Identification Of Novel Autoinducer-2 Receptors In Clostridia Reveals Plasticity In The Binding Site Of The LsrB Receptor Family

I. M. Torcato
M. R. Kasal
P. H. Brito
Stephen T. Miller
Swarthmore College, smiller1@swarthmore.edu
K. B. Xavier

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

Part of the Biochemistry Commons

Let us know how access to these works benefits you

Recommended Citation

This work is brought to you for free by Swarthmore College Libraries' 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 novel autoinducer-2 receptors in Clostridia reveals plasticity in the binding site of the LsrB receptor family

Received for publication, December 3, 2018, and in revised form, January 26, 2019 Published, Papers in Press, January 29, 2019, DOI 10.1074/jbc.RA118.006938

Inês M. Torcato,† Meghann R. Kasal,‡ Patrícia H. Brito,§ Stephen T. Miller,¶ and Karina B. Xavier

From the †Instituto Gulbenkian de Ciência, Rua da Quinta Grande, 6, 2780-156 Oeiras, Portugal, the ‡Instituto de Tecnologia Química e Biomolecular António Xavier, Universidade NOVA de Lisboa, Avenida da República, 2780-157 Oeiras, Portugal, the §Department of Chemistry and Biochemistry, Swarthmore College, Swarthmore, Pennsylvania 19081, and the ¶Chronic Disease Research Center (CEDOC), NOVA Medical School, Universidade NOVA de Lisboa, Rua Câmara Pestana, 6, 1150-082 Lisboa, Portugal

Edited by Chris Whitfield

Autoinducer-2 (AI-2) is unique among quorum-sensing signaling molecules, as it is produced and recognized by a wide variety of bacteria and thus facilitates interspecies communication. To date, two classes of AI-2 receptors have been identified: the LuxP-type, present in the Vibrionales, and the LsrB-type, found in a number of phylogenetically distinct bacterial families. Recently, AI-2 was shown to affect the colonization levels of a variety of bacteria in the microbiome of the mouse gut, including members of the genus Clostridium, but no AI-2 receptor had been identified in this genus. Here, we identify a noncanonical, functional LsrB-type receptor in Clostridium saccharobutylicum. This novel LsrB-like receptor is the first one reported with variations in the binding-site amino acid residues that interact with AI-2. The crystal structure of the C. saccharobutylicum receptor determined at 1.35 Å resolution revealed that it binds the same form of AI-2 as the other known LsrB-type receptors, and isothermal titration calorimetry (ITC) assays showed that binding of AI-2 occurs at a submicromolar concentration. Using phylogenetic analysis, we inferred that the newly identified noncanonical LsrB receptor shares a common ancestor with known LsrB receptors and that noncanonical receptors are present in bacteria from different phyla. This led us to identify putative AI-2 receptors in bacterial species in which no receptors were known, as in bacteria belonging to the Spirochaetes and Actinobacteria phyla. Thus, this work presents a significant step toward understanding how AI-2-mediated quorum sensing influences bacterial interactions in complex biological niches.

Bacteria are able to colonize and adapt to different environments, and the process of quorum sensing greatly enhances this ability. Quorum sensing is a mechanism of cell-to-cell communication mediated by exchange of small chemical molecules, named autoinducers, that allows bacteria to monitor their population density and regulate gene expression accordingly (1–3). Behaviors regulated by this process include biofilm formation, virulence factor expression, antibiotic production, and bioluminescence (4, 5). Autoinducer-2 (AI-2) is a signaling molecule produced and recognized by a wide variety of bacteria and thus facilitates interspecies communication (4, 6–8). Because bacteria often produce species-specific signals as well, the relative proportion of AI-2 depends on the identity and number of bacterial species present in the environment; thus AI-2-sensitive bacteria can, in principle, regulate their behavior according to the species composition of the community (9).

To date, all characterized AI-2 receptors have high sequence similarity to, and complete binding site identity with, one of two receptors: LuxP or LsrB (4, 6, 7). Both receptors belong to the high affinity substrate-binding protein family. Importantly, the two receptors bind chemically distinct derivatives of the AI-2 precursor molecule DPD ((S)-4,5-dihydroxypentane-2,3-dione) (6,7). Whereas LuxP binds S-THMF-borate ((2S,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran-borate), LsrB binds R-THMF ((2R,4S)-2-methyl-2,3,3,4-tetrahydroxytetrahydrofuran), with discrimination between these AI-2 adducts being presumably due to differences in the amino acid composition of the binding sites of these receptors (7). Regulatory mechanisms also differ by receptor type. In response to AI-2 binding, LuxP modulates the activity of a membrane-spanning sensor histidine kinase protein, thus regulating a phosphorylation-mediated signaling cascade. In contrast, LsrB-mediated signaling is believed to rely on a pathway involving a two-component system.
tion signal transduction cascade, whereas LsrB is part of an ATP-binding cassette (ABC) transport system that internalizes AI-2 (10–12). LsrB was first identified in *Salmonella enterica* serovar *Typhimurium* but, unlike LuxP, which has only been found in the Vibrionales, it is also present in other enteric bacteria and members of the *Rhizobiaceae* and *Bacillaceae* families (7, 13, 14).

The LsrB-mediated transport system is comprised of two transmembrane proteins, LsrC and LsrD, and an ATPase, LsrA, that are encoded in the *lsr* (for *LuxS* regulated) operon together with LsrB. As the density of AI-2–producing bacteria increases, AI-2 accumulates extracellularly. When the threshold concentration is reached, LsrB-bound AI-2 is internalized via this transport system (14,15). Internalized AI-2 is subsequently phosphorylated by the kinase LsrK (16–18), and phosphorylated AI-2 relieves the repression caused by the *lsr* regulator, LsrR (19, 20). This derepression leads to *lsrABCDGF* expression and thus rapid depletion of AI-2 from the extracellular medium. In most bacteria that possess the *lsr* operon, the enzymes responsible for further AI-2 processing, LsrG and LsrF, are also encoded in the *lsr* operon (see Fig. 1 for an explanatory scheme) (21, 22). It has been shown that Lsr-mediated AI-2 internalization influences the expression of downstream genes involved in regulating aggregation, attachment, and biofilm formation in *Escherichia coli* (23–25). Moreover, in pathogenic enterohemorrhagic *E. coli*, extracellular AI-2 acts as a chemoattractant; this activity, mediated by the LsrB receptor, is required for cell aggregation and biofilm formation (24–28). Thus, AI-2 internalization presumably regulates these behaviors via the different responses exerted by intracellular and extracellular AI-2.

In previous work, we identified LsrB receptors in several species (including *Bacillus anthracis* and *Bacillus cereus*, the first Firmicutes to be shown to have functional LsrB receptors) through an approach that used bioinformatics to identify candidate receptors and biochemical and genetic studies to confirm their function (13). Based on this work, we proposed the following criteria to identify functional LsrB receptors: (i) more than 60% sequence identity with the LsrB from *Salmonella* serovar *Typhimurium* but, unlike LuxP, which has only been found in the Vibrionales, it is also present in other enteric bacteria and members of the *Rhizobiaceae* and *Bacillaceae* families (7, 13, 14).

The LsrB-mediated transport system is comprised of two transmembrane proteins, LsrC and LsrD, and an ATPase, LsrA, that are encoded in the *lsr* (for *LuxS* regulated) operon together with LsrB. As the density of AI-2–producing bacteria increases, AI-2 accumulates extracellularly. When the threshold concentration is reached, LsrB-bound AI-2 is internalized via this transport system (14,15). Internalized AI-2 is subsequently phosphorylated by the kinase LsrK (16–18), and phosphorylated AI-2 relieves the repression caused by the *lsr* regulator, LsrR (19, 20). This derepression leads to *lsrABCDGF* expression and thus rapid depletion of AI-2 from the extracellular medium. In most bacteria that possess the *lsr* operon, the enzymes responsible for further AI-2 processing, LsrG and LsrF, are also encoded in the *lsr* operon (see Fig. 1 for an explanatory scheme) (21, 22). It has been shown that Lsr-mediated AI-2 internalization influences the expression of downstream genes involved in regulating aggregation, attachment, and biofilm formation in *Escherichia coli* (23–25). Moreover, in pathogenic enterohemorrhagic *E. coli*, extracellular AI-2 acts as a chemoattractant; this activity, mediated by the LsrB receptor, is required for cell aggregation and biofilm formation (24–28). Thus, AI-2 internalization presumably regulates these behaviors via the different responses exerted by intra- and extracellular AI-2.

In previous work, we identified LsrB receptors in several species (including *Bacillus anthracis* and *Bacillus cereus*, the first Firmicutes to be shown to have functional LsrB receptors) through an approach that used bioinformatics to identify candidate receptors and biochemical and genetic studies to confirm their function (13). Based on this work, we proposed the following criteria to identify functional LsrB receptors: (i) more than 60% sequence identity with the LsrB from *Salmonella* serovar *Typhimurium*, (ii) conservation of all six binding-site amino acid residues that interact with AI-2, and (iii) co-occurrence with orthologs to the other key transport proteins encoded by the *lsr* operon. With the constant increase in the number of sequenced genomes available, new strains of bacteria can now be studied. It became clear in the course of this work that, whereas these criteria do identify functional receptors, the identification of a more diverse range of AI-2 receptors requires the expansion of these criteria.

AI-2 receptors are relevant in complex communities such as the gut microbiota, a clinically significant community where species imbalance has been linked to bowel disease, obesity, and susceptibility to pathogen colonization. However, little is known about the interspecies communication mechanisms typically employed in this environment, partly due to the inability to identify the AI-2 receptors of certain members of the microbiota (29, 30). Recently, it was shown that AI-2 can influence the proportions of the major phyla in antibiotic-treated communities.
mice gut microbiota (29). Specifically, the Firmicutes phylum was shown to respond positively to this signal. Interestingly, the majority of Firmicutes, including many Clostridia, have orthologs of the AI-2 synthase (LuxS) and are therefore putative AI-2 producers (29); however, no AI-2 receptors were known or readily identified by prior criteria in the Clostridiales. These findings, together with increasing understanding of the importance of Clostridia members as both commensals and pathogens in the gut microbiota, led us to search for AI-2 receptors in this class (31–33).

In this work we identify new LsrB-type receptors in Clostridium saccharobutylicum and Clostridium autoethanogenum, species shown to exist in the microbiota of mammals (34, 35). We characterize the C. saccharobutylicum receptor and show that this novel receptor binds R-THMF like previously characterized LsrBs, despite the fact that two of the amino acids involved in AI-2 binding differ from those of the canonical LsrBs (13). This variation demonstrates that the plasticity of the LsrB-binding site is higher than previously thought and expands our understanding of the diversity among AI-2-binding receptors. To the best of our knowledge, this is the first report of the identification and characterization of AI-2 receptors in Clostridia. Moreover, this knowledge allowed us to identify new noncanonical putative AI-2 receptors in other bacterial species belonging to the Firmicutes phylum, including bacterial species isolated from human microbiota. We also identified putative noncanonical LsrB receptors in organisms belonging to other phyla where AI-2 receptors had not previously been identified and where AI-2-mediated quorum sensing can now be studied. The identification of AI-2 receptors in these species represents a key step in understanding the mechanisms by which AI-2 regulates the levels of Firmicutes in the mouse gut and, more broadly, how AI-2-mediated quorum sensing influences the behavior of this and other complex communities.

Results

Identification of LsrB orthologs and in vitro binding to AI-2

To identify novel LsrB receptors, we searched the complete genomes deposited in the Kyoto Encyclopedia of Genes and Genomes (KEGG) for LsrB orthologs. We were seeking to identify functional LsrB receptors in the Firmicutes phylum, but we only obtained hits in members of the Bacillus genus using the previous established criteria (13). However, we were intrigued by the identification of LsrB homologous proteins in the Clostridia class. These hits had sequence identities lower than 60%, the minimum observed in previously identified LsrB functional receptors, but still higher than 36%, the maximum observed for the homologs unable to bind AI-2. Proteins identified from C. saccharobutylicum and C. autoethanogenum shared a sequence identity with the LsrB from B. cereus of 39.2 and 37.9%, respectively (Table S2). Sequences for these proteins were submitted to the Phyre2 fold-recognition server and the top two hits for both were the LsrBs from B. anthracis and Salmonella serovar Typhimurium, suggesting that these proteins share the same overall tertiary structure as previously characterized LsrB receptors. We aligned the resulting predicted structures to the LsrB from B. anthracis (Protein Data Bank ID 4PZ0) and found that the putative binding site in the Clostridia orthologs varied from canonical LsrBs in two of the six AI-2-binding residues (P223N, A225S; numbering based on that of the PDB structure of CsLsrB). We also analyzed the genome context of these proteins and found the genome location of the genes encoding the putative LsrB receptors of C. saccharobutylicum and C. autoethanogenum to be identical to that of bacteria with functionally characterized LsrB receptors. The genes are located in a putative operon with homologs for an ABC transport system (Fig. 2A; Table S2). These findings, combined with the biological relevance of the Clostridia class, led us to test if these proteins were functional LsrB-like receptors despite their lower sequence identity and the absence of a fully conserved AI-2-binding site.

To determine the functionality of these proteins, in vitro AI-2-binding studies using Vibrio harveyi strain MM32 (a reporter strain for AI-2 activity) were conducted (10). Candidate receptors from C. saccharobutylicum and C. autoethanogenum were expressed in AI-2-producing E. coli strains, purified, and heat-denatured to release the bound ligand. The ligands released from the putative receptors induced light production in V. harveyi at levels similar to that of our positive control, ligand released from B. anthracis LsrB (Fig. 2B). As a negative control, receptors were also expressed in a LuxS−E. coli strain (i.e. non-AI-2–producing); as predicted these proteins elicited no response from V. harveyi showing that the response is AI-2–dependent (Fig. 2B). Given this demonstration of AI-2–binding ability by the candidate proteins, we concluded that they are functional AI-2 receptors. We named these receptors CsLsrB for C. saccharobutylicum LsrB and CaLsrB for C. autoethanogenum LsrB.

AI-2 internalization in C. saccharobutylicum

As C. saccharobutylicum and C. autoethanogenum have orthologs for the Lsr transporter system, we tested whether accumulation and internalization of AI-2 was similar to previously studied bacteria with the lsr operon (13, 15, 16). Due to the similarity of the putative LsrBs identified in C. saccharobutylicum and C. autoethanogenum, we focused our studies on C. saccharobutylicum because the strain is more amenable to manipulation. Cell-free supernatants from C. saccharobutylicum cultures were collected at different time points post-inoculation and the activity of extracellular AI-2 was assayed via a V. harveyi MM32 luminescence assay (Fig. 3A). We observed that the levels of extracellular AI-2 increased until a certain threshold level was reached, then rapidly decreased starting at mid-exponential growth. To further support the observation that C. saccharobutylicum is capable of internalizing AI-2, we added synthetic AI-2 to cultures of C. saccharobutylicum and E. coli at time 0. We then compared the concentration of extracellular AI-2 present in C. saccharobutylicum supernatants to that in supernatants of the E. coli mutant that can neither produce or internalize AI-2 (ΔlsrKΔluxS mutant) (Fig. 3B). As this E. coli mutant is not capable of internalizing AI-2, any decrease in AI-2 activity in the supernatants from these E. coli cultures would be due to degradation of AI-2 in these conditions. At 6 h, supernatants of C. saccharobutylicum only stimulated ~280-fold induction, whereas supernatants of E. coli ΔlsrKΔluxS...
caused a 2000-fold induction, supporting the conclusion that \textit{C. saccharobutylicum} depletes AI-2 from the culture by internalization. Moreover, AI-2 is not degraded extracellularly, as no significant degradation was observed in the course of the 6-h incubation of 40 \(\mu\)M AI-2 in cell-free supernatants (Fig. S1A) collected from a late stationary phase culture of \textit{C. saccharobutylicum} (culture shown in Fig. S1B). Thus, we concluded that \textit{C. saccharobutylicum} is able to internalize AI-2 as previously observed for other LsrB receptors.

\textbf{C. saccharobutylicum LsrB binds R-THMF}

To determine the form of AI-2 recognized by \textit{C. saccharobutylicum} LsrB and \textit{C. autoethanogenum} and the identity of the AI-2–binding residues, we determined the crystal structure of \textit{C. saccharobutylicum} LsrB. This receptor was expressed in a \textit{LuxS}\(^+\) \textit{E. coli} strain and crystallized; the structure was solved at 1.35 Å resolution by molecular replacement omitting the ligand from the molecular replacement model (Table S3). The crystal structure shows that \textit{C. saccharobutylicum} LsrB has a classic periplasmic binding protein-fold like the canonical LsrB receptors from \textit{Salmonella} serovar Typhimurium (7), \textit{Salmonella} enterica serovar Typhi (PDB ID 5GTA), \textit{Sinorhizobium meliloti} (36), \textit{Yersinia pestis} (37), and \textit{B. anthracis} (PDB ID 4PZ0), with two \(\alpha/\beta\) domains connected through a three-stranded hinge. The binding site is located near this hinge in the cleft between the two domains (Fig. 4A). Superimposition of the \textit{C. saccharobutylicum} structure with that of \textit{B. anthracis} shows that the structures of the two proteins are very similar (root mean square deviation 0.78 Å (Fig. 4A)). After refinement of the structure, clear electron density showed the ligand to be R-THMF, the same form of AI-2 recognized by the canonical LsrB receptors.

\textbf{Figure 2. LsrB-like receptors from \textit{C. saccharobutylicum} and \textit{C. autoethanogenum} bind AI-2.} A, comparison of the putative AI-2 transport and processing orthologs in \textit{C. saccharobutylicum} (top) and \textit{C. autoethanogenum} (bottom) with \textit{B. cereus} (middle), which has a functional Lsr transporter system. B, AI-2 binding was assessed by measurement of light production of \textit{V. harveyi} MM32 after addition of ligand released from the pure putative LsrBs from \textit{C. saccharobutylicum} and \textit{C. autoethanogenum} expressed in a LuxS\(^+\) \textit{E. coli} BL21 strain (white bars). The LsrB receptor from \textit{B. anthracis} was used as a positive control. As negative controls we tested the mentioned proteins expressed in a LuxS\(^-\) mutant (black bars). Results are shown as fold-induction relative to the light production induced by the growth medium.

\textbf{Figure 3. \textit{C. saccharobutylicum} LsrB receptor internalizes AI-2.} A, AI-2 internalization profile of \textit{C. saccharobutylicum} (black circles) and respective growth curve (gray circles). B, AI-2 internalization profile of \textit{C. saccharobutylicum} (black circles) and \textit{E. coli} \(\Delta\text{lsrK}\Delta\text{luxS}\)} (open squares) after addition of 40 \(\mu\)M AI-2 at time 0. Cell-free supernatants were collected at the indicated time points and activity of extracellular AI-2 was measured by assessing light produced by \textit{V. harveyi} MM32 in response to cell-free supernatants. AI-2 activity is reported as fold-induction relative to light production induced by the growth medium. The internalization curves are representative of three independent experiments shown in Fig. S2. Error bars represent the standard deviation of three technical replicates. Given the different lag phases of \textit{C. saccharobutylicum} growth it was not possible to join the biological replicates.
The crystal structure also revealed that, as predicted by Phyre2, the AI-2–binding residues in CsLsrB (and, presumably by extension, CaLsrB) differ from those of the canonical LsrBs in two positions: a proline replaced with an asparagine and an alanine substituted with a serine (P225N, A227S numbering using 
\textit{B. anthracis} LsrB as reference; compare Fig. 4, B with C). The structure has been deposited in the PDB under the description “LsrB from \textit{Clostridium saccharobutylicum} in complex with AI-2” with the PDB code 6DSP.

\textbf{The LsrB-binding site is more plastic than previously thought}

The observation that CsLsrB binds R-THMF with two of the six canonical AI-2–binding residues altered, led us to investigate whether other amino acid residues could be altered without loss of AI-2–binding ability. We performed alanine scanning mutagenesis on each of the six AI-2–interacting residues in the binding site (Lys-29, Asp-110, Asp-167, Gln-168, Asn-223, Ser-225). To assess AI-2 binding of these mutants, each alanine mutant protein was expressed, purified, and tested for AI-2–binding ability via the \textit{V. harveyi} MM32 assay. We observed that alanine substitutions in three of the four amino acid residues that bind AI-2 through side chains (Lys-29, Asp-110, and Gln-168) caused loss of AI-2–binding ability and are thus necessary for binding (Fig. 5A). The exception to this observation was Asp-167, which is thought to make a hydrogen bond with AI-2 through the carboxylic group of the γ carbon and did not cause CsLsrB to lose its ability to bind AI-2 when substituted by an alanine.

A D167N substitution was also constructed because the asparagine mutation at this site was previously observed as one of two recurring amino acid substitutions in putative LsrB homologs that failed to bind AI-2 (13); AI-2 binding was abolished in this mutant (Fig. 5A). The second substitution commonly observed in nonfunctional LsrB orthologues is a threonine at position 225. As CsLsrB has a serine in that position, we performed the single S225T substitution and observed that AI-2 binding was abolished (Fig. 5).

We also tested whether reverting the noncanonical binding site residues to the canonical ones present in previous characterized LsrBs would hinder AI-2 binding by replacing the serine at position 225 with an alanine (S225A) and the asparagine at position 223 with a proline (N223P). Both mutants were able to bind AI-2, showing that the binding site of the CsLsrB receptor is plastic enough to accept reversion to the canonical LsrB-binding site composition. Altogether, these results allow us to conclude that the residues interacting with AI-2 in the LsrB-binding site are not constrained to a single composition, as previously thought.

\textbf{LsrB receptors bind AI-2 with high affinity}

Both the canonical LsrB proteins and the noncanonical CsLsrB studied here share the fold with the proteins of the large
family of substrate-binding proteins. Proteins of this family usually bind their ligands with high affinity; in fact, the affinity is sufficiently strong as to allow their purification with bound substrates (38). Dissociation constants in this family of proteins and their cognate ligands typically range between 0.1 and 1 μM for sugars and around 0.1 μM for amino acids (38). Thus, we would expect that LsrB receptors bind AI-2 with a dissociation constant in the submicromolar range. Supporting this hypothesis, CsLsrB purified and crystallized with AI-2 bound (above). To quantify this interaction, we used isothermal titration calorimetry (ITC) to determine the affinity of CsLsrB for AI-2 and assess-binding proteins and are therefore better genetic markers for tracing the evolution of the genes in this operon (40, 41). The evolution of LsrA sequences corroborated all major conclusions obtained with LsrB (Fig. S3). Although some local differences in topology and bootstrap support between the two trees do occur, these differences are likely due to the different rates of protein evolution and do not cause conflict with the interpretation of the evolution of C. saccharobutylicum Lsr genes. The inferred evolution of LsrA sequences also corroborates the close relationship of the same species of Spirochaetes and Actinobacteria with the Clostridiales, which strongly suggests that events of operon lateral transfer are responsible for the observed taxonomic distribution of Lsr genes.

**CsLsrB shares a common ancestor with canonical LsrB receptors**

The identification of a functional LsrB receptor with naturally occurring variations in the amino acid residues of the binding site led us to inquire if CsLsrB arose from the same ancestor as the canonical LsrBs. To answer this question, a phylogenetic analysis was performed by querying the protein sequence of CsLsrB against the UniProt database of Reference Proteomes. Our analysis indicates that the large majority of LsrB proteins are distributed between two unrelated phyla, the Proteobacteria, particularly in Gammaproteobacteria and Alphaproteobacteria and the Firmicutes, from the orders Clostridiales and Bacillales (Fig. 7). The LsrB sequence of C. saccharobutylicum clusters in close proximity with other sequences from Clostridiales and in the same evolutionary lineage as the LsrB proteins from the B. cereus group previously shown to have functional AI-2 receptors (Fig. 7). There is clear evidence of lateral gene transfer events in the evolution of LsrB sequences, which explains the disjointed presence of LsrB orthologs within the two unrelated phyla (Proteobacteria and Firmicutes), as well as the presence of LsrB sequences from Actinobacteria and Spirochaetes clustering within a clade largely composed of Firmicutes species.

To obtain further support for the evolutionary position of CsLsrB, we performed a similar phylogenetic analysis focusing on the ATP-binding proteins (LsrA), a protein typically encoded by a gene co-transcribed with lsrB as part of the lsr operon. Given that the function of these two proteins is usually associated, we expected their evolutionary histories to be congruent. Importantly, the ATP-binding proteins from the ABC transporters are known to evolve more slowly than the substrate-binding proteins and are therefore better genetic markers for the interpretation of the evolution of C. saccharobutylicum Lsr genes. The inferred evolution of LsrA sequences also corroborates the close relationship of the same species of Spirochaetes and Actinobacteria with the Clostridiales, which strongly suggests that events of operon lateral transfer are responsible for the observed taxonomic distribution of Lsr genes.

**Figure 5. Determination of the amino acid residues essential for binding of AI-2 by C. saccharobutylicum LsrB.** A, binding of AI-2 to WT and binding-site mutants of CsLsrB. Binding assessed by measurement of light production of V. harveyi/MM32 after addition of ligands released from purified proteins by thermal denaturation. Proteins were expressed either in LuxS (3) or LuxS (3) E. coli BL21 strains. Results are shown as fold-induction relative to the light production induced by the growth medium. B, two-dimensional representation of the binding site of CsLsrB. Substitutions of aspartate 167 or serine 225 (which interfere with AI-2 binding) are boxed in red, whereas substitutions that allow AI-2 binding are boxed in green.
To determine which organisms have LsrB proteins with canonical (KDDQPA) or noncanonical (KDDQNS) predicted binding sites, we analyzed sequence conservation at the binding sites across species as shown in Fig. 7. Interestingly, some of the protein sequences in this phylogenetic tree have a serine instead of an alanine in position 225, as in CsLsrB, but retain the proline in position 223 observed for canonical receptors (KDDQPS) (Fig. 7). According to the results of our mutagenesis studies we infer that these noncanonical receptors with one substitution in the predicted AI-2–interacting residues will also be functional AI-2 receptors because the CsLsrB N223P mutant was able to bind AI-2 (Fig. 5A).

Overall, the evolutionary history of the LsrB receptors supports the conclusion that lateral gene transfer is responsible for the appearance of canonical LsrB receptors in the Firmicutes, which subsequently evolved to noncanonical receptors in a stepwise manner accumulating one, then two, substitutions in the AI-2–binding residues. Most noncanonical receptors cluster together, but events of lateral gene transfer are again evident, as members from the Spirochaetes and Actinobacteria...
Al-2 receptors in Clostridia

LsrB receptors are known to have acquired noncanonical LsrB receptors with one or two substitutions.

Identification of additional putative noncanonical LsrB receptors

Following the phylogenetic analysis of the LsrBs, it became clear that the noncanonical LsrB receptors were present in organisms other than *C. saccharobutylicum* and *C. autoetha-nogenum*. For this reason, we conducted a comprehensive bioinformatic analysis of proteomes currently available in the NCBI RefSeq nonredundant protein database using the sequence of CsLsrB to further identify organisms that possess noncanonical LsrB proteins with one (KDDQPS) or two substitutions (KDDQNS) in the putative binding site.

We identified 95 noncanonical putative LsrB receptors of which 26 have one substitution and 69 have two substitutions, as CsLsrB. The organisms encoding these proteins belong to the same phyla identified in our phylogenetic analysis with the large majority belonging to the Firmicutes phylum. To characterize these hits, all 95 protein sequences were submitted to the fold recognition server Phyre2. The predicted structures for all putative LsrB hits were consistent with the class I periplasmic binding protein fold shared by the canonical LsrBs and the noncanonical CsLsrB (Table S4). We also determined whether these hits had Lsr transport proteins next to LsrB in their genomes (Table S4). From the 95 hits, only the LsrB proteins from *Treponema primitia* and *Treponema azotonutriciu* did not have Lsr transport proteins next to LsrB protein (Table S4). In these genomes the LsrB homolog is located near a hybrid sensor histidine kinase/response regulator, although homologs for Lsr proteins were identified in another location in the genome (Table S4). The presence of the enzymes involved in AI-2 metabolism was also assessed. Even though LsrF was present in 93 of the 95 hits, LsrG was only present (with a query cover higher than 75%) in 9 proteomes. Finally, we checked the proteomes for the existence of the AI-2 synthase LuxS. We did not identify LuxS homologs in all these organisms (79 LuxS homologs in 95 hits; Table S4). However, this is also true for *Sinorhizobium meliloti*, a species that carries a functional *lsr* operon but is unable to produce AI-2 (36). Altogether, these results suggest that functional, noncanonical LsrB receptors are present in many bacterial species, particularly in the Firmicutes phylum.

Discussion

LsrB AI-2 receptors share an identical fold to substrate-binding proteins that recognize a wide variety of substrates including 5-carbon sugars similar to AI-2. Thus, the identification of LsrB receptors has been possible due to the conservation of the six amino acid residues that bind AI-2, which allows the distinction between LsrB receptors and other substrate-binding proteins. In this study, we provide the first report of a functional LsrB receptor in Clostridia; this receptor has a previously unobserved binding site variation and gives insights into the plasticity of the binding site of LsrB receptors. The crystal structure of this receptor identified in *C. saccharobutylicum*, CsLsrB, showed that it binds an R-THMF form of AI-2 with two variations in the amino acid residues that bind AI-2: an asparagine instead of a proline (P223N) and a serine instead of an alanine (A225S) (Fig. 4, B and C). The fact that the substitution of the nonpolar proline and alanine for the polar asparagine and serine do not abolish AI-2 binding was surprising both due to the dissimilar nature of the amino acids and the fact that the canonical AI-2–binding residues were thought to be essential. However, examination of both canonical and noncanonical AI-2–binding pockets reveals that the amino acids in these positions bind AI-2 through the backbone. Moreover, the crystal structure of CsLsrB shows that the amino and carboxyl groups are still positioned to form hydrogen bonds with similar geometry to canonical receptors; thus the local region must be plastic enough to accept side chain variation without significantly changing the overall structure of the protein. Still, not all amino acid residues can be accommodated in these positions as the S225T substitution caused loss of AI-2 binding. The side chain hydroxyl group of the threonine would possibly be shifted by the presence of the additional methyl group, causing a change in geometry that would likely impact the interaction with the ligand. In addition, this shift could alter the interaction between S225T and N33, potentially causing a conformational change that impacts the surrounding amino acids, altering the conformation of the binding site further (Fig. 5). As for D167N, which similarly to S225T, is a naturally occurring substitution in sugar-binding proteins with homology to LsrB, the inability to bind AI-2 might be a result of the change from a presumably negatively charged side chain to a side chain containing an amine. This could lead to unfavorable interactions with the ligand or the adjacent Gln-168, which also binds AI-2 through side chain interactions. Moreover, the introduction of the amine in this mutation could lead to more fundamental conformational changes in the binding site and surrounding environment than might be expected for a simple alanine mutation, explaining the loss of binding activity in the *Asn* but not the Ala mutant (Fig. 5). Accordingly, S225T and D167N substitutions also caused loss of AI-2 binding ability in *B. anthracis* LsrB (13), a canonical receptor, indicating that the presence of a threonine in position 225 and an asparagine in position 167 might be characteristic of sugar-binding proteins.

The two variations in AI-2–binding residues of CsLsrB might explain the slightly lower affinity of this receptor to AI-2 when compared with the canonical LsrBs tested. Nevertheless, CsLsrB bound AI-2 with high affinity (submicromolar *Kₐ*; Fig. 6). Interestingly, in a previous study by Zhu and Pei (42) a significantly higher *Kₐ* (160 μM) was reported for the binding of...
the LsrB from *Salmonella* serovar Typhimurium to AI-2. However, our results were obtained in the absence of boron and, as was shown in Miller *et al.* (7), the removal of boron is essential to avoid shifting the equilibrium of AI-2 molecules toward the S-THMF-borate form. Hence, the higher apparent *Kd* value obtained by Zhu and Pei (42) might result from the presence of boron. Moreover, the authors reported a *Kd* of 0.16 μM for LuxP from *V. harveyi* (42), a value closer to the ones we obtained for the LsrBs tested here and more consistent with the values expected for substrate-binding proteins toward their cognate ligands (38).

Similarly to canonical LsrB receptors, CsLsrB was able to internalize AI-2 (Fig. 3), likely through interaction with the ABC transporter encoded in the same operon as *lsrB* (Fig. 2A). Curiously, although AI-2 uptake started at mid-exponential phase (Fig. S1B) earlier than what is usually observed for *E. coli* K12 MG1655 that starts internalization in the transition from exponential to stationary phase (Fig. S1C and see Refs. 9 and 13). Additionally, *C. saccharobutylicum* supernatants induced 10-fold less light production, indicating that it accumulates lower concentrations of extracellular AI-2. The observed difference is likely because *C. saccharobutylicum* starts AI-2 uptake at lower extracellular AI-2 concentrations indicating that, at least under the laboratory conditions tested, *C. saccharobutylicum* has a lower concentration threshold than *E. coli*. This might indicate that the AI-2 concentration needed for induction of the *lsr* operon in *C. saccharobutylicum* is lower than in *E. coli* K12 MG1655. As *C. saccharobutylicum* and *E. coli* LsrB receptors have high affinity to AI-2, it is unlikely that differences in affinities of the receptors cause the variation in concentrations required for the induction of the operons and, thus, in the start of internalization. Moreover, *E. coli* can accumulate extracellular AI-2 to concentrations up to 40 μM without starting to internalize AI-2. This indicates that, in *E. coli*, additional mechanisms are preventing this bacterium from starting internalization of AI-2 at lower concentrations (43). In fact, in *E. coli*, the Lsr system is regulated not only by LsrR but also by catabolite repression and the phosphotransferase system, such that the availability of phosphotransferase system substrates inhibits Lsr-mediated AI-2 internalization (43, 44). Thus, it is possible that these additional regulatory mechanisms are either different or not present in *C. saccharobutylicum*, allowing internalization to start at lower concentrations of AI-2. Further work on the regulation of these systems is necessary to address this hypothesis. The search for new noncanonical receptors revealed that in most organisms where putative noncanonical LsrB receptors were identified, the *lsrB* neighboring genes encoded homologs for the components of the Lsr transport system. In contrast, homologs of the AI-2–processing enzymes, LsrG and LsrF, identified in *E. coli* and *Salmonella* serovar Typhimurium, were not as widely conserved. For example, in *C. saccharobutylicum* no LsrF ortholog was identified and in both *C. saccharobutylicum* and *C. autoethanogenum*, the gene that encodes the LsrG orthologs is present in another region of the genome (Fig. 2A). These results, together with previous reports of canonical LsrB receptors encoded in operons that also encode functional Lsr transport systems but have one or both LsrG and LsrF orthologs located in other regions of the genome (13, 36) indicate that AI-2 metabolism might not be as conserved as the steps required for the uptake, phosphorylation, and regulation of the *lsr* operon. Thus, the lack of orthologs for the proteins responsible for AI-2 metabolism (LsrG and LsrF) should not be an excluding factor when searching for new LsrB receptors. Curiously, in *T. primitia* and *T. azonutricium* we identified putative noncanonical LsrBs that were near a sensor kinase instead of an ABC transporter in the genome. LuxP, the other type of known AI-2 receptor known, interacts with a sensor kinase (LuxQ) that facilitates signal transduction through a phosphorylation cascade (2). However, when searching for homologs of the Lsr proteins, we found hits for all the proteins, except LsrR, located in a more distant region of the genome (Table S4). Hence, further studies are needed to understand if these putative LsrBs evolved to interact with the nearby sensor kinase or if they instead use an ABC transporter located in another region of the genome.

In summary, the characterization of the amino acid variation in the AI-2–binding site of CsLsrB fulfills our initial motivation as it allowed the identification of new putative LsrB receptors in microbes present in biologically relevant niches like *Roseburia inulinivorans* (45) and *Clostridium merdae* (46). Firmicutes isolated from human microbiota that appear to have a noncanonical receptor with two substitutions (like CsLsrB). This is a significant step toward understanding the molecular mechanisms involved in interspecies communication in these niches. In particular, we expect the identification of AI-2 receptors in Firmicutes to help in the characterization of the mechanism by which their colonization is favored in the presence of AI-2 in the mammalian gut microbiome, a clinically relevant niche where interspecies interactions are highly prevalent.

**Experimental procedures**

**Identification of LsrB orthologs**

The search for LsrB orthologs was performed as described previously (13). Briefly, we searched the complete genomes in the KEGG SSDB (Sequence Similarity Database, 3478 bacterial genomes in September 2015) for amino acid sequences with similarity to LsrB from *B. cereus*. All pairwise genome comparisons were performed using SSEARCH program and best bidirectional hits with a Smith–Waterman score of at least 120 were selected. From the obtained hits with sequence identity between 30 and 60% we focused on the putative LsrBs of *C. saccharobutylicum* and *C. autoethanogenum*. The genome context of the CsLsrB and CaLsrB genes was assessed through analysis of the neighboring genes. To determine the presence/absence of Lsr orthologs in *C. saccharobutylicum* and *C. autoethanogenum*, we looked for the best bidirectional hits using the Lsr proteins of *B. cereus* ATCC 10987 and *B. anthracis* Sterne as query. Simultaneously, we determined their similarity (Table S2). The amino acid sequences of the putative LsrBs from *C. saccharobutylicum* and *C. autoethanogenum* were submitted to the Phyre2-fold recognition server. The presence/absence of conserved amino acid residues in AI-2 binding was determined through structural alignment of the structures pre-
Al-2 receptors in Clostridia

dicted by Phyre2 with the LsrB from B. anthracis (PDB ID 4PZ0) using COOT (47).

Al-2 synthesis

The Al-2 precursor DPD was synthesized as previously described (48). Boron-free DPD was obtained by performing the synthesis with plastic material and using boron-free water (prepared by batch incubation with Amberlite IRA743 resin for 2 h at room temperature, as described before (49)).

Al-2–internalization assays

Aerobic frozen stocks of C. saccharobutylicum DSM13864 were revived on modified PY + X solid medium with galactose (see Table S1 for media composition). 13 ml of broth (modified PY + X with arabinose and 100 mM MOPS buffer, pH 7.0) was inoculated with 3–6 colonies of C. saccharobutylicum and grown overnight at 37 °C in an anaerobic chamber (PlasLabs, USA) under a gaseous mix of 80% nitrogen, 15% carbon dioxide, and 5% hydrogen to an A600 nm lower than 4. A 20% (v/v) inoculum was added to fresh medium and incubated at 37 °C without agitation in a 100-ml Hungate Schott flask. Culture suspensions were collected at the stated time points for optical density measurement at 600 nm (UV-visible spectrophotometer; Helios Delta, ThermoSpectronic) and for detection of Al-2.

E. coli ΔluxSΔluxK (E. coli ARO093) (29) was revived directly from the frozen stock to liquid medium and incubated at 22 °C. At this point, the temperature was decreased to 2 °C. At A595 = 0.9, 0.3 mM isopropyl 1-thio-β-D-galactopyranoside was added and the cells were induced for 6 h before being harvested by centrifugation. Cells were resuspended in 50 mM NaH2PO4 (pH 8.0), 300 mM NaCl, 10 mM imidazole, 2.5 μg/ml of Dnase, and 2.5 μg/ml of leupeptin and lysed using a M-110Y microfluidizer (Microfluidics, USA). The lysate was centrifuged and the tagged protein was purified from the clarified supernatants using nickel–nitrilotriacetic (Ni3+-NTA) acid affinity chromatography (Qiagen). The protein was eluted from the column in 50 mM NaH2PO4 (pH 8.0), 300 mM NaCl, and 250 mM imidazole and subsequently swapped into 25 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 1 mM dithiothreitol (DTT) using Sephadex G25-agarose. The His6 tag was removed by cleavage with tobacco etch virus (TEV) protease at a proportion of 1 mg of TEV protease per 250 mg of protein at 4 °C, overnight. A second round of Ni3+-NTA affinity chromatography was performed to remove the His6-tagged TEV protease, the cleaved tag, and any uncleaved fusion protein. The protein collected from the flow-through was buffer swapped into 25 mM Tris-HCl (pH 8.0), 50 mM NaCl, and 1 mM DTT via Sephadex G25-agarose and further purified by anion exchange chromatography (SourceQ column, GE Healthcare Life Sciences) using a NaCl gradient from 0 to 1 M. As a final purification step, the protein underwent size exclusion chromatography on a Superdex 75 (GE healthcare) column and was eluted in 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM DTT.

Crystallization studies

Crystals of C. saccharobutylicum LsrB expressed in E. coli BL21 LuxS+ were grown via the sitting drop method with a well solution of 0.1 M citric acid (pH 2.75) and 26% (w/v) PEG 3350 and developed in approximately 1 week at room temperature. Crystals were frozen after a 30-s soak in a solution of 27% (w/v) PEG 3350 plus 15% (v/v) glycerol. Diffraction data were collected on beamline BL14-1 at the Stanford Synchrotron Radiation Light Source. Data were processed with CCP4 software (50). A molecular replacement solution was determined via PHENIX (51) using PDB 1TJY as the search model. The initial model was built by PHENIX with subsequent manual building in Coot (47) and refinement in PHENIX. The resulting PDB file is used as a reference for amino acid numbering of CsLsrB throughout this article. Thus based on the structure, the amino acids of the binding site were numbered Lys-29, Asp-110, Asp-167, Gln-168, Asn-223, Ser-225, whereas the numbering based on the complete amino acid sequence of the protein was Lys-48, Asp-129, Asp-186, Gln-187, Asn-242, Ser-244.

Site-directed mutagenesis

The QuikChange Lightning Site-directed Mutagenesis kit (Agilent) was employed to make the single amino acid substitutions in pDEST527/C. saccharobutylicum LsrB constructs. Primers were designed using the QuikChange primer design program (Agilent) and sequences are given in Table S5. The
mutant proteins were expressed as described above. The His₆-tagged mutants were purified through Ni²⁺-NTA affinity chromatography (Qiagen) as above and then buffer swapped into 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 1 mM DTT using PD-10 desalting columns (GE Healthcare). The protein was concentrated in 10-kDa Centricons (Millipore) until a concentration of at least 10 mg/ml was achieved.

In vitro AI-2–binding assay

The in vitro AI-2–binding ability of the studied proteins was assessed as previously described (13). Briefly, LsrB proteins at a concentration of 10 μg/ml were heat-treated at 70 °C for 10 min. The denatured protein was pelleted via centrifugation and 10 μl of the supernatant were added to 90 μl of a 1:5000 dilution of an overnight culture of V. harveyi MM32, as described above. Bioluminescence after 5 h of incubation at 30 °C was measured. A 1420 Victor 2 plate reader (PerkinElmer Life Sciences) and a 1420 Victor 3 plate reader were used for data in Figs. 2 and 5, respectively. Binding data are representative of three independent experiments and the standard deviations are derived from three technical replicates. Propagation of uncertainty was used to calculate the standard deviation after normalizing the light produced by AI-2 by the light produced by buffer alone.

Isothermal titration calorimetry

ITC measurements were performed in a MicroCal iTC200 microcalorimeter (GE Healthcare Biosciences) at 25 °C. Boron-free AI-2 was diluted in boron-free buffer containing 25 mM sodium phosphate buffer (pH 8.0), 150 mM NaCl, 1 mM β-mercaptoethanol. Amberlite IRA743 resin was employed to remove the boron from the water used to prepare buffer as previously described (49). To avoid contamination by boron silicates present in glass, only plastic material was used. AI-2 at 800 μM was injected into 117.4 and 108 μM LsrB protein from C. saccharobutylicum and E. coli K-12 MG1655, respectively. For B. anthracis Sterne 34F2 LsrB, a solution of 750 μM LsrB was added to 106.9 μM protein. As a negative control, 3 mM AI-2 was injected into 114.7 μM of the LsrB ortholog from R. etli CNF42 (RHE-PE00289 in KEGG, annotated as substrate-bind-strap using an automatic frequency-based criterion (autoFC option) to determine the number of replicates. We identified two truncated proteins, one from Clostridium magnum str. DSM 2767 and the other from Klebsiella pneumoniae str. ISC21, which were eliminated from final analyses. The phylogenetic analysis of the 180 protein sequences homologous to CsLsrB indicated a clear separation between bona fide LsrB sequences and the paralogous sequences identified as rhamnose-binding proteins (Fig. S4). This result was used to define a smaller dataset focused on exclusively LsrB sequences keeping all sequences within the most inclusive and highly supported clade that includes all confirmed functional LsrB sequences and our target LsrB sequence from C. saccharobutylicum. This final dataset of 97 protein sequences was realigned and reanalyzed as before. All sequences included have an e-value smaller than 10⁻⁶³ and an average sequence identity of 43%. This smaller dataset produced a better alignment and a more accurate phylogenetic inference. For the phylogenetic study of LsrA, we performed a homology search by querying the ATP-binding protein of C. saccharobutylicum against the UniProt database of Reference Proteomes using phmmer algorithm. We extracted all significant hits using an e-value threshold of 10⁻⁹⁵, which retrieved 470 sequences. To this dataset we added all ATP-binding protein sequences that pertain to the same strain/operon as the functional LsrB (11 sequences), if not previously included, as well as the two ATP-binding proteins from the rhamnose-binding operon (Table S8). These sequences were analyzed as explained above for LsrB sequences to produce a phylogenetic analysis of bona fide LsrA sequences present in the Reference Proteome database (Fig. S3).

Identification of noncanonical LsrBs

The search for putative noncanonical LsrB receptors was accomplished using BLASTP against the NCBI RefSeq non-
**AI-2 receptors in Clostridia**

dundant protein database using CsLsrB as query sequence (in May 2018). RefSeq allows sequence submission and includes complete and incomplete genomes, thus having more genomes available than KEGG Genome, which is a collection of organisms with complete genome sequences or than UniProt database of Reference Proteomes, which is a balanced database to inference evolutionary history \((56, 57)\). The alignments between the amino acid sequence of CsLsrB and the hit proteins were examined to determine the identity of the six amino acid residues involved in AI-2 binding. To simplify the analysis, the search was conducted for each one of the 35 phyla described in KEGG separately. From the functional mutations identified in this study, only putative LsrBs with N223P and N223P/S225A were found. The annotation of the proteins in the vicinity of these noncanonical hits was assessed. For the hits with nearby Lsr orthologs, the amino acid sequences of these proteins were submitted to a comparative BLASTP with the respective ortholog in *C. saccharobutylicum*. For the absent orthologs and for LuxS, a BLASTP directed to the proteome of the organisms using the matching ortholog in *C. saccharobutylicum* as query was performed. As no LsrF homolog was identified in *C. saccharobutylicum*, the amino acid sequence of LsrF from *E. coli* K12 MG1655 was employed. Additionally, very few hits for LsrG homologs were obtained using the *C. saccharobutylicum* LsrG as query, so we also checked for homologs using the biochemically characterized LsrG from *E. coli* K12 MG1655. All of the 95 putative LsrB sequences were submitted to the Phyre2-fold recognition server.


**Acknowledgments**—We thank Ana Rita Oliveira and Roberto Balbontin for helpful suggestions, Oswaldo Ascenso, Vanessa Miranda, and Rita Ventura for the boron-free DPD, Steven Chen and Kendall Yoshii for help with the in vitro AI-2 binding assay of CsLsrB and CaLsrB, and Paula Chieco and Marco Domingues for technical support with ITC.

**References**


