11-15-2009

Identification Of Functional LsrB-Like Autoinducer-2 Receptors

C. S. Pereira

Anna Katherine De Regt, '09

P. H. Brito

Stephen T. Miller
Swarthmore College, smiller1@swarthmore.edu

K. B. Xavier

Follow this and additional works at: http://works.swarthmore.edu/fac-chemistry

Part of the Biochemistry Commons

Recommended Citation
http://works.swarthmore.edu/fac-chemistry/52

This Article is brought to you for free and open access by the Chemistry & Biochemistry at Works. It has been accepted for inclusion in Chemistry & Biochemistry Faculty Works by an authorized administrator of Works. For more information, please contact myworks@swarthmore.edu.
Identification of Functional LsrB-Like Autoinducer-2 Receptors

Catarina S. Pereira, Anna K. de Regt, Patrícia H. Brito, Stephen T. Miller and Karina B. Xavier


Updated information and services can be found at:
http://jb.asm.org/content/191/22/6975

These include:
SUPPLEMENTAL MATERIAL
Supplemental material

REFERENCES
This article cites 65 articles, 29 of which can be accessed free at: http://jb.asm.org/content/191/22/6975#ref-list-1

CONTENT ALERTS
Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»
Identification of Functional LsrB-Like Autoinducer-2 Receptors\textsuperscript{\textdagger,\textdaggerdbl}

Catarina S. Pereira,\textsuperscript{1,2} Anna K. de Regt,\textsuperscript{3} Patrícia H. Brito,\textsuperscript{1} Stephen T. Miller,\textsuperscript{3,\textdagger} and Karina B. Xavier\textsuperscript{1,2,\textdaggerdbl}

Instituto Gulbenkian de Ciência, 2781-901 Oeiras, Portugal; Instituto de Tecnologia Química e Biológica, 2781-901 Oeiras, Portugal; and Department of Chemistry and Biochemistry, Swarthmore College, Swarthmore, Pennsylvania 19081

Received 24 July 2009/Accepted 1 September 2009

Although a variety of bacterial species have been reported to use the interspecies communication signal autoinducer-2 (AI-2) to regulate multiple behaviors, the molecular mechanisms of AI-2 recognition and signal transduction remain poorly understood. To date, two types of AI-2 receptors have been identified: LuxP, present in Vibrio spp., and LsrB, first identified in Salmonella enterica serovar Typhimurium. In S. Typhimurium, LsrB is the ligand binding protein of a transport system that enables the internalization of AI-2. Here, using both sequence analysis and structure prediction, we establish a set of criteria for identifying functional AI-2 receptors. We test our predictions experimentally, assaying key species for their abilities to import AI-2 in vivo, and test their LsrB orthologs for AI-2 binding in vitro. Using these experimental approaches, we were able to identify AI-2 receptors in organisms belonging to phylogenetically distinct families such as the Enterobacteriaceae, Rhizobiaceae, and Bacillaceae. Phylogenetic analysis of LsrB orthologs indicates that this pattern could result from one single origin of the functional LsrB gene in a gammaproteobacterium, suggesting possible posterior independent events of lateral gene transfer to the Alphaproteobacteria and Firmicutes. Finally, we used mutagenesis to show that two AI-2-interacting residues are essential for the AI-2 binding ability. These two residues are conserved in the binding sites of all the functional AI-2 binding proteins but not in the non-AI-2-binding orthologs. Together, these results strongly support our ability to identify functional LsrB-type AI-2 receptors, an important step in investigations of this interspecies signal.

Supplemental material for this article may be found at http://jb.asm.org/.

\textsuperscript{\textdagger} Corresponding author. Mailing address for Karina B. Xavier: Instituto Gulbenkian de Ciência, 2781-901 Oeiras, Portugal. Phone: (351) 21 446 4655. Fax: (351) 21 440 7970. E-mail: kxavier@igc.gulbenkian.pt. Mailing address for Stephen T. Miller: Department of Chemistry and Biochemistry, Swarthmore College, Swarthmore, PA 19081. Phone: (610) 957-6063. Fax: (610) 328-7355. E-mail: smiller1@sawthmore.edu.

\textsuperscript{\textdaggerdbl} Published ahead of print on 11 September 2009.
Typhimurium LsrB AI-2 receptors is only approximately 11%). The function of the LsrB protein has been characterized for the two closely related enteric bacteria, S. Typhimurium (56, 57) and \textit{Escherichia coli} (65), the plant symbiont \textit{S. meliloti} (37), and the oral pathogen \textit{Aggregatibacter (Actinobacillus) actinomycetemcomitans} (48). In all these organisms it is thought that LsrB acts as the substrate binding protein of an ATP binding cassette (ABC) transport system responsible for AI-2 internalization. Due to the homology with other ABC transport systems, it is predicted that the Lsr transporter is composed of LsrB, two transmembrane proteins (LsrC and LsrD) that form a channel, and a cytoplasmic protein (LsrA) that contains an ABC binding motif and is thought to be responsible for ATP hydrolysis during transport. Once inside the cell, AI-2 is phosphorylated by the kinase LsrK and further processed by the enzymes LsrG and LsrF (56, 66). The genes encoding these proteins (with the exception of LsrK) are all in the same operon, which is regulated by the repressor LsrR. In the absence of phospho-AI-2 (P-AI-2), LsrR represses the transcription of the \textit{lsr} operon; however, when AI-2 is internalized and phosphorylated by LsrK, P-AI-2 binds LsrR, causing the derepression of the operon. Thus, the increased expression of the \textit{lsr} system leads to increased AI-2 import, resulting in a rapid depletion of AI-2 from the extracellular medium.

It does not appear that the AI-2 taken up by this system is used as a carbon source, since cultures of \textit{S. Typhimurium} and \textit{S. meliloti} were unable to grow when AI-2 was used as the sole carbon source (37, 57). Rather, AI-2 removal via the \textit{lsr} system enables these organisms to terminate their own AI-2 signaling system and to regulate the AI-2-dependent gene expression of other organisms in the vicinity. Thus, in cultures composed of different species, bacteria with a functional Lsr system are capable of interfering with the AI-2-mediated group behaviors of the other species (63).

Recently, two studies have undertaken database sequence analysis to identify LsrB orthologs (41, 50). Those studies showed that orthologs of the Lsr system are not broadly conserved across the kingdom \textit{Bacteria} while identifying hypothetical LsrB receptors in some organisms belonging to very distant families such as the \textit{Enterobacteriaceae}, \textit{Pasteurellaceae}, \textit{Rhizobiacae}, \textit{Rhodobacteriaceae}, and \textit{Bacillaceae}.

Here, we expand upon previous bioinformatic studies (41, 50) with additional analysis based not only on sequence but also on structure prediction that allow us to establish a set of criteria for predicting which orthologs of LsrB are functional AI-2 receptors. We then present experimental evidence that confirms a set of these predictions and demonstrates the presence of functional AI-2 receptors in the families \textit{Enterobacteriaceae}, \textit{Rhizobiacae}, and \textit{Bacillaceae}.

### MATERIALS AND METHODS

**Bacterial strains and growth conditions.** The strains used are listed in Table 1. Bacteria from the family \textit{Enterobacteriaceae} (\textit{E. coli} MG1655 and uropathogenic \textit{E. coli} [UPEC] UT89) and the family \textit{Bacillaceae} (\textit{Bacillus cereus} ATCC 10987 and vaccine strain \textit{Bacillus anthracis} Sterne 34F2) were grown in Luria-Bertani (LB) medium with shaking at 37°C. The bacteria from the family \textit{Rhizobiacae} (\textit{S. meliloti} Rm1021, \textit{Agrobacterium tumefaciens} C58, \textit{Rhizobium etli} C58, and \textit{Rhizobium leguminosarum} bv. \textit{viciae} 3841) were cultured with shaking at 30°C in their optimal cultured medium, LBMC (LB medium supplemented with 2.5 mM MgSO\textsubscript{4} \text{and} 2.5 mM CaCl\textsubscript{2}), LB medium, YEM (10 g liter\textsuperscript{-1} mannitol, 0.5 g liter\textsuperscript{-1} yeast extract, 0.2 g liter\textsuperscript{-1} MgSO\textsubscript{4}, 7H\text{O}, and 1 g liter\textsuperscript{-1} NaCl), and TYC (5 g liter\textsuperscript{-1} tryptone, 3 g liter\textsuperscript{-1} yeast extract, 0.5 g liter\textsuperscript{-1} CaCl\textsubscript{2}), respectively.

### Table 1. Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Strain Source and/or reference</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Salmonella Typhimurium</td>
<td>ATCC 14028</td>
</tr>
<tr>
<td>\textit{Escherichia coli} K-12 MG1655</td>
<td>7</td>
</tr>
<tr>
<td>\textit{Escherichia coli} UT89 (UPEC)</td>
<td>Jeffrey I. Gordon (40)</td>
</tr>
<tr>
<td>\textit{Bacillus anthracis} Sterne 34F2</td>
<td>(vaccine strain)</td>
</tr>
<tr>
<td>\textit{Bacillus cereus} ATCC 10987</td>
<td>Adriano O. Henriques</td>
</tr>
<tr>
<td>Sinorhizobium meliloti Rm1021</td>
<td>29</td>
</tr>
<tr>
<td>\textit{Agrobacterium tumefaciens} C58</td>
<td>James P. Shapleigh (3)</td>
</tr>
<tr>
<td>\textit{Rhizobium leguminosarum} bv. \textit{viciae} 3841</td>
<td>Gladys Alexandre (30)</td>
</tr>
<tr>
<td>\textit{Rhizobium etli} CFN42</td>
<td>ATCC 5125</td>
</tr>
</tbody>
</table>

### Protein expression and purification.

The genes encoding LsrB orthologs in \textit{R. etli}, \textit{R. leguminosarum}, \textit{A. tumefaciens}, \textit{E. coli} MG1655, and \textit{E. coli} UT89 were cloned from genomic DNA into plasmid pProEX HTB for expression as poly-histidine-tagged proteins. The \textit{B. anthracis} LsrB ortholog was cloned into plasmid pET15/p-TOPO using the Champion pET Directional TOPO expression kit (Invitrogen) for expression as a poly-histidine-tagged fusion protein. N-terminalsequence peptides for secretion, as determined by the program SignalP 3.0 (6), were excluded from the constructs. Plasmids were transformed into \textit{E. coli} strains BL21 and FED101 (BL21 Lee' null mutant), and expression was induced with 0.1 mM isopropyl-\textbeta-D-1-thiogalactopyranoside (IPTG) when the cultures reached an OD\text{max} of 0.9. The bacteria were harvested after expression for 5 h at 22°C. Pellets were resuspended in a solution containing 50 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 8.0), 300 mM NaCl, 10 mM imidazole, 1 mM diithiothreitol, 0.36 mg/ml leupeptin, 0.36 mg/ml aprotonin, and 0.36 mg/ml DNase and lysed using an M-110Y microfluidizer (Microfluidics). The lysate was centrifuged, and the tagged protein was purified using Ni-nitrotriacetic acid affinity chromatography (Qiagen). The protein was eluted from the column using a solution containing 50 mM NaH\textsubscript{2}PO\textsubscript{4} (pH 8.0), 300 mM NaCl, and 250 mM imidazole, and the buffer was then
swapped using Sephadex-G25 agarose into a solution containing 50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, and 1 mM dithiothreitol. Purified protein was concentrated to 10 mg ml\(^{-1}\). The genes encoding the S. Typhimurium and B. cereus LsrB orthologs were cloned into plGEX-4T1, transformed, expressed, and purified as described previously (31, 37). The primers used for the cloning of the respective genes are listed in Table 2.

**AI-2 binding assay.** Proteins tested for AI-2 binding were denatured (70°C for 10 min) to release any bound ligand and pelleted (12). As previously described (4, 5), the mutations D171N and A227T were introduced into two separate B. anthracis/pET151 constructs using the QuikChange Lightning site-directed mutagenesis kit (Strategene). Primers used for creating the mutations are given in Table 2. The same kit was then used to create the double mutant D171N A227T. The mutant proteins were expressed and purified as described above for the B. anthracis wild-type LsrB ortholog.

**Phylogenetic analyses.** The evolutionary history of the lsrB gene was studied by analyzing the phylogenetic relationship of the functional orthologs identified in this study and contrasting it with the phylogeny of rpoB (RNA polymerase \(\beta\)-subunit). rpoB is generally accepted to provide a good representation of the phylogenetic relationships among the Bacteria (11), as it provides a phylogenetic resolution comparable to that of 16S rRNA with the advantage of being a single-copy gene. To construct the organismal tree, the rpoB gene sequences from all the organisms in Table S1 in the supplemental material and representative species of all major phyla of the Bacteria were downloaded from the KEGG database and aligned with ClustalW (59) using the translated protein sequences. Alignments were carried out with default parameters and visually inspected by use of Molecular Evolutionary Genetics Analysis (MEGA), version 4 (58). Hypervariable regions with ambiguous alignment were excluded from the analysis. The lsrB tree was built with all the sequences identified as being functional lsrB orthologs (group I, Table 3), and the tree was inferred using the maximum likelihood in PAUP* 4.0b10 (53) using heuristic searches, 10 random taxon additions, and TBR (tree bisection and reconnection) branch swapping. MrBayes 3.1.2 (44) was used to infer branch support by running two simultaneous sets of four Markov chains for 1 million generations sampled every 100 generations. The distribution of the log likelihoods was used to evaluate the stationarity of this parameter and to determine burn-in values. Modeltest 3.7 (38) and MrModeltest 2.2 (34) were used to select the best-fitting evolutionary models for phylogenetic analyses. The rpoB phylogeny was estimated with a total data set of 83 species. This data set was translated to amino acids and analyzed using neighbor joining (45) with Poisson correction distances (68) and a gamma distribution rate variation among sites. Nodal support was estimated with non-parametric bootstrapping (1,000 replicates). The rpoB trees were rooted with Thermotoga maritima (Thermotogales). These analyses were carried out using MEGA.

### RESULTS

**LsrB orthologs in completely sequenced bacterial genomes.** To search for orthologs of LsrB, we carried out a reciprocal best-hit analysis against all 809 completely sequenced bacterial genomes present in the KEGG database as of January 2009 using the protein sequence of LsrB from S. Typhimurium (STM4077). The reciprocal best-hit strategy of sequence similarity comparisons was employed previously for this type of study because it allows the distinction between orthologs and paralogs (10). The organisms with proteins identified as being orthologs are shown in Table 3 (KEGG protein identities and E values are provided in Table S1 in the supplemental material). Sorting these organisms in order of percentage of identity of the LsrB orthologs with the S. Typhimurium AI-2 receptor clearly revealed two distinct groups of LsrB orthologs, a first group with a high percentage of identity (>60%; E value below 1E-103) and a second group with a percentage of identity below 36% (E value higher than 1E-44), which we termed group I and group II, respectively.

We then performed the reciprocal best-hit analysis against all genomes using each LsrB protein sequence from group I as a reference (i.e., instead of LsrB from S. Typhimurium), in all cases, the only hits with greater than 57% identity were the other protein sequences included in group I from the first analysis. Thus, the group I orthologs are consistent regardless of the LsrB sequence used as a reference.

The genomes of the organisms with LsrB orthologs were further analyzed to identify orthologs of the other proteins encoded by the lsr operon. As shown in Table 3, all species of group I have orthologs of all the proteins encoded by the lsr operon (with the exception of LsrF, a putative AI-2-processing protein, in Rhodobacter sphaeroides), whereas none of the group II organisms have orthologs of the complete operon, lacking at least two proteins encoded by genes from this operon in all cases. LsrE was not included in this analysis because the protein seems to be exclusive to the Salmonella genus, and an LsrE knockout mutant in S. Typhimurium showed no phenotype related to the regulation of the lsr operon or AI-2 production (56, 57). Reasoning that a conservation of the residues that formed hydrogen bonds with AI-2 (31) would be crucial to LsrB function, we next used a fold recognition-based server to predict structures for the LsrB orthologs. The sequences of the LsrB orthologs were submitted to the PHYRE Web server (22), which returned structure predictions and structure-based alignments based on each of the 10 best-scoring template Pro-

### Table 2. Primers used in this study

<table>
<thead>
<tr>
<th>Construct (purpose)</th>
<th>5' (sense) sequence</th>
<th>3' (antisense) sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>R. leguminosarum/pPro (PCR)</td>
<td>CGCGGATCGTCGAAGACGTTGGAGAACTG</td>
<td>CCGGCTCGAGCAGCTCGAGATCGT</td>
</tr>
<tr>
<td>A. tumefaciens/pPro (PCR)</td>
<td>CGCGGATCGTCGAAGACGTTGGAGAACTG</td>
<td>CCGGCTCGAGCAGCTCGAGATCGT</td>
</tr>
<tr>
<td>E. coli MG1655/pPro (PCR)</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
</tr>
<tr>
<td>B. anthracis/pET151 (D171N)</td>
<td>CGCGGATCGTCGAAGACGTTGGAGAACTG</td>
<td>CCGGCTCGAGCAGCTCGAGATCGT</td>
</tr>
<tr>
<td>B. anthracis/pET151 (A227T)</td>
<td>CGCGGATCGTCGAAGACGTTGGAGAACTG</td>
<td>CCGGCTCGAGCAGCTCGAGATCGT</td>
</tr>
<tr>
<td>S. Typhimurium/pGEX (PCR)</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
</tr>
<tr>
<td>B. cereus/pGEX (PCR)</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
</tr>
<tr>
<td>B. anthracis/pET151 (PCR)</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
<td>CTTAATTGCAGTATTTGGTTCGAGATATTGTC</td>
</tr>
</tbody>
</table>

*a See reference 31.*
tein Data Bank structures available in the PHYRE library. For all of the orthologs in group I and two-thirds of the orthologs in group II, the structure of S. Typhimurium LsrB was returned as one of these top 10 templates. The alignments with S. Typhimurium LsrB were then examined to determine if residues previously shown to form hydrogen bonds with R-THMF in S. Typhimurium LsrB (K35, D116, D166, Q167, P220, and A222) (31) were predicted to be structurally conserved. Strikingly, as shown in the last column of Table 3, these six residues were completely conserved in all of the orthologs in group I and differed in at least two positions in all cases for group II. Residue D166 (numbering based on S. Typhimurium LsrB)
was not conserved in any of the group II orthologs, most typically being replaced with an asparagine. The other most common substitution was A222T (full listing of the nonconserved amino acids is given in Table S1 in the supplemental material).

Based on these results, we hypothesize that the species in group I, which have >60% identity, orthologs to the proteins encoded by the lsr operon, and all six AI-2 binding-site residues conserved, have functional LsrB-like AI-2 receptors, whereas group II proteins are likely to have a different function.

Profiles of AI-2 removal from extracellular medium. Previous studies of S. Typhimurium (57), E. coli (65), S. meliloti (37), and A. actinomycetemcomitans (48) revealed that the lsr operons in these organisms encode proteins involved in an ABC transport system that imports extracellular AI-2. Thus, in the presence of these organisms, AI-2 does not persist in the extracellular medium but is internalized by the cells and further modified. Our analysis, described above, indicated that all the organisms predicted to have functional LsrB receptors (group I, Table 3) also had orthologs to all the proteins encoded by the lsr operon. Thus, we predicted that the organisms in group I have a functional Lsr system for AI-2 internalization and that these organisms would rapidly remove AI-2 from culture fluids. In contrast, for organisms from group II, which lack orthologs to some of the proteins encoded by the lsr operon and presumably do not have a functional AI-2 transport system, we predicted that AI-2 would persist in the extracellular media. To test these predictions, we compared the profiles of AI-2 removal for a set of organisms from groups I and II.

Our analysis revealed that almost all E. coli strains (11 out of 13) analyzed belong to group I. However, two E. coli strains (avian-pathogenic E. coli and UPEC UT189) have LsrB orthologs with very low sequence identity and lack orthologs to several of the proteins from the lsr operon, and they are therefore classified as members of group II (Table 3). We tested an E. coli strain (MG1655) from group I for AI-2 uptake and found, as was previously shown (65), that this strain removed AI-2 from culture fluids (Fig. 1A). We then compared the AI-2 removal profile of UPEC strain UT189 (from group II) with the profile from E. coli strain MG1655 and observed that while E. coli MG1655 efficiently cleared AI-2 from culture fluids by 6 h, UPEC strain UT189 cleared little, if any, AI-2 by 10 h (Fig. 1A). This supports our prediction that UPEC strain UT189, although belonging to the same species as MG1655, is a member of group II and accordingly does not have a functional Lsr transport system for AI-2 uptake.

Like E. coli MG1655, two Bacillus strains, B. cereus (ATCC 1087) and B. anthracis (vaccine strain Sterne 34F2), have orthologs classified as belonging to group I. Putative AI-2 receptors were identified in these species previously (41, 50) but not confirmed experimentally. We tested these strains for AI-2 removal, and as expected, they were able to completely remove AI-2 from culture fluids (Fig. 1B), supporting the premise that organisms in group I have functional AI-2 transporters.

To further test the premise that group II organisms are unable to incorporate AI-2, we compared AI-2 removal profiles for organisms from the family Rhizobiaceae from group II (R. etli, R. leguminosarum, and A. tumefaciens) with AI-2 in the only member of the Rhizobiaceae from group I (S. meliloti). None of the members of the Rhizobiaceae shown in Table 3 has LuxS orthologs, and thus, we expected that none of these species would produce AI-2. This was confirmed by the fact that cell-free culture fluids collected from these bacteria produced only low levels of bioluminescence induction in a V. harveyi BB170 bioassay (data not shown). However, as we have previously shown in the case of S. meliloti, non-AI-2-producing species can still be capable of taking up AI-2 produced synthetically or by other species (37). Thus, in order to compare AI-2 removal profiles for these species, we cultured these bacteria to the same cell density (OD_{600} of 1), supplied chemically synthesized AI-2, and measured AI-2 activity in the culture fluids over time (Fig. 2). Over the time of the measurements, S. meliloti effectively removed the exogenously provided AI-2, while the other three species did not, supporting the prediction that the bacteria tested from group II (R. etli, R. leguminosarum, and A. tumefaciens), and likely all group II species, do not have Lsr systems capable of taking up AI-2.

In vitro AI-2 binding to LsrB orthologs. While the abovedescribed results support our ability to identify species with functional AI-2 transporters, they do not directly show that the identified LsrB ortholog is responsible for AI-2 binding. In order to directly test for AI-2 binding ability, we cloned the LsrB orthologs from the same organisms tested as described.
above (three belonging to group I and four belonging to group II) and compared their abilities to bind AI-2 with that of LsrB from S. Typhimurium. The candidate proteins were overexpressed in both an E. coli strain that produces AI-2 and, as a negative control, a luxS mutant E. coli strain that does not make AI-2. These proteins were then purified and tested for the ability to bind AI-2 using a previously developed assay (12) in which the protein is heat denatured to release any bound ligand. The denatured protein is then pelleted, and the resulting supernatants are added to a reporter strain of V. harveyi that bioluminesces in response to AI-2. As shown in Fig. 3, all three orthologs from group I (i.e., that have >60% identity, a complete set of orthologs to the lsr genes, and the six amino acids from the binding pocket conserved), E. coli MG1655, B. cereus, and B. anthracis, showed a LuxS-dependent AI-2 binding ability similar to that observed for the previously characterized S. Typhimurium LsrB protein (Fig. 3). Conversely, no AI-2 binding activity was detected in the candidates from group II (R. etli, R. leguminosarum, A. tumefaciens, and UPEC UT189) (Fig. 3). Thus, as predicted from sequence analysis and structure prediction (above), LsrB orthologs from group I demonstrate AI-2 binding ability, while group II orthologs lack this ability.

The amino acids aspartate 166 and alanine 222 are required for AI-2 binding. Based on predicted structure-based sequence alignments (see above), the amino acids that form hydrogen bonds with AI-2 are completely conserved in all of the LsrB orthologs that demonstrated the ability to bind AI-2. In contrast, all the proteins that were unable to bind AI-2 in our in vitro assays lacked at least two of these residues. Specifically, in

FIG. 2. Removal of exogenously supplied AI-2. S. meliloti (triangles), R. leguminosarum (circles), R. etli (squares), and A. tumefaciens (crosses) were cultured to an OD$_{600}$ of 1 in their optimal culture media (LBMC, TYC, YEM, and LB, respectively). Chemically synthesized AI-2 was then added to all the cultures, and aliquots were taken at the specified times. AI-2 activity in cell-free culture fluids is reported as the change from the induction of light produced by V. harveyi BB170.

![Graph showing AI-2 activity in the culture fluids over time](http://example.com/graph.png)

FIG. 3. Binding of AI-2 to potential LsrB-like orthologs. Proteins were expressed in either LuxS$^+$ (black bars) or LuxS$^-$ (white bars) E. coli strains (BL21 and FED101, respectively), purified, and denatured to release the ligand. The released ligand was added to a V. harveyi AI-2 reporter strain (BB170) to determine AI-2 activity. AI-2 activity is reported as the change in the induction of light production by V. harveyi BB170 supplemented with protein supernatant compared to that of the appropriate buffer. Error bars represent the standard deviations for three independent cultures.
R. etli, R. leguminosarum, and A. tumefaciens, there are predicted to be two substitutions: D166N and A222T (numbering follows that of LsrB from S. Typhimurium). Indeed, the majority of the proteins in group II have these substitutions, although other substitutions are observed (see Table S1 in the supplemental material for detailed information). The complete conservation of AI-2 hydrogen binding residues in orthologs of group I but not group II is apparent in a multiple-sequence alignment of all of the LsrB orthologs for which we have experimental data (purple in Fig. S1 in the supplemental material). It is worth noting that 29 non-binding-site residues are completely conserved across all groups in this alignment (yellow in Fig. S1 in the supplemental material). However, structural analysis shows that these residues are not clustered. Moreover, these residues are disproportionately Gly and Pro (10 and 4 conserved occurrences, respectively), suggesting that unlike the six residues in the binding site, these residues are conserved for structural rather than functional reasons.

We interpreted this to indicate that residues D166 and A222 are essential for the AI-2 binding ability, and to test this idea, we introduced the mutations described above (D166N and A222T) into the B. anthracis LsrB ortholog, both individually and together, and assayed for AI-2 binding ability. As shown in Fig. 4, while the wild-type protein is capable of binding AI-2, no AI-2 activity was present in the binding pockets of any of the mutants as measured by the V. harveyi bioassay. As a complementary experiment, we tested the ability to create AI-2 binding capacity in the distantly related LsrB ortholog of R. etli by mutating the putative binding-site residues to mimic the binding site of the proteins from group I. These mutants failed to show AI-2 binding in the V. harveyi bioassay (data not shown), indicating that these proteins have already diverged to such a degree that other aspects of the protein structure important for AI-2 binding are missing.

These results show that D166 and A222, conserved in all the LsrB orthologs we have shown to bind AI-2, are necessary (though not sufficient) for the ability of these proteins to bind AI-2 and thus provide a useful criterion for the identification of other LsrB-like AI-2 receptors. It is possible that more conservative mutations would still allow AI-2 binding, but such mutations were not observed in our list of orthologs. Furthermore, these results support the hypothesis that the proteins in group II are incapable of AI-2 binding and are therefore very unlikely to function as AI-2 receptors in vivo.

Evolution of functional LsrB-like AI-2 receptors. Our sequence/structural and functional studies lead us to predict that all the organisms from group I have LsrB orthologs that function as LsrB–AI-2 receptors. This group contains members from the evolutionarily distant orders of the Enterobacteriales, Pasteuromellales, Rhizobiales, Rhodobacteriales, and Bacillales. To infer the evolutionary history of the lsrB gene, we determined the phylogenetic tree of all the lsrB gene orthologs from group I (Fig. 5) and compared it to the rpoB housekeeping gene organismal tree constructed with representatives of all major phyla of the Bacteria (Fig. 6). Importantly, the organismal tree recovers all major phyla and classes with high bootstrap support. The relationship among phyla has a lower bootstrap support, but this does not influence our analysis because the phylogenetic relationship between all species with functional lsrB genes (Fig. 6) is also well supported in this tree.

This analysis indicates that the phylogenies of lsrB and rpoB largely overlap in their diversification patterns although with some important exceptions. The majority of the species included in group I of Table 3 clustered within the Enterobacteriales and Pasteuromellales (both Gammaproteobacteria), and the diversification pattern of the lsrB gene mimics the phylogenetic relationships obtained in the rpoB organismal tree within this group (compare distributions in Fig. 5 and 6); that is, the lsrB gene tree recovers all species groups, and the relationship among members of the Enterobacte- riales and Pasteuromellales is largely congruent between gene trees. Additionally, the widespread occurrence of LsrB among the Enterobacteriales and Pasteuromellales strongly suggests a single origin for this AI-2 receptor that occurred in an ancestor of these organisms after the diversification of the Enterobacteriales and Pasteuromellales from the Vibrionales. Nonetheless, the presence of lsrB genes in the Enterobacte- riales and Pasteuromellales is not ubiquitous, as shown by Er-
Carotovora and two E. coli strains (UTI89 and APECO1), suggesting independent events of gene loss (Fig. 6).

The major discordance between the lsrB and rpoB phylogenies relates to the occurrence of functional LsrB in S. meliloti (Rhizobiales, Alphaproteobacteria), R. sphaeroides (Rhodobacterales, Alphaproteobacteria), and three species of Bacillus (Bacillales, Firmicutes). Specifically, lsrB genes from these species cluster with strong nodal support (Bayesian posterior probability of 1.0) (Fig. 5) with specific clades of the Enterobacterales and Pasteurellales. Thus, these species appear “misplaced” in the lsrB gene phylogeny (Fig. 5), in contrast with the organis-

**FIG. 5.** Evolutionary history of genes encoding functional LsrB orthologs inferred with maximum likelihood. The lsrB gene tree was constructed with the sequences from all organisms in group I. This is an unrooted phylogram oriented to show maximum congruence with the organismal tree. Numbers on the nodes indicate posterior probability as estimated with MrBayes.
FIG. 6. Molecular phylogeny of Bacteria estimated with the rpoB gene. The rpoB gene tree was constructed with all the organisms listed in Table 3 (groups I and II) and representative species of all major phyla of the Bacteria. This represents our best inference of the organismal tree. Gray boxes indicate species with functional lsrB genes (group I, Table 3), and the dashed box locates the species with protein sequences in Table 3 that are likely to function as a rhamnose binding protein. The numbers after species names indicate the number of strains analyzed for the respective species. Taxonomic classifications (phyla) are shown on the right. This tree was inferred with neighbor joining, and the branch lengths are scaled to the number of amino acid substitutions per site. Thickened branches indicate high bootstrap support (higher than 75%). This is a measurement of phylogenetic strength between nodes, and this value reflects a high confidence in the inferred relationships between species. Branch lengths are scaled in terms of the number of nucleotide substitutions per site.
mal phylogeny (rpoB tree) (Fig. 6). This type of incongruence is consistent with lateral gene transfer (LGT) events (9, 54).

In the case of Bacillus species, the phylogenetic pattern of the lsrB gene tree reveals that these species cluster with the Pasteurellales. Thus, the occurrence of the lsrB gene in the Bacillus lineage could be explained by a putative LGT event from bacteria of the family Pasteurellaceae. The occurrence of this gene within so many Bacillus species indicates that if such a transfer occurred, enough time has passed for the lineage to diversify into at least three different species (Fig. 5 and 6).

The two species from the Alphaproteobacteria (S. meliloti and R. sphaeroides) are nested within the Enterobacteriales, clustering with the Klebsiella and Enterobacter species. Given the phylogenetic distance that separates S. meliloti and R. sphaeroides (Fig. 6), it is surprising that lsrB gene topology clusters these two species together. The most likely explanation for this occurrence requires at least more than one LGT event. Such a pattern could be obtained if two sequential LGT events had occurred, for example, from one enterobacterium (most likely an ancestor of Klebsiella and Enterobacter) first to a Sinorhizobium species and then to a Rhodobacter species or from one Enterobacter species first to a Rhodobacter species and then to a Sinorhizobium species. However, with the data at hand, it is difficult to predict the specific order of these events. Furthermore, we predict that the proposed LGT to S. meliloti and R. sphaeroides must have been a quite recent event given that no further alphaproteobacterium species were identified with group I LsrB orthologs. Alternatively, we could postulate one LGT event to the ancestor of these Alphaproteobacteria with a massive number of gene losses, but we find this possibility to be very unlikely.

**DISCUSSION**

A variety of bacterial species have been shown to be capable of responding to AI-2 by the regulation of a range of niche-specific functions, but the mechanisms for AI-2 detection have been characterized in only a few cases (17, 64). This constitutes a major obstacle in work toward an understanding of the function of AI-2. While sequence analysis of bacterial genomes reveals the presence of orthologs of LsrB-like AI-2 receptors in gram-negative as well as gram-positive bacteria (41, 50; this study), establishing which orthologs are, in fact, functional as AI-2 receptors is important for determining if and how these species use AI-2 as a chemical signal. Thus, after analyzing the sequences and predicted structures of LsrB orthologs, we identified criteria for predicting which LsrB orthologs are functional AI-2 receptors and assayed the AI-2 binding ability of selected candidates to test our criteria. Our results not only support our predictions but also provide the first biochemical confirmation of the presence of functional AI-2 receptors in gram-positive bacteria, specifically in B. anthracis and B. cereus.

Our sequence and structural analyses allowed us to categorize the organisms with LsrB orthologs into two different groups. Members of group I have (i) LsrB orthologs with greater than 60% sequence identity with S. Typhimurium LsrB, (ii) orthologs to the other key transport proteins encoded by the lsr operon, and (iii) complete conservation of all six residues that hydrogen bond with AI-2 in S. Typhimurium LsrB (based on structure predictions). On the other hand, in organisms belonging to group II, the LsrB orthologs have a sequence identity below 36%, are missing orthologs to key proteins encoded by the lsr operon, and lack at least two of the six residues in the AI-2 binding pocket. These characteristics led us to hypothesize that the organisms from group I had functional AI-2 binding proteins, whereas the LsrB orthologs in group II were likely to have a different function. In all organisms where the function of either the LsrB protein or its gene has been studied, LsrB has been shown, along with other proteins that form the Lsr transport system, to participate in the uptake of AI-2 (37, 48, 57, 65); thus, we further predicted that organisms with a functional LsrB and orthologs to all the proteins from the Lsr system would take up AI-2. Accordingly, all the organisms from group I tested for the binding of AI-2 by LsrB or for in vivo AI-2 removal (S. Typhimurium, S. meliloti [37], E. coli K-12 [MG1655], B. cereus, and B. anthracis) were capable of binding AI-2, nor were these organisms able to take up AI-2. In addition, our analysis of predicted structures of the LsrB orthologs identified key binding-site residues that are not conserved in group II organisms. Mutagenesis of the B. anthracis LsrB ortholog (classified as group I and demonstrated to bind AI-2) with the two most common group II substitutions (D166N and A222T) confirmed that these residues are critical for AI-2 binding. This result strongly supports our use of binding-site conservation as a key criterion in identifying group I orthologs.

These results offer experimental evidence that functional LsrB–AI-2 receptors are present in particular members of the Enterobacteriaceae (S. Typhimurium and E. coli), Rhizobiaceae (S. meliloti), and Bacillaceae (B. cereus and B. anthracis), and given the correlation of our experimental results with our classification scheme, we predict that all the other LsrB orthologs from group I are functional AI-2 receptors and that these organisms are competent for AI-2 uptake. Accordingly, we expect that the members of the families Pasteurellaceae and Rhodobacteraceae in group I (Table 3) also have functional AI-2 transporters. On the other hand, we believe that it is likely that all group II members have orthologs that are not involved in AI-2 transport and thus that these organisms do not take up AI-2 via an LsrB-type mechanism. The criteria described here can be used to predict the presence (or absence) of functional LsrB-like AI-2 receptors in newly sequenced species, and as new species are sequenced, we expect the number of organisms in group I to increase.

The large majority of the organisms from group I belong to the Enterobacteriales and the Pasteurellales. This, coupled with the fact that the diversification pattern of the lsrB gene largely mimics the bacterial phylogenetic relationships within this group, is consistent with a single origin for the LsrB–AI-2 receptor that likely occurred in an ancestor of these organisms after the diversification of the Enterobacteriales and the Pasteurellales from the Vibrionales. Thus, the occurrence of LsrB receptors in one species of the Rhizobiiales (S. meliloti) and the Rhodobacterales (R. sphaeroides) and in three species of the Bacillales was very surprising and immediately raised the possibility of LGT. The hypothesis of LGT between organisms from the Enterobacteriales or the Pasteurellales and these three
orders was supported by the comparison of the \(lsrB\) gene tree and the \(rpoB\) organismal tree. Specifically, in the \(lsrB\) gene tree, \textit{Bacillus} species are clustered with the \textit{Pasteurellales}, and \textit{S. meliloti} and \textit{R. sphaeroides} are nested within the \textit{Enterobacte- riaeales}. These are nested patterns where species appeared to be “misplaced” in the gene phylogeny and can be interpreted as an indication of events of LGT. Often, genes that have been acquired by LGT have an atypical nucleotide distribution (reflected in GC content or codon usage) compared with the rest of the genome (25). However, in this case, analysis of GC usage and codon bias provided no information to argue for or against the hypothesis of LGT (data not shown). Certainly, other occurrences such as convergent evolution by natural selection or the ancient origin of \(lsrB\) at the base of the \textit{Bacteria} tree with a large number of events of gene loss could also explain the observed patterns, but since we do not have specific data to support a particular explanation over the others, we favor LGT as the most parsimonious explanation, as it requires the minimum number of assumptions. LGT events are now well accepted as a major force in the evolution of bacterial genomes (8, 23) leading to an increment in the number of genes (35) and pathways (19) and often enabling bacteria to acquire new functions, such as traits associated with pathogenicity, that allow adaptation to novel environments. In the specific cases of \textit{S. meliloti} and \textit{R. sphaeroides}, it is intriguing that these organisms have acquired the AI-2 receptor but not its synthase (\(LuxS\)); thus, these organisms have potentially gained the ability to eavesdrop on their neighbor’s signal, as previously suggested (37, 41). It will also be interesting to determine the adaptive value of this new function and explore its impact on the physiology of these organisms. LGT has been proposed for other autoinducer receptors and regulators from the LuxI/LuxR family of species-specific quorum-sensing proteins, where it was previously proposed that the acquisition of this family of proteins has benefited certain bacterial species by allowing them to gain an efficient mechanism for regulating virulence genes (8, 16, 26).

Interestingly, the \textit{LsrB} ortholog in \textit{R. leguminosarum} \textit{bv. trifolii}, which we identified as belonging to group II, has been shown to be essential for rhamnose (a methyl-pentose sugar) uptake and growth in this sugar and is thus likely to be a rhamnose binding protein (42, 43). Motivated by this finding, we used the protein sequence of \(R. leguminosarum\) \textit{bv. trifolii} (KEGG identification number pRL110413) to carry out a reciprocal best-hit analysis against all the genome sequences used in the previous analysis. We found that there are 12 orthologs of the \textit{R. leguminosarum} binding protein (along with the proteins from the rhamnose transport operon) present in group II (shown in Table 3). Thus, these 12 binding proteins are orthologs of both \textit{LsrB} of \textit{T. Pyhimurium} and the rhamnose binding protein of \textit{R. leguminosarum}. These proteins have more than 65% sequence identity with the \textit{R. leguminosarum} protein but less than 36% identity with \textit{S. Typhimurium} \textit{LsrB}.

We interpret this as strong evidence that these 12 proteins in group II are functioning as rhamnose binding proteins, in agreement with our prediction that they are not AI-2 receptors (these proteins are highlighted in the Table S1 in the supplemental material). These 12 organisms correspond to species belonging to \textit{Alphaproteobacteria} that cluster together in the organismal \textit{rpoB} tree (Fig. 6). Interestingly, \textit{S. meliloti} is the only organism that has an \textit{LsrB} ortholog belonging to group I and also a different set of proteins that are orthologs to the \textit{R. leguminosarum} proteins from the rhamnose transport operon, further corroborating our hypothesis that the acquisition of \textit{LsrB} occurred by LGT in \textit{S. meliloti}.

While the presence of a functional \textit{LsrB} ortholog does not prove that AI-2 import is involved in the control of AI-2-mediated behavior, it is suggestive. Accordingly, the function of the \(Lsr\) system in AI-2 signaling has already been shown for a member of the \textit{Pasteurellaceae}, \textit{A. actinomycetemcomitans} (an organism not present in Table 3 because, to date, its genome is not present in the KEGG database). Shao and coworkers previously showed that this oral pathogen is capable of internalizing AI-2 via the \(Lsr\) system and, importantly, that \textit{LsrB} is required to mediate the complete AI-2-dependent activation of biofilm formation in this organism (48, 49). In other cases such as that of \textit{Photobacterium luminescens}, an insect pathogen belonging to the \textit{Enterobacteriaceae}, the transcription of the \textit{lsr} operon was shown to be induced by AI-2, and AI-2 was also implicated in the regulation of biofilm formation and motility (24). However, it remains to be demonstrated whether or not the \textit{Lsr} system is involved in mediating these AI-2-regulated behaviors. Likewise, it will be interesting to determine whether the \(Lsr\) system is involved in mediating AI-2 signal transduction in \textit{B. cereus} and \textit{B. anthracis}, where AI-2 has been implicated in regulating biofilm formation (2) and growth rate (21). Certainly, the results presented here give support to that possibility.

This study, along with the two previous studies based on sequence analysis (41, 50), also reveals that certain bacteria such as \textit{Helicobacter pylori} (39), \textit{Streptococcus mutans} (55), \textit{Staphylococcus epidemidis} (27), \textit{Porphyromonas gingivalis} (20, 67), \textit{Pseudomonas aeruginosa} (14), and \textit{Bacillus subtilis} (28), which have been shown to respond to AI-2, do not have either of the known types of AI-2 receptors (neither LuxP nor \textit{LsrB}), and thus we expect that other receptors for AI-2 remain to be discovered. These receptors may be of entirely new classes or may be promiscuous receptors for other small molecules. Novel receptor classes are likely to be identified by approaches that rely on genetic screens to isolate mutants involved in modulating AI-2-regulated phenotypes, and as shown here, integration with approaches that use sequence analysis coupled with biochemical assays may prove very useful. Clearly, an elucidation of the proteins involved in AI-2 recognition and signal relay is essential for studying the potential functions of this class of signal molecule in intra- and interspecies cell-to-cell communication and/or intra- and intercellular signal transduction. The identification and experimental confirmation of functional \textit{LsrB} receptors in this study open the door to the understanding of the molecular basis of AI-2-mediated behavior regulation in a variety of new species.

ACKNOWLEDGMENTS

The work performed in the laboratory of K.B.X. was supported by Fundação para a Ciência e Tecnologia, Portugal (FCT), grant PTDC/ BIA-BCM/73676/2006. C.S.P. and P.H.B. were supported by FCT awards SFRH/BD/28543/2006 and SFRH/BPD/26852/2006, respectively. The laboratory of S.T.M. gratefully acknowledges grants from the NIH (grant AI074041), the Camille and Henry Dreyfus Foundation, and Swarthmore College. We also thank program FLAD/NSF
project 600-10/2006 for sponsoring the stay of S.T.M. in the laboratory of K.B.X. during his sabbatical.

We thank Chengetai Mahomva and Bella Liu for technical assistance. We are very grateful to Adrian O. Henriques, Gladys Alexander, James P. Shapleigh, Jeffrey I. Gordon, and Marcus B. Jones for providing the strains used in this study. We thank Bonnie Bassler, Michiko Taga, Brian Hammer, and Isabel Gordo for critically reading the manuscript. We also thank Bonnie Bassler for receiving C.S.P. in her laboratory.

REFERENCES


