Journal of Supramolecular Structure 6:411–417 (1977)
Molecular Aspects of Membrane Transport 273–279

L-Arabinose Transport and the L-Arabinose Binding Protein of Escherichia coli

Robert W. Hogg

Department of Microbiology, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106

The active accumulation of L-arabinose by arabinose induced cultures of Escherichia coli is mediated by 2 independent transport mechanisms. One, specified by the gene locus *araE*, is membrane bound and possesses a relatively "low affinity." The other, specified in part by the genetic locus *araF*, contains as a functional component the L-arabinose binding protein and functions with a "high affinity" for the substrate. The L-arabinose binding protein has been purified, partially characterized, crystallized, and sequenced.

Key words: L-arabinose, transport, binding proteins, sequence

L-arabinose is a pentose which can serve as the sole source of carbon and energy for the enteric bacteria Escherichia coli. The transport proteins and enzymes engaged in the utilization of arabinose are inducible (by arabinose) and regulated in a positive manner by the product of the gene *araC* (Fig. 1, Ref 1).

The active accumulation of arabinose by induced E. coli cells is mediated by 2 genetically distinct transport systems. One, the product of the gene *araE*, was described by Novotny and Englesberg (2) and will be referred to as the "low-affinity" L-arabinose transport system. The second transport system contains as a functional component the L-arabinose binding protein, the product of the gene *araF*, and represents the "high-affinity" L-arabinose transport system. *AraE*, at 61 minutes, and *araF*, tentatively located at 45 minutes on the *E. coli* chromosomal map, are both regulated in a positive manner by the *araC* gene product which is located adjacent to the contiguously linked arabinose operon *araD, A,B,I,O* at 1 minute (3, 1).

"Low-Affinity," *araE*, L-arabinose Transport

Novotny and Englesberg first described the active accumulation of L-arabinose by E. coli (2). Accumulation of arabinose was observed to be an energy dependent, inducible function with a temperature optimum of 25°C. The K^{en} (K_m of entry) for L-arabinose was determined to be 1.25×10^{-4} M and D-fucose, D-xylose, and D-galactose were found to competitively inhibit L-arabinose uptake. The accumulated arabinose was recovered

Received April 1, 1977; accepted June 6, 1977

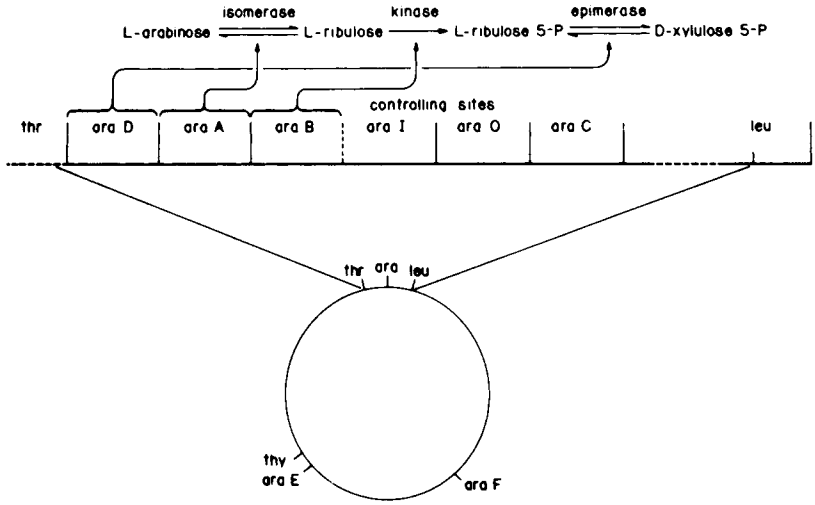

Fig. 1. The L-arabinose operon of Escherichia coli

unmodified from the cytoplasm of the cell. Mutants defective in L-arabinose transport were isolated and described by Isaacson and Englesberg (4). The transport defect was designated *araE* and was observed to cotransduce with the *thyA* locus.

We have since determined that "low-affinity" L-arabinose uptake capacity is retained by membrane vesicles prepared from induced E. coli cells (unpublished). Vesicles prepared from *araE* mutant strains do not demonstrate this capacity. Energization of this transport function appears to be via a chemiosmotic mechanism as described by Henderson (5). "Low-affinity" transport can thus be concluded to represent a membrane-bound *araE* gene product. A partial purification of this protein has been achieved using double isotope techniques and membrane solubilization with Brij 36T (unpublished).

"High-Affinity," *araF*, L-arabinose Transport

The L-arabinose binding protein was isolated from E. coli B by Hogg and Englesberg (6) and from E. coli K12 by Schleif (7). The binding protein was implicated in L-arabinose transport when it was observed that the specificity of ligand binding by the purified protein and the pattern of inhibition of binding by various arabinose analogues were identical to those of the in vivo transport process. Mutations in all genes specifying L-arabinose metabolic enzymes (L-arabinose isomerase, *araA;* L-ribulo kinase, *araB;* and L-ribulose-5-phosphate-4-epimerase, *araD*) did not alter in vivo transport or in vitro binding (6). In addition, Schleif isolated a mutant of E. coli K12 which demonstrated reduced transport capacity and also a lower content of binding protein than the parental strain (7).

The observation that *araE* mutations exerted no apparent effect on the binding protein led to a search for binding protein mutants and a reevaluation of L-arabinose uptake kinetics. Non-*crm*-producing binding protein mutants were obtained using the antibody plate assay described by Hogg (8). This technique permits rapid screening of large numbers

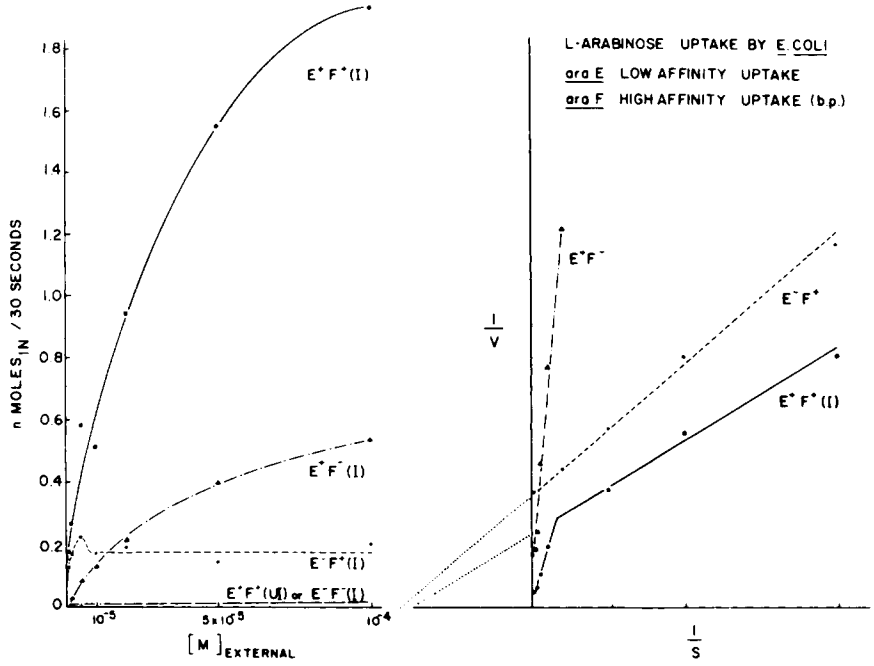

Fig. 2. The kinetics of L-arabinose transport in Escherichia coli. The reciprocal plot on the right panel represents the same data plotted over the same concentration range and units used in the figure on the left.

of suspected mutants by growing such mutants on plates containing arabinose and antibody directed against the binding protein. Subsequent lysis with chloroform and toluene releases cell contents and *crm*$^+$ or *crm*$^-$ colonies are readily distinguished after 24-h incubation at 4°C. A detailed kinetic analysis of induced cells which had been extensively washed indicated that the kinetics of L-arabinose uptake were biphasic (Fig. 2). Mutant strains defective in *araE* and/or *araF* indicate that "low-affinity" L-arabinose uptake is associated with *araE* and demonstrates a K^{en} of 1×10^{-4} M. "High-affinity" uptake was associated with *araF* and the K^{en} was determined to be 1×10^{-6} M. Transport defects could be transferred, *araE* cotransducing with thymine, *araF* cotransducing with histidine. It was also noted that the strain used by Novotny and Englesberg in their earlier studies possessed only "low-affinity," *araE*, transport. This condition probably facilitated the later isolation of *araE* mutants from the same genetic background by Isaacson and Englesberg.

The fact that the uptake of *araE*$^+$ *araF*$^-$ and *araE*$^-$ *araF*$^+$ strains is not additive to the the uptake observed in *araE*$^+$ *araF*$^+$ strains suggests that although both systems are capable of independent function, a synergistic effect occurs when both are present. Rotman has reported binding-protein-independent galactose uptake by the methylgalactoside system (9, 10); however, this uptake is low and appears restricted to components of a binding-protein-mediated complex. In our case 2 apparently independent systems are capable of interaction.

Transport studies conducted on cells subjected to osmotic shock (11) indicated that "low-affinity" transport capacity was retained by shocked cells and "high-affinity" transport functions were lost.

PURIFICATION AND PROPERTIES OF THE L-ARABINOSE BINDING PROTEIN

Cell-free extracts or osmotic shock supernatants of induced E. coli cells served as an abundant source of the L-arabinose binding protein. E. coli *araA39,* a strain containing a nonsense mutation in the isomerase gene, produces approximately 2 mg of binding protein protein per gram wet weight of cells. The procedure for purification, from cell-free extracts, involves passing the supernatant of a 50% ammonium sulfate fractionation through a 0.01 M DEAE column. The eluant of this column is concentrated and further fractionated by column chromatography on DEAE (0.001 M KPO_4, pH 7.8, gradient from 0 to 0.03 M KCl). This procedure separates the L-arabinose binding protein from the D-galactose binding protein. Arabinose binding material was concentrated and further purified by column chromatography on Sephadex G75 equilibrated and eluted with 0.01 M KPO_4, pH 7.8 (10). The purified protein can be crystallized from 2-methyl-2, 4-pentanediol. The molecular weight of the purified protein is 33,300.

Sedimentation velocity studies and optical rotatory dispersion analysis in the presence and absence of arabinose indicated that no major structural reorientation occurred when arabinose was bound. Arabinose binding did result in minor structural rearrangements as indicated by a 5-nm blue shift in the fluorescence emission spectra suggesting that some tryptophan residues enter a more hydrophobic environment when arabinose is bound. Arabinose was also observed to protect some tryptophan residues from N-bromosuccinimide oxidation (12). Sulfhydryl reactive agents were found to eliminate binding capacity; however, this inhibition was completely reversible for those agents which can be inactivated by thiols such as 2-mercaptoethanol (unpublished). Binding of arabinose was uneffected by ionic strengths between 0 and 1.5 M KCl and occurred over a pH range of 6.4–9.5.

Studies of this nature suggest that a tryptophan residue or residues is involved in or near the L-arabinose binding site and the single cysteine is required to maintain an active configuration. The insensitivity of binding to ionic strength and pH suggest a minimum role for charge interactions in the binding process; however, the in vivo environment may provide considerably different conditions.

SEQUENCE ANALYSIS OF THE L-ARABINOSE BINDING PROTEIN

The complete amino acid sequence of the L-arabinose binding protein has been determined (13). The protein contains 306 amino acid residues in a single peptide chain and is presented on Fig. 3. The sequence of the protein contains few distinguishing features. The single cysteine residue is located at position 64 and a highly charged region, -Arg-Arg-Arg-, occurs at residues 149, 150, and 151. The sequence analysis and the crystallographic analysis (18) will provide a detailed molecular model of the L-arabinose binding protein.

BINDING PROTEIN HOMOLOGY

Earlier studies by Parsons and Hogg have indicated that antibody directed against the L-arabinose binding protein of E. coli will form an immunoprecipitate with the D-

GLU - asn - leu - lys - leu - gly - phe - leu - val - lys -
gln - pro - glu - glu - pro - trp - phe - gln - thr - glu - 20
TRP - lys - phe - phe - asp - lys - ala - gly - lys - asp -
leu - gly - phe - GLU - VAL - ile - lys - ile - ala - val - 40
pro - asp - gly - GLU - LYS - thr - leu - asn - ala - ile -
asp - ser - leu - ala - ala - ser - gly - ala - lys - gly - 60
phe - val - ile - cys - thr - pro - ASP - PRO - lys - leu -
gly - ser - ala - ile - val - ala - lys - ala - ARG - gly - 80
tyr - asp - MET - lys - val - ile - ala - val - asp - asp -
gln - phe - val - asn - ala - lys - gly - lys - pro - MET - 100
asp - thr - val - pro - leu - val - MET - MET - ala - ala -
thr - lys - ile - gly - glu - ARG - gln - gly - gln - glu - 120
leu - TYR - lys - glu - MET - gln - lys - ARG - gly - TRP -
asp - val - lys - glu - ser - ala - val - MET - ala - ile - 140
thr - ala - asn - glu - leu - asp - thr - ala - ARG - ARG -
ARG - thr - thr - gly - ser - MET - asp - ala - leu - lys - 160
ala - ala - gly - phe - pro - glu - lys - gln - ile - tyr -
gln - val - pro - thr - lys - ser - asn - asp - ile - pro - 180
gly - ala - phe - asp - ala - ala - asn - ser - MET - leu -
val - gln - his - pro - glu - val - lys - his - TRP - leu - 200
ile - val - gly - MET - asn - asp - ser - thr - val - leu -
gly - gly - val - ARG - ala - thr - glu - gly - gln - gly - 220
phe - lys - ala - ala - asp - ile - ile - gly - ile - gly -
ile - ASN - GLY - val - asp - ala - val - ser - glu - leu - 240
ser - lys - ala - gln - ala - thr - gly - phe - tyr - gly -
ser - leu - leu - pro - ser - pro - asp - val - his - gly - 260
tyr - lys - ser - ser - glu - MET - leu - tyr - asn - TRP -
val - ala - lys - asp - val - glu - pro - pro - lys - phe - 280
thr - glu - val - thr - asp - val - val - leu - ile - thr -
ARG - asp - asn - phe - lys - glu - glu - leu - glu - lys - 300
lys - gly - leu - gly - gly - lys

Fig. 3. The amino acid sequence of the L-arabinose binding protein

galactose binding protein isolated from the same source (14). Using an immunocompetition assay the degree of homology was roughly quantitated and found to be approximately 20%. The fact that binding proteins function in similar environments and carry out similar processes led to the supposition that any existing homology may represent structural requirements for active function. Perhaps there exists a common or similar structure for interaction with the membrane or a membrane-bound translocating protein. Conceivably such units might even specify the periplasmic nature of these proteins.

An extensive consideration of molecular homology must await the completion of additional sequence analysis; however, a comparison of the amino terminal residues of the L-arabinose-, D-galactose-, and D-ribose-binding proteins is possible. The amino terminal sequences of the ribose- and galactose-binding proteins isolated from Salmonella typhimurium were determined by Koshland and Bradshaw (manuscript in preparation). Sequence analyses of the arabinose- and galactose-binding proteins isolated from E. coli and S. typhimurium were determined by Hogg and Hermodson (manuscript in preparation).

The data tentatively suggest that some homology exists between the 3 proteins; however, more extensive sequence analysis will be required to substantiate these possibilities.

The possibility that homology exists between the L-arabinose-binding protein and

the D-galactose-binding protein (for review of galactose-binding proteins see Ref. 15) is supported by our observation that the gene locus for the arabinose-binding protein is located at approximately 44.5 minutes on the E. coli linkage map and the gene locus for the galactose-binding protein is located at 45.5 minutes (16). The possibility of gene duplication is under consideration.

MUTANTS DEFECTIVE FOR "HIGH-AFFINITY" TRANSPORT OF L-ARABINOSE

A number of mutants defective in the "high-affinity" transport of L-arabinose have been isolated. Initially, mutants which were unable to produce a protein that would cross-react with antibody directed against the L-arabinose binding protein were obtained using the *crm* plate technique described above (8). Such mutants are of interest for genetic mapping and for the demonstration of cotransfer of loss of function and genetic lesion; however, they are of little value for biochemical studies of binding protein structure and function. An extensive search for mutants producing defective L-arabinose binding protein was undertaken using as a parental strain E. coli K12 containing a mutation at the *araE* locus to effectively eliminate "low-affinity" L-arabinose uptake. Following mutagenesis, colonies demonstrating a more arabinose-negative phenotype than the parental strain when grown on EMB arabinose or TPTC arabinose, were isolated. Mutations in the structural genes *araD*, *araA*, and *araB*, the regulatory gene *araC*, and regulatory sites *araI* and *araO* were eliminated by enzymatic assay for *araA* and *araB* gene products and by growth in the presence or absence of L-arabinose. The remaining mutants were screened for "high-affinity" L-arabinose uptake using concentrations of L-arabinose [10^{-5}M] where only binding-protein-mediated uptake would be observed.

Mutants obtained in this manner fall into 2 classes. One class has lost arabinose-binding-protein-mediated uptake though they retain D-galactose uptake (presumably mediated by the methyl galactoside uptake system). The second class appears to be defective for both L-arabinose and D-galactose uptake. D-galactose uptake was originally included to eliminate pleiotrophic mutations; however, the doubly defective mutants are of considerable interest. Amongst the *ara*$^-$, *gal*$^+$ mutations, 3 subclasses exist. The first class appears to bind L-arabinose normally in vitro, however, no in vivo uptake occurs. The second class exibits altered binding properties in vitro and does not transport in vivo. The third class does not bind arabinose in vitro and does not transport in vivo. Since binding proteins are recognized as such because they bind their respective ligands, we assume that mutants in the second and third classes consist of alterations which affect the binding site proper or the molecular configuration which insures the binding site. Some of the mutations in class one effect those regions of the molecule which are required for functions other than the recognition of substrate. These regions of the molecule could be envisioned as determining, membrane interactions, interactions with an energized translocating protein embedded in the membrane, sites involved in oligimer formation, or portions of the molecule which specify the periplasmic nature of an immature variety of the functional protein.

Mutations which result in the loss of both arabinose- and galactose-binding-protein-mediated uptake may possibly represent lesions in a common unit analogous to the *hisP* mutations described by Kustu and Ames (17).

CONCLUSIONS

The L-arabinose-binding protein and the role it plays in the active accumulation of L-arabinose by Escherichia coli is still under consideration.

ACKNOWLEDGMENTS

RWH was supported by United States Public Health Service Grant 1-R01-AM-13791 from the National Institute of Arthritis and Metabolic Diseases.

REFERENCES

1. Englesberg E, Irr J, Power J, Lee N: J Bacteriol 90:946, 1965.
2. Novotny C, Englesberg E: Biochem Biophys Acta 117:217, 1966.
3. Brown C, Hogg RW: J Bacteriol 111:606, 1972.
4. Isaacson D, Englesberg E: Bacteriol Proc 113:114, 1964.
5. Henderson PT, Skinner A: Biochem Soc Trans 2:543, 1974.
6. Hogg R, Englesberg E: J Bacteriol 100:423, 1969.
7. Schleif R: J Mol Biol 46:185, 1969.
8. Hogg RW: J Bacteriol 105:604, 1971.
9. Robbins A, Rotman B: Proc Natl Acad Sci USA 72:423, 1975.
10. Robbins A, Guzman R, Rotman B: J Biol Chem 251:3112, 1976.
11. Neu H, Heppel L: J Biol Chem 240:3685, 1965.
12. Parsons R, Hogg RW: J Biol Chem 249:3602 1974.
13. Hogg RW, Hermodson M: J Biol Chem 252:5135–5141, 1977.
14. Parsons R, Hogg RW: J Biol Chem 249:3608, 1974.
15. Boos W: Annu Rev Biochem 43:123, 1974.
16. Boos W: J Biol Chem 247:5414, 1972.
17. Kustu SG, Ames GF: J Bacteriol 116:107, 1973.
18. Quiocho F: J Supramol Struct 6:503, 1977.