Gene araC (RDBECOLIGNC00050)

Gene Information

Name: araC
Synonyms: ['EG10054']
Bnumber: b0064
Position: 70387 -> 71265
Size: 878 bp
Strand: forward
gc content: 52.55972696245734%
Centisome Position: 1.5164213

multifunTerms

1 - metabolism --> 1.1 - carbon utilization --> 1.1.1 - carbon compounds

2 - information transfer --> 2.2 - RNA related --> 2.2.2 - Transcription related

3 - regulation --> 3.1 - type of regulation --> 3.1.2 - transcriptional level --> 3.1.2.2 - activator

3 - regulation --> 3.1 - type of regulation --> 3.1.2 - transcriptional level --> 3.1.2.3 - repressor

3 - regulation --> 3.3 - genetic unit regulated --> 3.3.1 - operon

External Cross References

STRING ECOLIHUB ASAP ECHOBASE OU-MICROARRAY CGSC ECOCYC REFSEQ

Sequence:

```
ATGGCTGAAGCGCAAAATGATCCCCTGCTGCCGGGATACTCGTTTAACGCCCATCTGGTGGCGGGTTTAACGCCGATTGAGGCCAACGGTTATCTCGATTTTTTTATCGACCGACCGCT
GGGAATGAAAGGTTATATTCTCAATCTCACCATTCGCGGTCAGGGGGTGGTGAAAAATCAGGGACGAGAATTTGTCTGCCGACCGGGTGATATTTTGCTGTTCCCGCCAGGAGAGATTC
ATCACTACGGTCGTCATCCGGAGGCTCGCGAATGGTATCACCAGTGGGTTTACTTTCGTCCGCGCGCCTACTGGCATGAATGGCTTAACTGGCCGTCAATATTTGCCAATACGGGTTTC
TTTCGCCCGGATGAAGCGCACCAGCCGCATTTCAGCGACCTGTTTGGGCAAATCATTAACGCCGGGCAAGGGGAAGGGCGCTATTCGGAGCTGCTGGCGATAAATCTGCTTGAGCAATT
GTTACTGCGGCGCATGGAAGCGATTAACGAGTCGCTCCATCCACCGATGGATAATCGGGTACGCGAGGCTTGTCAGTACATCAGCGATCACCTGGCAGACAGCAATTTTGATATCGCCA
GCGTCGCACAGCATGTTTGCTTGTCGCCGTCGCGTCTGTCACATCTTTTCCGCCAGCAGTTAGGGATTAGCGTCTTAAGCTGGCGCGAGGACCAACGCATTAGTCAGGCGAAGCTGCTT
TTGAGCACTACCCGGATGCCTATCGCCACCGTCGGTCGCAATGTTGGTTTTGACGATCAACTCTATTTCTCGCGAGTATTTAAAAAATGCACCGGGGCCAGCCCGAGCGAGTTTCGTGC
CGGTTGTGAAGAAAAAGTGAATGATGTAGCCGTCAAGTTGTCATAA
```

Regulation Info

[promoters: 1 | regulators: 5 | Regulatory Interaction: 2]

Operon araC

Promoters	Regulators	Transcription Unit
None	[]	araC

Regulators

Type	Function	Name
regulatoryComplex	repressor	XylR-Xylose DNA-binding transcriptional dual regulator
regulatoryComplex	activator	AraC-L-arabinose DNA-binding transcriptional activator
regulatoryComplex	activator	CRP-cAMP DNA-binding transcriptional dual regulator
regulatoryComplex	repressor	DNA-binding transcriptional dual regulator AraC
regulatoryComplex	repressor	AraC-L-arabinose DNA-binding transcriptional activator

Products

DNA-binding transcriptional dual regulator AraC

anticodon: None
cellularLocations ['cytosol']
isRegulator None
isoelectricPoint None
molecularWeight 33.384
synonyms ['AraC']
type None

External Cross References

INTERPRO PFAM PRIDE DIP PDB SMR SMART PROSITE PRINTS PRODB ECOLIWIKI MODBASE UNIPROT REFSEQ ECOCYC

Notes

The "**ara**binose regulator," AraC, is a transcription factor that regulates transcription of several genes and operons involved in arabinose catabolism and transport. It coregulates with another transcriptional regulator, CRP; both are transcription factors involved in l-arabinose degradation. These regulators bind cooperatively to activate transcription of five operons related to transport, catabolism, and autoregulation of l-arabinose. Transcription of these operons is induced when *E. coli* is grown in the absence of glucose and when the physiological inducer, l-arabinose, binds to the AraC regulator. In the absence of glucose, cellular cyclic AMP levels are high and cyclic AMP forms a dimeric complex with CRP to coregulate with AraC [14] Hofmann K et al.1997 [23] Reeder T et al.1991 [10] Schleif R et al.1983 [8] Flaherty C et al.1992 [12] Schleif RF et al.1998 [1] Hendrickson W et al.1990 [4] Wilcox G et al.1984 AraC binds to five target sites in the *araB*p region. AraC binds to the less-conserved site (-42.5) with less strength; this binding occurs only in the presence of arabinose, and it is absolutely required for expression of *araB*p [13] Lobell RB et al.1990 [17] Hamilton EP et al.1988 [19] Francklyn C et al.1987 [16] Huo L et al.1986 [24] Carra JH et al.1993 AraC binding to the distal site (-123.5) has been shown to down-regulate expression of *araB*p and *araC*p [20] Huo L et al.1992 [17] Hamilton EP et al.1988 In the absence of arabinose, AraC is unable to activate *araB*p, but it regulates its own expression by repressing *araC*p and *araB*p simultaneously [13] Lobell RB et al.1990 [17] Hamilton EP et al.1988 [16] Huo L et al.1986 Arabinose triggers AraC-dependent activation of *araB*p and relieves AraC-dependent repression of *araC*p [17] Hamilton EP et al.1988 [16] Huo L et al.1986 The *araBAD* operon is located upstream of *araC* and in the opposite direction. In the presence of arabinose, this regulator activates transcription by overlapping the -35 box of the core promoters, and the central position of the binding site is located near bp -41.5. The binding targets for AraC consist of 17-nucleotide-long direct repeat sequences that possess conserved motifs; each monomer binds to one of these conserved sequences [14] Hofmann K et al.1997 [18] Huhne R et al.1996 Studies by A. Tischer et al. suggest that when arabinose binds AraC, a helix-capping motif is formed at the dimerization domain in a region close to the interdomain linker [15] Auton M et al.2019 This capping motif appears to induce destabilization of the amide hydrogen-bonded structure of two residues in the linker, which could allow the optimum quaternary rearrangement of the DNA-binding domain that leads to an inducing state of AraC [15] Auton M et al.2019 The AraC regulator belongs to the AraC/XylS family and occurs as both a monomer and a homodimer. It is composed of two domains. The solution structure of the C-terminal DNA-binding domain has been solved. It consists of two helix-turn-helix regions connected by an α-helix [5] Rodgers ME et al.2009 The N-terminal domain is responsible for dimerization and L-arabinose binding [14] Hofmann K et al.1997 [3] Ramos JL et al.1993 A new domain, between the dimerization domain and DNA-binding domains of the dimeric AraC protein, facilitates repression of the *araBAD* operon by AraC in the absence of arabinose [11] Cole SD et al.2012 Its crystal structure [6] Wolberger C et al.1997 [9] Wolberger C et al.1997 | reveals that the sugar molecule is bound within a β-barrel, buried by the N-terminal arm of the protein. It has been suggested that this N-terminal arm plays a key role in the regulation of the arabinose-dependent DNA-binding properties of the protein. In the absence of arabinose it interacts with the DNA-binding domain and constrains this domain, and it releases it in the presence of arabinose [9] Wolberger C et al.1997 [21] Dirla S et al.2009 This interaction appears to be affected by a mutation in the interdomain linker [7] Seedorff J et al.2011 On the other hand, the function of AraC is largely unaffected by several other mutations in the interdomain linker [22] Schleif RF et al.2016

Sequence

```
MAEAQNDPLLPGYSFNAHLVAGLTPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQGREFVCRPGDILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSI
FANTGFFRPDEAHQPHFSDLFGQIINAGQGEGRYSELLAINLLEQLLLRRMEAINESLHPPMDNRVREACQYISDHLADSNFDIASVAQHVCLSPSRLSHLFRQQLGISVLSW
REDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAGCEEKVNDVAVKLS
```

Gene Ontology Terms

Biological Process

GO:0005975 - carbohydrate metabolic process [25] UniProt-GOA et al.2012 **[ICWHO]**

GO:0006355 - regulation of transcription, DNA-templated [26] GOA et al.2001 **[ICWHO]**

GO:0019568 - arabinose catabolic process [27] Englesberg E et al.1959 **[IMP]** [25] UniProt-GOA et al.2012 **[ICWHO]**

GO:0045892 - negative regulation of transcription, DNA-templated [28] GOA curators et al.2016 **[ICWHO]**

Cellular Component

GO:0005737 - cytoplasm [25] UniProt-GOA et al.2012 **[ICWHO]** [29] UniProt-GOA et al.2012 **[ICWHO]**

GO:0005829 - cytosol [30] Mann M et al.2008 [IDA] [31] Babu M et al.2009 **[ICWHO]**

GO:0032993 - protein-DNA complex [32] Wade JT et al.2014 **[IMP]**

Molecular Function

GO:0000976 - transcription cis-regulatory region binding [32] Wade JT et al.2014 **[IMP]**

GO:0001217 - DNA-binding transcription repressor activity [32] Wade JT et al.2014 **[IMP]**

GO:0003677 - DNA binding [25] UniProt-GOA et al.2012 **[ICWHO]**

GO:0003700 - DNA-binding transcription factor activity [26] GOA et al.2001 **[ICWHO]**

GO:0005515 - protein binding [33] Mosca R et al.2014 **[IPI]**

GO:0042802 - identical protein binding [33] Mosca R et al.2014 **[IPI]** [9] Wolberger C et al.1997

GO:0043565 - sequence-specific DNA binding [26] GOA et al.2001 **[ICWHO]**

Motifs

Type: Sequence-Conflict **Source:** UniProt 2015-08

Position: 7 - 7

sequence:

D

Note: None

Type: Conserved-Region **Source:** UniProt 2019-08

Position: 180 - 279

sequence:

EACQYISDHLADSNFDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG

Note: None

Type: DNA-Binding-Region **Source:** UniProt 2019-08

Position: 198 - 219

sequence:

SVAQHVCLSPSRLSHLFRQQL

Note: None

Type: DNA-Binding-Region **Source:** UniProt 2019-08

Position: 246 - 269

sequence:

ATVGRNVGFDDQLYFSRVFKKCT

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2019-08

Position: None - None

sequence:

P

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2019-08

Position: None - None

sequence:

T

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2019-08

Position: None - None

sequence:

R

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2019-08

Position: None - None

sequence:

Y

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2019-08

Position: None - None

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 31 - 31

sequence:

Y

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 209 - 209

sequence:

S

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 213 - 213

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 213 - 213

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 257 - 257

sequence:

D

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 258 - 258

sequence:

Q

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2019-08

Position: 262 - 262

sequence:

S

Note: None

Type: Pfam-Feature **Source:** Pfam 33.1

Position: 200 - 278

sequence:

AQHVCLSPSRLSHLFRQQLGISVLSWREDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRA

Note: None

Type: Pfam-Feature **Source:** Pfam 33.1

Position: 23 - 160

sequence:

TPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQGREFVCRPGDILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSIFANTGFFRPDEAHQPHFSDLFGQ
IINAGQGEGRYSELLAINLLEQLL

Note: None

Type: Conserved-Region **Source:** UniProt 2021-03

Position: 180 - 279

sequence:

EACQYISDHLADSNFDIASVAQHVCLSPSRLSHLFRQQLGISVLSWREDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRAG

Note: None

Type: DNA-Binding-Region **Source:** UniProt 2021-03

Position: 198 - 219

sequence:

SVAQHVCLSPSRLSHLFRQQL

Note: None

Type: DNA-Binding-Region **Source:** UniProt 2021-03

Position: 246 - 269

sequence:

ATVGRNVGFDDQLYFSRVFKKCT

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2021-03

Position: None - None

sequence:

P

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2021-03

Position: None - None

sequence:

T

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2021-03

Position: None - None

sequence:

R

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2021-03

Position: None - None

sequence:

Y

Note: None

Type: Amino-Acid-Sites-That-Bind **Source:** UniProt 2021-03

Position: None - None

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 31 - 31

sequence:

Y

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 209 - 209

sequence:

S

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 213 - 213

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 213 - 213

sequence:

H

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 257 - 257

sequence:

D

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 258 - 258

sequence:

Q

Note: None

Type: Mutagenesis-Variant **Source:** UniProt 2021-03

Position: 262 - 262

sequence:

S

Note: None

Type: Sequence-Conflict **Source:** UniProt 2021-03

Position: 7 - 7

sequence:

D

Note: None

Type: Pfam PF12833 **Source:** Pfam 35.0

Position: 201 - 278

sequence:

QHVCLSPSRLSHLFRQQLGISVLSWREDQRISQAKLLLSTTRMPIATVGRNVGFDDQLYFSRVFKKCTGASPSEFRA

Note: None

Type: Pfam PF02311 **Source:** Pfam 35.0

Position: 23 - 160

sequence:

TPIEANGYLDFFIDRPLGMKGYILNLTIRGQGVVKNQGREFVCRPGDILLFPPGEIHHYGRHPEAREWYHQWVYFRPRAYWHEWLNWPSIFANTGFFRPDEAHQPHFSDLFGQ
IINAGQGEGRYSELLAINLLEQLL

Note: None

Citations

[1] Hendrickson W, Schleif R, Stoner C, 1990, Characterization of the Escherichia coli araFGH and araJ promoters..
[2] Johnson CM, Schleif RF, 1995, In vivo induction kinetics of the arabinose promoters in Escherichia coli..
[3] Ramos JL, Michan C, Gallegos MT, 1993, The XylS/AraC family of regulators..
[4] Wilcox G, Stoltzfus L, Miyada CG, 1984, Regulation of the araC gene of Escherichia coli: catabolite repression, autoregulation, and effect on araBAD expression..
[5] Rodgers ME, Schleif R, 2009, Solution structure of the DNA binding domain of AraC protein..
[6] Wolberger C, MacDougall-Shackleton B, Soisson SM, Schleif R, 1997, The 1.6 A crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose..
[7] Seedorff J, Schleif R, 2011, Active role of the interdomain linker of AraC..
[8] Flaherty C, Hendrickson W, Molz L, 1992, Sequence elements in the Escherichia coli araFGH promoter..
[9] Wolberger C, MacDougall-Shackleton B, Soisson SM, Schleif R, 1997, Structural basis for ligand-regulated oligomerization of AraC..
[10] Schleif R, Stoner C, 1983, The araE low affinity L-arabinose transport promoter. Cloning, sequence, transcription start site and DNA binding sites of regulatory proteins..
[11] Cole SD, Schleif R, 2012, A new and unexpected domain-domain interaction in the AraC protein..
[12] Schleif RF, Seabold RR, 1998, Apo-AraC actively seeks to loop..
[13] Lobell RB, Schleif RF, 1990, DNA looping and unlooping by AraC protein..
[14] Hofmann K, Bairoch A, Gallegos MT, Ramos JL, Schleif R, 1997, Arac/XylS family of transcriptional regulators..
[15] Auton M, Tischer A, Schleif RF, Brown MJ, 2019, Arabinose Alters Both Local and Distal H-D Exchange Rates in the Escherichia coli AraC Transcriptional Regulator..

[16] Huo L, Schleif RF, Martin K, 1986, The DNA loop model for ara repression: AraC protein occupies the proposed loop sites in vivo and repression-negative mutations lie in these same sites..
[17] Hamilton EP, Lee N, 1988, Three binding sites for AraC protein are required for autoregulation of araC in Escherichia coli..
[18] Huhne R, Niland P, Muller-Hill B, 1996, How AraC interacts specifically with its target DNAs..
[19] Francklyn C, Hamilton EP, Lee N, 1987, Arabinose-induced binding of AraC protein to araI2 activates the araBAD operon promoter..
[20] Huo L, Schleif R, Lee DH, 1992, Repression of the araBAD promoter from araO1..
[21] Dirla S, Rodgers ME, Holder ND, Schleif R, 2009, Functional modes of the regulatory arm of AraC..
[22] Schleif RF, Mayberry O, Rodgers ME, Toptygin D, Malaga F, Park DJ, 2016, A genetic and physical study of the interdomain linker of E. Coli AraC protein-a trans-subunit communication pathway..
[23] Reeder T, Schleif R, 1991, Mapping, sequence, and apparent lack of function of araJ, a gene of the Escherichia coli arabinose regulon..
[24] Carra JH, Schleif RF, 1993, Variation of half-site organization and DNA looping by AraC protein..
[25] UniProt-GOA, 2012, Gene Ontology annotation based on UniProtKB/Swiss-Prot keyword mapping, accompanied by conservative changes to GO terms applied by UniProt..**[ICWHO]**
[26] GOA, ZFIN, DDB, MGI, FB, 2001, Gene Ontology annotation through association of InterPro records with GO terms..**[ICWHO]**
[27] Englesberg E, Gross J, 1959, Determination of the order of mutational sites governing L-arabinose utilization in Escherichia coli B/r bv transduction with phage Plbt..**[IMP]**
[28] GOA curators, 2016, Automatic assignment of GO terms using logical inference, based on on inter-ontology links..**[ICWHO]**
[29] UniProt-GOA, 2012, Gene Ontology annotation based on UniProtKB/Swiss-Prot Subcellular Location vocabulary mapping, accompanied by conservative changes to GO terms applied by UniProt..**[ICWHO]**
[30] Mann M, Hartl FU, Rappsilber J, Schmidt T, Kerner MJ, Frishman D, Ishihama Y, 2008, Protein abundance profiling of the Escherichia coli cytosol..[IDA]
[31] Babu M, Diaz-Mejia JJ, Emili A, 2009, Computational and experimental approaches to chart the Escherichia coli cell-envelope-associated proteome and interactome..**[ICWHO]**
[32] Wade JT, Baranowski C, Reilly AA, Petrone BL, Stringer AM, Bonocora RP, Currenti S, Erill I, Zhang Z, Palumbo MJ, 2014, Genome-scale analyses of Escherichia coli and Salmonella enterica AraC reveal noncanonical targets and an expanded core regulon..**[IMP]**
[33] Mosca R, Vlasblom J, Wuchty S, Emili A, Uetz P, Pakala SB, Siszler G, Babu M, Aloy P, Ceol A, Arnold R, Rajagopala SV, Pieper R, Hauser R, Franca-Koh J, Sikorski P, Kumar A, Phanse S, 2014, The binary protein-protein interaction landscape of Escherichia coli..**[IPI]**
[34] Wolberger C, MacDougall-Shackleton B, Soisson SM, Schleif R, 1997, Structural basis for ligand-regulated oligomerization of AraC..**[IPI]**