effector Specificity

Shuang-Yan Tang, Hossein Fazelinia, and Patrick C. Cirino

J. Am. Chem. Soc., **2008**, 130 (15), 5267-5271 • DOI: 10.1021/ja7109053 • Publication Date (Web): 21 March 2008

Downloaded from http://pubs.acs.org on February 8, 2009



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

ACS Publications
High quality. High impact.

Journal of the American Chemical Society is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

JACS ARTICLES

Published on Web 03/21/2008

AraC Regulatory Protein Mutants with Altered Effector Specificity

Shuang-Yan Tang, Hossein Fazelinia, and Patrick C. Cirino*

Department of Chemical Engineering, The Pennsylvania State University, University Park, Pennsylvania 16802

Received December 6, 2007; E-mail: cirino@engr.psu.edu

Abstract: The AraC regulatory protein of the *Escherichia coli ara* operon has been engineered to activate transcription in response to D-arabinose and not in response to its native effector L-arabinose. Two different AraC mutant libraries, each with four randomized binding pocket residues, were subjected to FACS-mediated dual screening using a GFP reporter. Both libraries yielded mutants with the desired switch in effector specificity, and one mutant we describe maintains tight repression in the absence of effector. The presence of 100 mM L-arabinose does not influence the response of the reported mutants to D-arabinose, and the mutants are not induced by other sugars tested (D-xylose, D-fucose, D-lyxose). Co-expression of the FucP transporter in *E. coli* enabled induction by D-arabinose in the 0.1 mM range. Our results demonstrate the power of dual screening for altering AraC inducer specificity and represent steps toward the design of customized *in vivo* molecular reporters and gene switches for metabolic engineering.

Introduction

Engineered regulatory proteins which control transcription in response to non-native small molecule stimuli find use as gene switches in a wide range of applications including gene therapy, metabolic engineering, biosensing, and environmental remediation.[1–7] Several bacterial regulatory proteins have been engineered to respond to novel small molecules, as reviewed by Galvao et al.[3] Mutations that alter molecular recognition in proteins generally result in mutants with relaxed specificity,[2,3,8–12] so specificity must be intentionally retained (or screened for) during the design process.[13–15] Whereas other popular bacterial transcriptional regulatory proteins only repress (e.g., LacI and TetR) or activate (e.g., LuxR) gene expression from their cognate promoters, the AraC protein of the *Escherichia coli ara* operon functions as a dual regulator, repressing transcription in the absence of effector (L-arabinose) and activating transcription in the presence of L-arabinose.[16] The resulting stringent control over gene expression and high induction ratios are attractive features for biotechnological applications.[17,18]

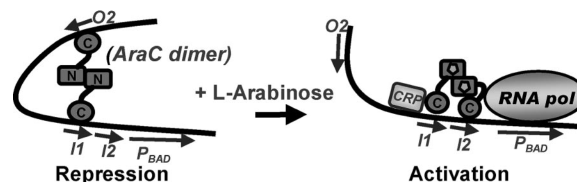


Figure 1. Mechanism of dual regulation by AraC at the P_{BAD} promoter, adapted from Schleif.[19] I1, I2, and O2 represent DNA binding half-sites for the AraC DBDs. CRP is a coactivator (requiring cAMP) and RNA pol represents RNA polymerase.

Schleif and co-workers have characterized AraC and the mechanisms of *ara* operon regulation and proposed the "light-switch" mechanism depicted in Figure 1.[19] In the absence of L-arabinose, the DNA-binding domains (DBDs) of an AraC dimer bind the I1 and O2 half-sites (separated by 210 bases), repressing transcription through the formation of a DNA loop upstream of the P_{BAD} promoter. Upon binding L-arabinose, the dimer changes conformation such that the DBDs bind the adjacent I1 and I2 half-sites, resulting in transcriptional activation via interactions with RNA polymerase at P_{BAD}. Induction

(1) Cebolla, A.; Sousa, C.; de Lorenzo, V. *Nucleic Acids Res.* **2001**, *29*, 759–766.
(2) Cebolla, A.; Sousa, C.; deLorenzo, V. *J. Biol. Chem.* **1997**, *272*, 3986–3992.
(3) Galvao, T. C.; de Lorenzo, V. *Curr. Opin. Biotechnol.* **2006**, *17*, 34–42.
(4) Krueger, C.; Danke, C.; Pfleiderer, K.; Schuh, W.; Jack, H. M.; Lochner, S.; Gmeiner, P.; Hillen, W.; Berens, C. *J. Gene Med.* **2006**, *8*, 1037–1047.
(5) Krueger, C.; Schmidt, A.; Danke, C.; Hillen, W.; Berens, C. *Gene* **2004**, *331*, 125–131.
(6) Ramos, J. L.; Wasserfallen, A.; Rose, K.; Timmis, K. N. *Science* **1987**, *235*, 593–596.
(7) Tropel, D.; van der Meer, J. R. *Microbiol. Mol. Biol. Rev.* **2004**, *68*, 474–500.
(8) Collins, C. H.; Arnold, F. H.; Leadbetter, J. R. *Mol. Microbiol.* **2005**, *55*, 712–723.
(9) Derr, P.; Boder, E.; Goulian, M. *J. Mol. Biol.* **2006**, *355*, 923–932.
(10) Scholz, O.; Kostner, M.; Reich, M.; Gastiger, S.; Hillen, W. *J. Mol. Biol.* **2003**, *329*, 217–227.
(11) Varadarajan, N.; Gam, J.; Olsen, M. J.; Georgiou, G.; Iverson, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6855–6860.
(12) Wise, A. A.; Kuske, C. R. *Appl. Environ. Microbiol.* **2000**, *66*, 163–169.
(13) Chockalingam, K.; Chen, Z. L.; Katzenellenbogen, J. A.; Zhao, H. M. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5691–5696.
(14) Collins, C. H.; Leadbetter, J. R.; Arnold, F. H. *Nat. Biotechnol.* **2006**, *24*, 708–712.
(15) Henssler, E. M.; Scholz, O.; Lochner, S.; Gmeiner, P.; Hillen, W. *Biochemistry* **2004**, *43*, 9512–9518.
(16) Schleif, R. *Trends Genet.* **2000**, *16*, 559–565.
(17) Guzman, L. M.; Belin, D.; Carson, M. J.; Beckwith, J. *J. Bacteriol.* **1995**, *177*, 4121–4130.
(18) Keasling, J. D. *Trends Biotechnol.* **1999**, *17*, 452–460.
(19) Schleif, R. *BioEssays* **2003**, *25*, 274–282.

of the *ara* operon is specific to L-arabinose: structurally and chemically similar sugars such as D-xylose, D-arabinose, and D-fucose (6-deoxy-D-galactose) fail to act as wild-type AraC effectors.[20] Numerous studies point to critical interactions between an N-terminal AraC arm and the C-terminal DBD in the absence of inducer, and between the arm and ligand binding pocket in the presence of L-arabinose.[21–26] N-terminal point mutations that confer constitutivity or uninducibility are common,[22,23,27] and weakened interactions with the C-terminal domain result in induction by both L-arabinose and D-fucose.[26] The dual regulatory properties of AraC therefore rely on a sensitive switch between two conformations and tailored interactions with the effector.

Challenges in the generation of useful AraC variants that respond to different small molecules lie in retention of repressibility and specificity. We sought AraC mutants that are insensitive to L-arabinose and instead induce P_{BAD} specifically in response to the rarer isomer D-arabinose, and repress transcription in the absence of the effector. Saturation mutagenesis in the AraC ligand binding pocket coupled with FACS-based dual screening yielded multiple mutants with the desired change in specificity. Here we report the detailed *in vivo* characterization of two such mutants. Regulatory proteins that respond to rare sugars could serve as reporters in the design of enzymes or metabolic pathways that synthesize these compounds.[28] A D-arabinose-inducible AraC analogue can also replace the native protein in P_{BAD}-based gene expression systems for use in strains that consume L-arabinose and when L-arabinose is present as a component of a natural sugar mixture (e.g., in hemicellulosic hydrolysates).

Materials and Methods

Plasmid Construction. Plasmid maps and a cloning diagram are given in Figure S1 of Supporting Information. Primer sequences are listed in Table S1 of Supporting Information. AraC and AraC mutants were expressed from plasmid pPCC423 (Apr^R, pBR322 origin). Plasmid construction was as follows: *araC* was amplified from *E. coli* K-12 strain W3110 using primers *araC*-K12-for and *araC*-K12-rev, and the PCR product was "TOPO cloned" into vector pCR2.1-TOPO (Invitrogen), resulting in plasmid pPCC400. The cloned *araC* gene (sequenced to verify no mutations) was isolated from pPCC400 by enzyme digestion at the *Nde*I (5′) and *Kpn*I (3′) restriction sites and subsequently ligated into vector pFLAG-CTC (Sigma-Aldrich Co.), resulting in plasmid pPCC402 (gene expression is controlled by a *tac* promoter and a Shine Dalgarno RBS sequence upstream of the multiple cloning site that includes *Nde*I at the start codon). The *bla* gene (ampicillin resistance marker) in pPCC402 was next replaced with the *aac* gene expressing apramycin resistance, resulting in plasmid pPCC416A. To facilitate AraC N-terminal domain mutant library construction, silent mutations were made in the *araC* sequence at base positions 187 and 542, resulting in addition of *Eco*RI and *Hin*dIII restriction sites. The resulting plasmid pPCC423 was used for library construction.

Reporter plasmid pPCC442 was used for AraC library screening and wild-type (WT) and mutant characterization. The regulatory region of the native *araBAD* operon includes promoter P_C, controlling AraC expression opposite of P_{BAD}, as shown in Figure S1 (Supporting Information). The O2 half-site lies between the *araC* reading frame and the P_C promoter. AraC expression from this control region is autoregulated in a complex fashion.[16] To achieve uniform AraC expression levels among mutants, we chose the dual plasmid approach, with AraC expressed independent of the *ara* operon (from plasmid pPCC423). As depicted in Figure S1 (Supporting Information), the *araC* open reading frame was removed from plasmid pBAD-GFPuv (Clontech)[29] (derived from the pBAD plasmid series)[17] by PCR amplification of the entire plasmid (minus *araC*) using primers that flank either end of *araC* (pbad-GFPfor and pbad-GFPrev). The PCR product was digested with *Bgl*II and self-ligated to form plasmid pPCC438, containing reporter construct "O2-P_{BAD}-*gfpuv*". pPCC438 was next digested with *Kpn*I, and the P_{BAD}-*gfpuv* fragment was ligated into high-copy RSF1030-derivative cloning vector pDHC29[30] resulting in plasmid pPCC442. pPCC423 and pPCC442 are compatible *E. coli* vectors. As shown in Figure S2 (Supporting Information), the dual plasmid system with WT AraC (expressed from pPCC423) transformed into strain HF19 gives a fluorescence response to L-arabinose that is nearly identical to that with the single pBAD-GFPuv plasmid (repression and activation by AraC are conserved, and the apparent affinity for L-arabinose is unchanged).

Deletion of the O_2 Site of P_{BAD}. Reporter plasmid pPCC446 is identical to pPCC442 except the O2 half-site (sequence GAAACCAATTGTCCATA, as reported by Seabold and Schleif)[31] has been deleted. To accomplish this, pPCC442 was PCR-amplified with primers O2-del-for and O2-del-rev. These primers flank the O2 half-site such that the entire pPCC442 plasmid except for the O2 site is amplified. The product was digested with *Afl*II and self-ligated, resulting in pPCC446.

Strain Construction. Strain HF19 was constructed by deleting the *araC* gene from *E. coli* strain BW27786.[32] The *araC* deletion was accomplished via P1 phage transduction using a phage library of strain JW0063 genomic DNA (*araC* is replaced by the FRT-*kan*-FRT sequence in JW0063).[33] The correct genotype of a kanamycin-resistant phage transductant was verified using PCR, and the FRT-flanked *kan* gene was subsequently removed using Flipase-mediated recombination as described.[34] The resulting strain HF19 was used in library screening and analysis of AraC mutants.

Strain HF24 was derived from HF19 by deleting the *E. coli fucPIKUR* operon and integrating the *fucP* permease gene under control of a *tac* promoter. The *fucPIKUR* operon is responsible

(20) Doyle, M. E.; Brown, C.; Hogg, R. W.; Helling, R. B. *J. Bacteriol.* **1972**, *110*, 56–65.
(21) Ghosh, M.; Schleif, R. F. *Anal. Biochem.* **2001**, *295*, 107–112.
(22) Reed, W. L.; Schleif, R. F. *J. Mol. Biol.* **1999**, *294*, 417–425.
(23) Ross, J. J.; Gryczynski, U.; Schleif, R. *J. Mol. Biol.* **2003**, *328*, 85–93.
(24) Saviola, B.; Seabold, R.; Schleif, R. F. *J. Mol. Biol.* **1998**, *278*, 539–548.
(25) Soisson, S. M.; MacDougall-Shackleton, B.; Schleif, R.; Wolberger, C. *Science* **1997**, *276*, 421–425.
(26) Wu, M.; Schleif, R. *J. Mol. Biol.* **2001**, *307*, 1001–1009.
(27) Wu, M.; Schleif, R. *J. Biol. Chem.* **2001**, *276*, 2562–2564.
(28) van Sint Fiet, S.; van Beilen, J. B.; Witholt, B. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 1693–1698.
(29) Crameri, A.; Whitehorn, E. A.; Tate, E.; Stemmer, W. P. *Nat. Biotechnol.* **1996**, *14*, 315–319.
(30) Phillips, G. J.; Park, S. K.; Huber, D. *BioTechniques* **2000**, *28*, 400–406.
(31) Seabold, R. R.; Schleif, R. F. *J. Mol. Biol.* **1998**, *278*, 529–538.
(32) Khlebnikov, A.; Datsenko, K. A.; Skaug, T.; Wanner, B. L.; Keasling, J. D. *Microbiology* **2001**, *147*, 3241–3247.
(33) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H *Mol. Syst. Biol.* **2006**, *2*, 2006.0008.
(34) Datsenko, K. A.; Wanner, B. L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 6640–6645.

for L-fucose metabolism in *E. coli* K-12, and these genes contribute to D-arabinose metabolism in *E. coli* mutants capable of growing on this substrate.[35] The operon was deleted using lambda Red technology, following a standard protocol[34] with primers *fucP*-KO-for and *fucP*-KO-rev. The resulting strain is HF23. Chromosomal integration of *fucP* was accomplished using the CRIM method.[36] *fucP* was cloned from BW27786 using primers *fucP*-for and *fucP*-rev. The PCR product was digested with *Nde*I and *Bam*HI and ligated into CRIM plasmid pAH55, yielding a construct in which *fucP* is under control of a *tac* promoter, adjacent to the *kan* resistance gene, and flanked by regions homologous to the *E. coli* λ integration site. This construct was integrated into the HF23 chromosome using helper plasmid pINT-ts as described[36] and integration was verified by PCR.

Library Construction. Overlap extension PCR[37] was performed for AraC saturation library construction as follows:

1. Saturation Library 1 ("Lib1"). Three parallel PCR reactions were performed to amplify three *araC* segments (L1A, L1B, L1C) using the following three sets of primers: *araC*-P8-for and *araC*-T24-rev; *araC*-comp-T24-for and *araC*-H80-Y82-rev; and wt-for-III and *araC*-rev-4. The first two amplified segments above correspond to site saturation at residue positions 8, 24, 80, and 82.

2. Saturation Library 2 ("Lib2"). Three parallel PCR reactions were performed to amplify three *araC* segments (L2A, L2B, L2C) using the following three sets of primers: *araC*-for-I and *araC*-T24-rev; *araC*-comp-T24-for and *araC*-H80-rev; and wt-for-III and *araC*-rev-4, respectively. The first two segments above correspond to site saturation at residue positions 8, 15, 24, and 80.

PCR products were gel-purified and equimolar aliquots of every pair of adjacent DNA fragments (1.6 nmol each) were combined (L1A + L1B and L1B + L1C; L2A + L2B and L2B + L2C) and PCR-assembled without primers. These assemblies resulted in PCR products L1AB and L1BC (for Lib1) and L2AB and L2BC (for Lib2). The two fragments for each respective library were PCR-assembled. Finally, outer primers wt-for-I and *araC*-rev-4 were added to each assembly reaction and the products were PCR-amplified.

Gene libraries were ligated into pPCC423 after digestion with *Nde*I and *Hind*III. Ligation products were transformed into *E. coli* DH10B (2×10^6 unique transformants were recovered from each library), and the plasmid library was prepared after amplification in the presence of antibiotic. Ten randomly picked clones from each library were sequenced, and these sequences revealed the expected random mutations at the targeted nucleotide positions, with no additional point mutations.

3. Error-Prone PCR Library. A third AraC library composed of random point mutations in the N-terminal (ligand binding) domain was generated using error-prone PCR.[38] Primers *arac*-for1 and *araC*-rev-4 were used to amplify the AraC fragment from pPCC423 (used as template), and PCR was carried out using GeneChoice Taq polymerase. Seven randomly picked clones from this library were sequenced and contained an average of 3.8 nucleotide mutations per clone.

(35) LeBlanc, D. J.; Mortlock, R. P. *J. Bacteriol.* **1971**, *106*, 90–96.
(36) Haldimann, A.; Wanner, B. L. *J. Bacteriol.* **2001**, *183*, 6384–6393.
(37) An, Y.; Ji, J.; Wu, W.; Lv, A.; Huang, R.; Wei, Y. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 774–778.
(38) Cirino, P. C.; Mayer, K. M.; Umeno, D. *Methods Mol. Biol.* **2003**, *231*, 3–9.

Fluorescence-Based Positive and Negative Screening. Strain HF19 does not express a D-arabinose transporter, so to ensure sufficient transport of this sugar (via nonspecific uptake by other transporters) screening was performed in the presence of 100 mM D-arabinose. AraC plasmid libraries were transformed into HF19 carrying reporter plasmid pPCC442 (2.5×10^8 transformants). Cells were prepared for screening by preculturing overnight in LB medium containing chloramphenicol and apramycin and 0.4 mM IPTG, followed by dilution to $OD_{600} = 0.2$ in the same medium containing appropriate inducer. Induced cells were then grown for 15 h. Fluorescence-activated cell-sorting (FACS) was performed on an inFlux V-GS Cytometry Workbench (Cytopeia) using Spigot software. Fluorescence was excited at 488 nm, and emission was collected using a 531/40 nm filter. In the first round of screening, the most fluorescent 10^4 cells were sorted from a total of 10^7 cells (i.e., the top 0.1% were selected). Flow cytometry analysis was performed on an FC500 flow cytometer (Beckman-Coulter). Flow cytometry of libraries resulting from the first round of positive screening revealed two subpopulations of cells: a majority were highly fluorescent in the absence of any inducer (constitutive or nonspecific phenotypes) and a smaller fraction were significantly less fluorescent in the absence of D-arabinose. The low-fluorescent cells were collected and subjected to another negative screen in the presence of L-arabinose (10 mM) to eliminate clones that were still induced by this isomer. This procedure was repeated in a second round of positive and negative FACS screening except the top 1% of cells were selected in the positive screen, yielding at most 100 different mutants induced by D-arabinose and not by L-arabinose. Ten clones from each library were selected for rescreening. The mutations identified in the five most highly induced clones are provided in the Supporting Information.

GFP Expression Fluorescence Assays. Cells (HF19 or HF24) harboring reporter plasmid (pPCC442) and AraC expression plasmid (pPCC423) were grown overnight in LB medium containing chloramphenicol and apramycin and 0.4 mM IPTG, then diluted to $OD_{600} = 0.01$ in the same medium containing an appropriate concentration of inducer, and allowed to grow under inducing conditions for 10 h. A total of 100 μL of culture was centrifuged, and the cells were washed with 10 mM potassium phosphate buffer (pH 7.5) and resuspended in 200 μL of the same buffer. The cell suspension optical density (OD_{600}) was measured with a SPECTRAmax microplate spectrophotometer (Molecular Devices Corporation), and fluorescence emission was measured with a GENios FL fluorescence spectrometer (Tecan Austria GmbH) (360 nm excitation filter, 536/50 nm emission filter). The data were normalized with respect to optical density (OD_{600}). The background fluorescence due to buffer served as the blank in all measurements. All reported data represent the mean of at least three independent data points. Error bars in Figure 3 represent standard deviations. The coefficient of variation (CV) for data reported in Table 1 was always less than 15% and less than 10% for most reported values.

Results and Discussion

The relative orientations of hydroxyl groups in α- and β-D-arabinopyranose differ significantly from those in α-L-arabinopyranose (the anomer bound by WT AraC[25]) (Figure 2a), making it difficult to rationally design D-arabinose binding in AraC via specific point mutations. Analysis of the binding pocket of wild-type AraC ("WT") in complex with L-arabinose

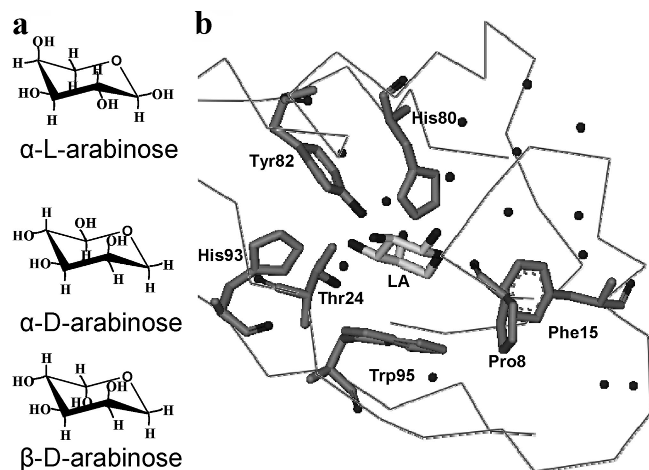


Figure 2. (a) Structures of α-L-arabinose and α- and β-D-arabinose. (b) Crystal structure of wild-type AraC binding pocket in complex with L-arabinose (LA).[25] Included are residues that play important roles in ligand binding. Water molecules are shown as spheres.

Table 1. Fluorescence of Strain HF19 Harboring P_{BAD}-GFP Reporter Plasmids and Expressing WT, Mut1, or Mut2, in the Presence of the Indicated Concentration of Effector[a]

	WT	Mut1	Mut2
no effector	170	2200	270
Δ*O2*, no effector	470	2800	930
100 mM D-arabinose	210	37000	16000
10 mM L-arabinose	66000	2600	310
100 mM L-arabinose	58000	2500	340
10 mM D-xylose (Δ*xylA*)	340	2500	450
100 mM D-xylose (Δ*xylA*)	360	2700	490
100 mM D-fucose	200	2500	350
100 mM D-lyxose	880	2900	340

[a] Values represent total cell suspension fluorescence normalized with respect to cell optical density (no correction for background fluorescence). All cells were induced for 10 h. Fluorescence of control cells not harboring GFP reporter plasmids is ∼140. Δ*O2* represents cells carrying reporter plasmid with deleted O2 half-site. Strain HF25 (HF19, Δ*xylA*) cannot metabolize xylose and was used for measuring induction by D-xylose.

(depicted in Figure 2b) combined with mechanistic insights provided by Schleif's studies and results from a previous computational AraC design study[39] resulted in the choice of two sets of binding pocket residues to target for saturation mutagenesis. To ensure exhaustive library cloning and FACS screening, a targeted library size of $\sim 10^6$ mutants (at the nucleotide level) was chosen, corresponding to the simultaneous saturation of four different residue positions. One library (Lib1) consisted of all possible mutations at positions Pro8, Thr24, His80, and Tyr82, while residues Pro8, Phe15, Thr24, and His80 were randomized in the second library (Lib2).

araC gene libraries were expressed in *E. coli* strain HF19 (Δ*araC*, Δ*araBAD*, Δ*araFGH*) harboring a P_{BAD}-GFP reporter plasmid and FACS-screened for high fluorescence in the presence of D-arabinose and low fluorescence in the absence of inducer or in the presence of L-arabinose (10^7 clones were screened per library). Neither sugar was metabolized by strains used in this study. Following two rounds of dual screening, five clones from each saturation mutagenesis library were selected for more detailed analysis. Mutants from Lib1 showed elevated

(39) Fazelinia, H.; Cirino, P. C.; Maranas, C. D. *Biophys. J.* **2007**, *92*, 2120–2130.

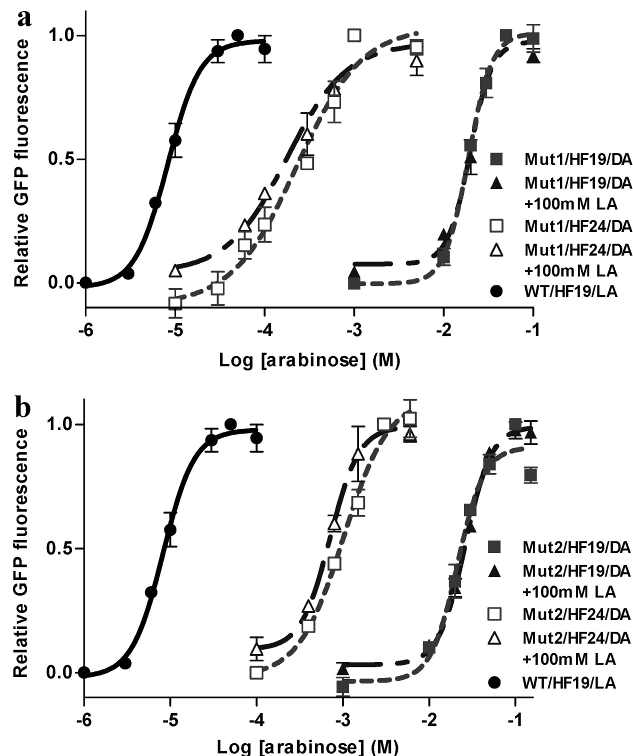


Figure 3. Relative GFP expression from strains (HF19 or HF24) co-expressing Mut1 (a), or Mut2 (b), reported as a function of D-arabinose (DA) or L-arabinose (LA) concentration. The response curve for WT AraC is included for reference. Strain HF24 constitutively expresses the FucP transporter to mediate D-arabinose uptake. Data represent cell suspension fluorescence normalized w.r.t. cell density, corrected by subtracting background fluorescence in the absence of inducer (constitutivity is therefore not portrayed) and divided by the fluorescence value under the corresponding saturating condition.

basal expression of GFP in the absence of inducer and also higher total expression upon full induction, compared to Lib2 mutants. Two out of five variants from Lib1 carried identical mutations (mutant "Mut1": P8R, T24D, H80L, Y82Q), while the other three sequences showed similar mutations (Supporting Information). Sequencing the five variants from Lib2 revealed two different mutants: "Mut2" (P8G, F15W, T24P, H80A) and "Mut3" (P8W, F15F, T24R, H80H).

A similar number of mutants ($\sim 10^7$) from a third library consisting of random point mutations in the AraC N-terminal domain was also subjected to dual FACS-based screening. While the random point mutation library also yielded variants responsive to D-arabinose and not L-arabinose, none of the recovered clones conferred inducibility to the same extent as Mut1 or Mut2 (the maximum fluorescence of fully induced cells was at least 2-fold lower). These mutants were not further characterized in this study. Note, however, that starting from any of the improved mutants as the parent gene, further rounds of random mutagenesis and screening are likely to yield variants with further improvements in D-arabinose sensitivity and/or reduced basal expression.

Here we present results for Mut1- and Mut2-controlled GFP expression (measured as whole-cell fluorescence) from P_{BAD} under various induction conditions. Table 1 reports the normalized fluorescence values for HF19 cells harboring reporter plasmid and expressing WT, Mut1, or Mut2, following incubation with various concentrations of effector. Whereas cells expressing WT AraC are fully induced in 10 mM L-arabinose,

and are not responsive to 100 mM D-arabinose, both Mut1 and Mut2 are induced in 100 mM D-arabinose and not in 10 mM or 100 mM L-arabinose. A high level of basal expression is evident for Mut1 (>10-fold higher background fluorescence relative to WT). Mut2 shows low basal expression (similar to WT) that is not elevated in the presence of L-arabinose, and a maximum induction response to D-arabinose that is ~25% of the maximum WT response to L-arabinose.

The dual selection employed here only guards against constitutivity and induction by L-arabinose. Therefore, it was possible that Mut1 or Mut2 are actually mutants with relaxed substrate specificity, with the exception of L-arabinose. The induction response in the presence of other sugars was therefore also studied. Of primary importance is the growth substrate D-xylose. To prevent D-xylose metabolism during the induction study (resulting in sugar depletion and acid secretion), the *xylA* gene encoding xylose isomerase was first deleted from HF19 (resulting in strain HF25). Table 1 shows that neither 10 mM nor 100 mM D-xylose significantly induces Mut1 or Mut2 (note that xylose induces the expression of xylose transporters in *E. coli*, and xylose transport is saturated above 1 mM[40]). D-Fucose and the pentose D-lyxose were also tested, and 100 mM concentrations of these sugars did not significantly induce either mutant (Table 1). While it is likely that some compounds other than D-arabinose can act as inducers of Mut1 and Mut2, these results indicate that the mutants are not nonspecifically induced by monosaccharides other than L-arabinose.

To study the role of DNA looping via interactions with the O2 half-site upstream of P_{BAD}, we constructed a variant of the GFP reporter plasmid in which the O2 half-site has been removed. As shown in Table 1 (labeled "Δ*O2*"), the increase in background fluorescence as a result of the O2 deletion is similar with all three proteins. The relative increase in basal expression from P_{BAD} as a result of deleting O2 is comparable to that previously reported for WT AraC.[31] The similarity in response between WT and Mut2 suggests that Mut2 retains interactions with O2. The increase in basal expression with the O2 deletion for both WT and Mut2 does not approach that observed for Mut1 in the presence of O2, suggesting that the high basal expression observed with Mut1 is not strictly due to a loss of contacts with O2. Rather, Mut1 may slightly activate transcription in spite of DNA looping, or in the absence of effector, this mutant may alternate between a conformation that interacts with O2 and an activating conformation, resulting in net elevated transcription. Studies are in progress to better characterize the repression and activation properties of Mut1 and Mut2.

Figure 3 depicts relative GFP expression levels for cells expressing Mut1 (Figure 3a) or Mut2 (Figure 3b) as a function of L- or D-arabinose concentration, with the response curve for WT AraC included for reference. The half-maximal induction response for cells expressing WT AraC occurs near 10 μM L-arabinose. In contrast, HF19 expressing Mut1 or Mut2 reaches half-maximum induction at ~20 mM D-arabinose. Whereas HF19 constitutively expresses an L-arabinose symporter (AraE),[32] this strain does not express a D-arabinose-specific transporter. We reasoned that poor uptake of D-arabinose was partly responsible for the low apparent affinity for this effector and therefore constructed strain HF24, which constitutively expresses the FucP permease, known to mediate D-arabinose uptake.[41] When HF24 is used, half-maximal induction for Mut1 (0.23 mM) and Mut2 (1.0 mM) occurs in D-arabinose concentrations 1 to 2 orders of magnitude lower than those observed with HF19.

(40) Sumiya, M.; Davis, E. O.; Packman, L. C.; McDonald, T. P.; Henderson, P. J. *Recept. Channels* **1995**, *3*, 117–128.

While the data in Table 1 demonstrates that L-arabinose does not induce Mut1 or Mut2, it does not indicate whether these mutants still bind L-arabinose, in which case L-arabinose would competitively inhibit induction by D-arabinose. (This would resemble the binding of WT AraC to D-fucose, which does not act as an inducer but inhibits induction by L-arabinose.[42]) This is particularly important if D-arabinose is to induce expression in cells growing on media containing L-arabinose. We therefore tested the influence of L-arabinose on the induction response of Mut1 and Mut2 to D-arabinose. As shown in Figure 3, 100 mM L-arabinose has essentially no impact on the GFP expression profiles for both strains HF19 and HF24, indicating that L-arabinose does not competitively inhibit D-arabinose binding.

Conclusion

The AraC mutants described here allow D-arabinose-inducible gene expression from the P_{BAD} promoter. These proteins (or similar analogues) can replace WT AraC for regulating expression from P_{BAD} during growth on substrates containing D-xylose and L-arabinose. Expression of the FucP transporter enables the use of "practical" in-culture concentrations of D-arabinose as inducer. While sugars other than L-arabinose also did not significantly induce Mut1 or Mut2 in these studies, multiple decoy ligands could easily be included during future negative screening, to ensure that particular specificities are achieved. Random mutagenesis could additionally fine-tune specific properties of AraC variants, such as reducing leaky expression from Mut1 or eliminating the weak response to D-xylose.

These results demonstrate the power of dual screening for altering AraC inducer specificity, and represent steps toward the development of orthogonal AraC−promoter pairs. Given the significant difference in orientation of hydroxyl groups between L-arabinose and D-arabinose, it should be possible to engineer a variety of customized AraC mutants that are induced by specific monosaccharides, and perhaps other small molecules. The natural coupling between molecular recognition and transcriptional activation ensured by this design process makes regulatory proteins such as AraC ideal targets as molecular reporters that can be implemented *in vivo* for the design of microbes and enzymes that overproduce effectors of interest.

Acknowledgment. We thank the Penn State Center for Quantitative Cell Analysis and the National Science Foundation for financial support (CAREER Award ID 0644678).

Supporting Information Available: Additional experimental details (primer sequences, PCR temperature programs); explanation of the chosen mutagenesis strategies. This material is available free of charge via the Internet at http://pubs.acs.org.

JA7109053

(41) Bradley, S. A.; Tinsley, C. R.; Muiry, J. A.; Henderson, P. J. *Biochem. J.* **1987**, *248*, 495–500.

(42) Soisson, S. M.; MacDougall-Shackleton, B.; Schleif, R.; Wolberger, C. *J. Mol. Biol.* **1997**, *273*, 226–237.