Genomic Rearrangements and Functional Diversification of lecA and lecB Lectin-Coding Regions Impacting the Efficacy of Glycomimetics Directed against Pseudomonas aeruginosa

Amine M. Boukerb 1, Aude Decor 2, Sébastien Ribun 1, Rachel Tabaroni 2, Audric Rousset 3, Loris Commin 4, Samuel Buff 4, Anne Doléans-Jordheim 1, Sébastien Vidal 3, Annabelle Varrot 2, Anne Imberty 2 and Benoit Cournoyer 1*

1 Equipes de Recherche, Bactéries Pathogènes Opportunistes et Environnement, Centre de Ressources Biologiques – Environnement Microbiologie Lyon, UMR Centre National de la Recherche Scientifique 5557 Ecologie Microbienne, Université Lyon 1 and VetAgro Sup, Lyon, France, 2 Centre de Recherche sur les Macromolécules Végétales (UPR 5301), Centre National de la Recherche Scientifique et Université Grenoble Alpes, Grenoble, France, 3 Laboratoire de Chimie Organique 2 – Glycochimie, Institut de Chimie et Biochimie Moléculaires et Supramoléculaires, UMR Centre National de la Recherche Scientifique 5246, Université Lyon 1, Lyon, France, 4 Université de Lyon, VetAgro Sup, UPSP 2011-03-101, Interactions Cellules Environnement and CRB-ANIM (ANR-INBS11-0003), Marcy-L’Étoile, France

LecA and LecB tetrameric lectins take part in oligosaccharide-mediated adhesion-processes of Pseudomonas aeruginosa. Glycomimetics have been designed to block these interactions. The great versatility of P. aeruginosa suggests that the range of application of these glycomimetics could be restricted to genotypes with particular lectin types. The likelihood of having genomic and genetic changes impacting LecA and LecB interactions with glycomimetics such as galactosylated and fucosylated calix[4]arene was investigated over a collection of strains from the main clades of P. aeruginosa. Lectin types were defined, and their ligand specificities were inferred. These analyses showed a loss of lecA among the PA7 clade. Genomic changes impacting lec loci were thus assessed using strains of this clade, and by making comparisons with the PAO1 genome. The lecA regions were found challenged by phage attacks and PAGI-2 (genomic island) integrations. A prophage was linked to the loss of lecA. The lecB regions were found less impacted by such rearrangements but greater lecB than lecA genetic divergences were recorded. Sixteen combinations of LecA and LecB types were observed. Amino acid variations were mapped on PAO1 crystal structures. Most significant changes were observed on LecBPAT7, and found close to the fucose binding site. Glycan array analyses were performed with purified LecBPAT7. LecBPAT7 was found less specific for fucosylated oligosaccharides than LecBPAT1, with a preference for H type 2 rather than type 1, and Lewisα rather than LewisX. Comparison of the crystal structures of LecBPAT7 and LecBPAT1 in complex with LewisX showed these changes in specificity to have resulted from a modification of the water network between the lectin, galactose and GlcNAc residues. Incidence of these modifications on the interactions
INTRODUCTION

Opportunistic infections are of major concern around the world. They can lead to community or nosocomial infections among individuals showing weaknesses in their protective barriers. *Pseudomonas aeruginosa* is one of the main human opportunistic pathogens, and can be found in multiple outdoor habitats leading to human exposures. It can be the etiological agent of several infections such as folliculitis, keratitis, otitis but also pneumopathies in cystic fibrosis patients which can evolve into severe septicemia and death (Forkner et al., 1958).

*P. aeruginosa* is divided into multiple lineages but has a panmictic organization (Kidd et al., 2012; Dettman et al., 2015). *P. aeruginosa* sequenced genomes showing the most significant differences with the classical laboratory strains such as PAO1 and PA14 are those of the PA7 clade (Boukerb et al., 2015; Freschi et al., 2015). This lineage was initially defined as a taxonomic outlier (Roy et al., 2010; Valot et al., 2015), and has diverged early from the other lineages (Gomila et al., 2015; Hilker et al., 2015). The multidrug-resistant PA7 strain was isolated from a patient in Argentina, and found to harbor several pathogenicity islands and putative virulence factors (Roy et al., 2010).

Among its key properties involved in host colonization, *P. aeruginosa* has selected a “glyco-strategy” favoring oligosaccharide-mediated recognition and adhesion to host cells. This strategy involves carbohydrate binding proteins such as lectins and other adhesins like fimbrial proteins and flagella. The lectin-carbohydrate interactions are characterized by their high specificity and multivalency to generate higher affinity of binding. Two soluble lectins, LecA and LecB, involved in adherence and biofilm formation have been described in *P. aeruginosa*. These lectins act as virulence factors through their carbohydrate binding ability. They promote the adhesion of *P. aeruginosa* to epithelial cells and cause alveolar damages (Chemani et al., 2009; Boukerb et al., 2014). They also inhibit ciliary beating of epithelial cells (Mewe et al., 2005). LecA and LecB bind to galactose and fucose, respectively (Gilboa-Garber, 1982), and their glycan-binding spectra cover a wide range of antigens (ABH, Lewis, P and I systems) present on human tissues or bacterial cell walls (Imberty et al., 2004). Glycomimetics have been synthesized to prevent lectin-dependent adhesion to host cells. Calixarene scaffolds decorated with galactose and fucose were found particularly efficient (Bernardi et al., 2013; Cecioni et al., 2015). Such molecules were found to increase *P. aeruginosa* aggregation and to reduce damages of alveolar tissues (Boukerb et al., 2014).

However, considering the many environmental matrices and living hosts that can be colonized by *P. aeruginosa*, one could hypothesize strong selective pressures on lectins that could lead to modified sugar affinities. Such changes could prevent therapeutic uses of some glycomimetics. In order to identify such events, genetic diversity analyses of lecA and lecB from a large panel of environmental and clinical *P. aeruginosa* strains were conducted. The relation between the observed lec types and sugar-based calixarene glyco-clusters induced aggregations was investigated. Most significant modifications at the lec loci were observed in the PA7 clade. This clade was thus used as a reference to investigate genomic instability around these loci. Three PA7-related genomes were sequenced and annotated for this work (Boukerb et al., 2015). The crystal structure of LecBPA7 was resolved to investigate the incidence of some amino acids changes on the structural organization of the LecB tetramers and sugar binding site. This work showed genomic and genetic changes at the lec loci indicative of ongoing adaptive processes among *P. aeruginosa* which can lead to their loss or changes in the sugar ligand affinities of their encoded lectins.

**Keywords**: *Pseudomonas aeruginosa*, soluble lectins, PA7 clade, region of genomic plasticity (RGP), glycoclusters, flow cytometry, glycan array, crystallography

**MATERIALS AND METHODS**

**Bacterial Strains, Growth Conditions, and DNA Extractions**

A total of 148 *P. aeruginosa* strains were used in this study (Supplementary Tables S1, S2). Environmental strains (*n* = 74) (Lavenir et al., 2007, 2014; Petit et al., 2013) belonging to *P. aeruginosa* were obtained from the EML-BRC (Environmental Microbiology Lyon—Biological Resource Center) collection (http://eml-brc.org) of the French Network of Biological Resource Centers (FBRCMi; www.fbrcmi.fr). Clinical strains (*n* = 70) coming from non-CF infections (*n* = 40; recovered from cases of otitis, urinary tract, and wound infections) were provided by the French “Collège de Bactériologie, de Virologie et d’Hygiène des Hôpitaux” (Paris, France). Isolates from CF patients (*n* = 30) were also included, and previously reported in (Doléans-Jordheim et al., 2009). Reference strains, PAO1 (Holloway, 1955) and PA7 (Roy et al., 2010) and isogenic mutants PAO1ΔlecA and PAO1ΔlecB (Boukerb et al., 2014) were used to validate the PCR screenings and aggregation assays. *P. aeruginosa*...
total DNA was extracted and visualized according to Johnson et al. (1994).

**PCR Screenings and Genetic Diversity Analyses**

MLST analyses were performed according to Curran et al. (2004). PCR primers used to detect lecA and lecB among the *P. aeruginosa* collection were designed using conserved DNA regions among PAO1, PA14, PA7, and LESB58 genomes (Supplementary Table S3). These lecA and lecB PCR screenings were validated on PAO1 and PA14 DNA extracts. PCR reactions were performed in 25 µL using a *Taq* DNA polymerase from Invitrogen (Cergy-Pontoise, France) by following the manufacturer's instructions. Annealing temperatures are indicated in Supplementary Table S3. Each PCR test was, at least, duplicated, and when needed, bacterial colony blots were analyzed to validate the presence or absence of targeted genes. Colony blots DNA hybridizations were performed using radioactively labeled (32P) lecA and lecB PCR products from strain PAO1 as described previously (Pallud et al., 2001). PCR products were visualized after electrophoresis using a 2% agarose gel, staining with ethidium bromide (at 5 mg/L for 15 min) and UV light exposure. PCR products were sequenced by the Biofidal (Villeurbanne, France) sequencing platform. Their DNA sequences were analyzed using the BioEdit software. Similarity and identity searches of the DNA/amino acid sequences were performed using the NCBI website (http://www.ncbi.nlm.nih.gov/).

Multiple alignments of DNA and protein sequences were computed using CLUSTALX v2.1. These alignments were visualized using Seaview (Gouy et al., 2010). BLAST analyses were run at NCBI and insertion sequences were characterized using the IS biotoul website (http://www.is.biotoul.fr/). The phylogenetic relationships between DNA or amino acid sequences were computed by distance matrix methods according to Galtier et al. (1996). Neighbor-joining trees and 1000 bootstrap replicates were generated. The adjusted Wallace coefficients enabling comparisons of partitions between genetic loci (lec and MLST types) were computed according to Severiano et al. (2011) using UMMI web site at Universidade de Lisboa (http://www.comparingpartitions.info/). Homoplasy test was performed using Splitstree4 (Huson and Bryant, 2006). The DNA and amino acid sequences generated in this work have been assigned GenBank accession numbers KT182468 to KT182473.

**Comparative Genomics**

Assembled genomes of three PA7-related strains, i.e., EML528, EML545 and EML548 used in this study have been deposited at DDBJ/EMBL/GenBank under the following accession numbers: LFXX0000000.1 for EML528, LGJE0000000.1 for EML545, and LFXR0000000.1 for EML548 (Boukerb et al., 2015). These genomes were aligned against the PAO1 and PA7 genome sequences using Mauve (v2.3.1; http://asap.genetics.wisc.edu/software/mauve/). The Magnifying Genome platform (MaGe, http://www.genoscope.fr/agc/mage/) was used to compare CDS contents and analyze some other features (Vallenet et al., 2013). Comparative analyses of lecA and lecB genomic regions among PAO1, PA7, and related strains were performed using ACT (Artemis Comparison Tool) at http://www.webact.org. RGP (regions of genomic plasticity) among these regions were detected by looking for synteny breaks due to an integration of DNA pieces above 900 bp. Small indels of a few nucleotides were not considered and not observed among the *lec* loci.

**P. aeruginosa PA7 lecB Sub-Cloning, Over-Expression, and Glycan Specificity Tests**

Cloning of the lecB<sub>PA7</sub> gene into pET25b vector (Novagen) was carried out by a PCR strategy using primers described in Supplementary Table S3. LecB<sub>PA7</sub> over-expression was performed in *Escherichia coli* BL21 (DE3) (Novagen). Cells were grown in LB (Luria-Bertani broth, Sigma-Aldrich, USA) under orbital shaking (200 rpm) at 37°C until reaching an OD<sub>600</sub> (optical density at 600 nm) of 0.6 prior induction with 0.5 mM isopropylthiogalactoside (IPTG) for 3 h at 30°C. Cell lysates were applied onto a D-mannose-agarose column (Sigma-Aldrich, USA) which was equilibrated with a buffer (20 mM TrisHCl pH 7.5, 100 mM NaCl, 100 µM CaCl₂). After washing out unbound proteins, LecB<sub>PA7</sub> was eluted using the same buffer but supplemented with 300 mM D-mannose. The purified protein was dialyzed for 1 week at 4°C. Purified recombinant LecB<sub>PA7</sub> was labeled with Alexa Fluor 488 (Invitrogen, France) according to the manufacturer's instructions and repurified on a D-Salt polyacrylamide desalting column (Pierce, France). Alexa-labeled LecB<sub>PA7</sub> was used for glycan array screening with the standard procedure of the Protein-Glycan Interaction Core (H) of the Consortium for Functional Glycomics, USA. The Glycan Array version 5.2 that contains 609 glycans was screened.

**Microcalorimetry**

Recombinant LecB<sub>PA7</sub> was dissolved in a 20 mM TrisHCl, pH 7.5, NaCl 150 mM, 100 µM CaCl₂ buffer. Carbohydrate ligands were dissolved in the same buffer, degassed, and loaded in the injection syringe. ITC was performed using ITC200 microcalorimeter (Malvern Instruments, UK). The LecB<sub>PA7</sub> solution was placed in a 200 µL sample cell at 25°C. Titration was performed with 20 injections of 2 µL carbohydrate ligands every 120 s. The experimental data were fitted on a theoretical titration curve using the “Origin” software (RITME Informatique, France), with ΔH (enthalpy change), K<sub>d</sub> (association constant) and n (number of binding sites per monomer) as adjustable parameters. Free energy change (ΔG) and entropy contributions (TΔS) were derived from the equation ΔG = ΔH − TΔS = −RT ln K<sub>d</sub> (with T as the absolute temperature and R = 8.314 J mol⁻¹ K⁻¹). Two independent titrations were performed for each tested ligand.

**Protein Crystallography**

Crystals of LecB<sub>PA7</sub> complexed with Lewis<sup>a</sup> tetrasaccharide (Elicityl, Crolles, France) were obtained by the hanging drop vapor diffusion method using 2 µL of drops containing a 50:50 (v/v) mix of protein and reservoir solution at 19°C. The protein at 10 mg/mL in 20 mM HEPES pH 7.5 and 1 µM CaCl₂ was incubated with the ligand at 10 mM during 1 h at
room temperature prior to co-crystallization. Crystal plates were obtained in 2 days from solution containing 14% PEG6000, 0.2 M lithium chloride, and 0.1 M sodium acetate pH 4.6. Plates were transferred in a solution where PEG6000 concentration was increased to 26% for cryoprotection prior mounting in a cryoloop and flash-freezing in liquid nitrogen. Diffraction data were collected at 100 K at the European Synchrotron Radiation Facility (Grenoble, France) on beamline BM30A using a ADSC Q315 CCD detector. The data were processed using iMosflm (Battye et al., 2011). All further computing was performed using the CCP4 suite (Winn et al., 2011). Data quality statistics are summarized in Supplementary Table S4. The structure was solved by molecular replacement using PHASER and the monomer coordinates of 1GZT as search model, according to McCoy et al. (2007). Five percent of the observations were set aside for cross-validation analysis, and hydrogen atoms were added in their riding positions and used for geometry and structure-factor calculations. The structure was refined using restrained maximum likelihood refinement in REFMAC 5.8 (Murshudov et al., 2011) iterated with manual rebuilding in Coot (Emsley et al., 2010). Incorporation of the ligand was performed after inspection of the 2Fo-DFc weighted maps. Water molecules, introduced automatically using Coot, were inspected manually. The stereochemical quality of the models was assessed with the program Molprobyt (Chen et al., 2010), and coordinates were deposited in the Protein Data Bank under code 4UT5.

**Bacterial Aggregation Assays through Imaging Flow Cytometry Analysis**

Aggregation assays were performed according to Boukerb et al. (2014) but aggregates were visualized and analyzed by Amnis FlowSight® cytometry. Image acquisitions and analyses were performed with the ImageStream system (Amnis Corporation, Seattle, WA, USA) and the ImageStream Data Exploration and Analysis Software (IDEAS, Amnis). *P. aeruginosa* cells were exposed to 100 µM glucosylated, mannosylated, galactosylated, and fucosylated glycoconjugates calix[4]arene ligands obtained from Boukerb et al. (2014) (Supplementary Figure S1). Results were compared with those obtained from PBS and monovalent ligands. For each condition, the aggregation properties were analyzed over 20,000 cellular events. Kruskal–Wallis rank sum tests were used to determine conditions favoring the formation of the highest number of bacterial aggregates. P < 0.05 were considered statistically significant. These tests were computed using the R (R Development Core Team, 2015) software.

**RESULTS**

**Prevalence and Diversity Of lecA and lecB**

lecA and lecB databases were built by retrieving the *P. aeruginosa* sequences available at GenBank and at the *Pseudomonas* Genome database (Winsor et al., 2015) and by performing lecA and lecB PCR screenings over a collection of environmental and clinical strains (Figure 1 and Supplementary Figure S2). A total of 292 PCR products from 147 strains were sequenced (Supplementary Tables S1, S2). DNA blot analyses were performed to verify PCR false-negative results (data not shown). Most strains showed PCR products of the expected size. Two strains (poe1196 and poe1293) yielded lecA PCR products of about 2 kb, and were found harboring an IS30p4-like element which is part of the IS30 family (Supplementary Table S1). This IS was inserted at lecA nucleotide position 288 (according to PA2570 numbering). In addition, the multi-drug resistant PA7 strain and strains EML528, EML545, and EML548 were found lacking lecA. These latter strains were found to be part of the PA7 clade (see below).

These data sets allowed the identification of genetic variants that were classified into Lec types according to their amino acid sequences (Supplementary Figure S3). Low divergence was observed between LecA types (<3%). More divergences were observed between LecB ones (<16%). LecA was divided into 17 types (Supplementary Figure S2), and LecB into 14 ones (Figure 1). It is noteworthy that the LecA typings did not perfectly match phylogenetic groupings inferred from MLST DNA sequences (Supplementary Figure S4). Adjusted Wallace coefficient analysis showed that ST defined by MLST will have a same LecA type for 84% of the strains while this probability was of 98% for LecB. LecA types 2 and 4 were found restricted to the PAO1 MLST group but types 1 and 3 were distributed among both the PAO1 and PA14 groups (Supplementary Figures S3, S4). However, the phi test did not find statistically significant evidence for recombination at this locus among *P. aeruginosa* (p = 0.23). These changes are thus likely to have occurred independently. Three significant LecB clusters could be inferred by these analyses (Figure 1). LecB type 1 (67%) and 2 (1%) were strictly found in the PAO1 phylogenetic cluster (Supplementary Figure S4). The PA7 and alike strains (EML528, EML545, and EML548) were grouped in a same phylogenetic cluster that was found strictly harboring LecB cluster III sequences (types 8, 11, and 12; Figure 1). LecB cluster II sequences, mainly shared by STs of the PA14 phylogenetic group, showed signs of genomic instability, with types 5 and 6 being also recorded in the PAO1 group (Supplementary Figure S4). Their respective pattern of distribution suggested horizontal transfers between the PA14 and PAO1 sub-clades. The phi test was in agreement with this observation, and found statistically significant evidence for recombination events among the *P. aeruginosa* collection between the lecB loci (p = 0.009). A total of 16 LecA and LecB combinations were observed (Supplementary Figure S4). Nine LecA/LecB combinations were observed in the PAO1 cluster with the dominant 1/1 combination at 61%. Four LecA/LecB combinations were observed in the PA14 cluster, and three combinations in the PA7 one.

**Inferred Incidence of Amino Acid Changes on Lectin Structures**

The amino acid variations of LecA and LecB were mapped on the crystal structures of lectins derived from PAO1 (Mitchell et al., 2002, 2005; Cioci et al., 2003). Amino acid changes observed among LecA lineage I (16% of the tested strains) are located on loops that are not directly involved in structure stabilization or ligand binding (Figure 2A). However, LecA type 5 showed an Asn115Ser change (Asparagine to Serine at position 115) that can prevent a hydrophobic contact with Phe82 on the
FIGURE 1 | LecB types among *P. aeruginosa*. (A) Alignment of the 14 types of LecB amino acid sequences detected in this work (bold text) or recovered from the GenBank database (faint text). Amino acids different from those of the PAO1 sequence (total of 25 positions) are boxed and in gray. Black triangles indicate the amino acids participating in sugar binding, and asterisks the amino acids participating in calcium coordination. (B) Neighbor-Joining phylogenetic tree of the 14 LecB sequence types. One hundred and fourteen sites were analyzed. Horizontal lines represent the divergence % between pairs of sequences. Bootstrap values are indicated on branches. Representative strains of the 12 types identified in this work are indicated. Types 9 and 10 are not indicated but showed, respectively, a single change with type 1, and a single one with type 6.
neighboring monomer, therefore destabilizing the LecA tetramer. LecB cluster I observed changes do not have any functional consequences. However, LecB cluster II variants showed Val49Ile and Thr52Ser changes that can affect the stability of the LecB tetramer (Figure 2B), and a Ser23Ala modification demonstrated to double the association constant for fucose (Adam et al., 2007). LecB type 6 and type 7 showed a Gly97Ser close to the fucose-binding site, affecting the binding of fucosylated oligosaccharides. LecB type 8 of cluster III showed a Phe89Val which is involved in dimer formation (Figure 2C), and Gly97Ser and Gly24Ala changes in the binding site area. Other LecB cluster III amino acid changes should not impact the tetramers.

Comparison of Affinities between LecBPA7 and LecBPAO1

Purified LecBPA7 produced in E. coli were assayed on the Glycan Array v5.2 from the Consortium for Functional Glycomics (n = 609 glycans). Only glycans with terminal fucose or mannose residues were recognized by the lectin. Stronger signals were obtained for biantennary glycans that present fucosylated epitopes at the top of two long chains, then to multiple branched glycans with several fucose residues such as Lewis^y epitope (Supplementary Table S6). Data previously obtained for LecBPAO1 (on 465 glycans, available on CFG web site) on epitopes containing only one terminal fucose or mannose residue were extracted and compared with those of LecBPA7 (Figure 3 and Supplementary Figure S5). LecBPA7 recognized more efficiently fucose when present on position 2 of galactose as in blood group O/H, and on position 3 and 4 of N-acetylglucosamine (GlcNAc) as in Lewis^x (Le^x) and Lewis^α (Le^α) epitopes. Comparisons with LecBPAO1 showed H-type 2 (αFuc1-2βGal1-4GlcNAc) to be preferred by LecBPA7 rather than H-type 1 (αFuc1-2βGal1-3GlcNAc), and Lewis^α (αFuc1-4]βGal1-3]GlcNAc) rather than Lewis^y (αFuc1-3]βGal1-4]GlcNAc). However, a substitution by a Gal or a GalNAc on H-disaccharide was not tolerated; so, blood group A and B were better recognized by LecBPA7 rather than LecBPA7. On the opposite, sialic acid substitution on Lewis^α was well tolerated, and siayl Lewis^α (sLe^a) bound as efficiently as Lewis^α. LecBPA7 binds more strongly to mannose than LecBPAO1, and appears less specific for fucose.

FIGURE 2 | Predicted incidence of amino acids changes on LecAPO1 and LecBPO1 crystal structures co-crystallized with galactose (LecA: code 1OKO) and fucose (LecB: code 1GZT). One protein monomer in each structure of LecA and LecB is represented by a blue ribbon. Calcium ions are represented as green spheres, and monosaccharides as sticks. (A) LecAPO1 structure with all amino acids change observed in lineage I (red sticks) and lineage II (blue sticks; Supplementary Figure S1). (B) LecBPO1 structure showing the incidence of the observed amino acids changes among cluster II sequences (Figure 1). (C) LecBPO1 structure showing the amino acids changes of cluster III sequences (Figure 1). All structural figures were drawn with PyMOL Molecular Graphics System.
Affinity data sets were recorded for the interaction of LecB<sub>PA7</sub> with monosaccharides and oligosaccharides using titration microcalorimetry (Supplementary Figure S6 and Supplementary Table S5) and compared with data previously obtained with LecB<sub>PAO1</sub> (Supplementary Table S5) indicates that both lectins have the same affinity for mannoside, and same thermodynamic contribution to the free energy of binding with slightly favorable entropy of binding. However, LecB<sub>PAO1</sub> bound FucOMe and Lewis<sup>a</sup> more efficiently than LecB<sub>PA7</sub>, with factor 5 and 10 respectively. This confirms that LecB<sub>PA7</sub> is less specific for fucosylated oligosaccharides and can likely bind to other glycans.

**Structural Basis for the Differences Between LecB<sub>PA7</sub> and LecB<sub>PAO1</sub>**

One crystal of LecB<sub>PA7</sub> complexed with Lewis<sup>a</sup> tetrasaccharide diffracted to 1.75 Å resolution (Supplementary Table S4) and the structure could be solved in the P2<sub>1</sub> space group, with one tetramer per asymmetric unit. All of 114 amino acids could be located in the electron density, as well as the 8 calcium ions (2 per sites). The whole tetrasaccharide was clearly visible in three sites, and only one reducing glucose was disordered in (2 per sites). The whole tetrasaccharide was clearly visible in three sites, and only one reducing glucose was disordered in (2 per sites).

The structure of the LecB<sub>PA7</sub>/Lewis<sup>a</sup> complex was compared to the one previously obtained for LecB<sub>PAO1</sub>/Lewis<sup>a</sup> complex (Perret et al., 2005). Both structures share the same overall organization with rmsd (root-mean-square deviation) of 0.4 Å over the 456 amino acids of the tetramer. When looking at the binding site, the fucose residue adopts the same orientation in both lectins. No amino acids change involved in calcium coordination or in the hydrogen bonding to fucose was recorded (Figures 2B,C). The other two sugar residues of Lewis<sup>a</sup>, i.e., galactose and GlcNAc, present a slight shift in response to the amino acids changes and the modified network of bridging water molecules (Figure 4). In LecB<sub>PA7</sub>, Ser97 (that replaces Gly97 in LecB<sub>PAO1</sub>) establishes a hydrogen bond with one of the two structural water molecules that bridges between the glycan ligand and the protein surface. The water molecule (Wat1 in Figure 4) is moved away from the protein backbone. As a result, the water molecule network is disturbed in this area, which can correlate with the stronger entropy barrier measured for LecB<sub>PA7</sub> when binding to sugars. The other variation, i.e., Ala24Gly results in an additional hydrophobic contact with the methyl group of Ala24 close to the CH<sub>2</sub> group at C6 of the GlcNAc residue. As displayed in Figure 4, the two mutations can therefore play a role on the fucose affinity by disturbing the water network and on the Lewis<sup>a</sup> affinity by modifying the shape of the binding site on the surface.

**Genome-Wide Analyses of LecB Cluster III Strains**

Assembled draft genomes of three PA7-related strains (EML528, EML545, and EML548) were generated (Boukerb et al., 2015) in order to investigate the LecB cluster III strains showing major changes at their lecA (deleted) and lecB genomic regions. MAUVE comparisons were performed (Figure 5). Two large
regions of genomic instability named α (2,390,740–3,038,628) and β (4,569,960–5,542,308) were detected. One of these regions was found to harbor RGP26 (Mathee et al., 2008) which is containing lecA among PAO1. Among the α region, at least 14 RGP defined in other studies were recorded (e.g., Mathee et al., 2008; Roy et al., 2010; and Klockgether et al., 2011). Five novel RGP were identified and named RGP90, RGP91, RGP92, RGP93, and RGP94 (Figures 5, 6). Some of these RGP matched DNA segments smaller than 5 kb but were confirmed, in some instances, to be targets for longer genetic elements e.g., RGP90. The lecB loci genomic environments were not subjected to similar large scale rearrangements but rather to small ones (up to 16 kb).

The α genomic region harboring lecA showed: (1) an inverted and distantly positioned DNA region between the PAO1-UW (3,028,630–3,077,236) and PA7 (4,096,626–4,144,187) and alike genomes (top arrowhead on Figure 5) including an additional 116 kb DNA segment among the EML545 genome that did not match other sequences in the databases; (2) RGP29 (see asterisk on Figure 5) in EML528 harboring the PAGI-2 genomic island (Supplementary Figure S8A); PAGI-2 was also observed in EML548 but at a distinct genomic position (2,254,558–2,359,617; Supplementary Figure S8B); (3) RGP94 shown to be colonized by a DNA segment of 49.4 kb in EML548 which was located in a different position in PA7 (with 22% of sequence cover and 84% of sequence identity; see bottom left arrowhead on Figure 5), and (4) an unstable region (dark green block, bottom middle arrowhead of Figure 5) of 19.1 kb absent in PAO1 and positioned differently among PA7 and alike genomes. The β variable region was found to match the area involved in the major DNA inversion observed among PAO1 strains, with the localization of RGP5-7, RGP42, and RGP60 (Mathee et al., 2008), and RGP78 (Roy et al., 2010). A large indel was observed in this area (green block, bottom right arrowhead; Figure 5). This block was shared between PA7 (4,584,072–4,620,550; 76% of sequence cover and 96% of sequence identity) and closely related genomes, but at a different position in EML545. It was found to match pUM505, a conjugative plasmid described by Ramirez-Diaz et al. (2011) which is encoding a chromate resistance determinant.

Around 80 kb surrounding lecA (or RGP26 in PA7 and related genomes) was subjected to deeper investigations of DNA
signatures indicative of genomic plasticity such as integrase CDS and breaks in synthy. The borders and size of the DNA segments found to match large and small genomic indels (insertions/deletions) are presented in Supplementary Table S7. These were often found encoding phage-like proteins, integrases, transcriptional regulators, and various enzymes, factors and transporters. Interestingly, the lecA area was found to contain several tRNA gene sequences (tRNA^{Leu}, tRNA^{Cys}, tRNA^{Gly}, and tRNA^{Ser}), and these are often targeted by integrative elements. The lecA CDS was in frame and proximal to a tRNA^{Leu} gene. This zone was previously defined as RGP26 in PA7 (Supplementary Table S7). This RGP was shown to contain CDS indicative of a likely integration of phage DNA which could have led to the observed lecA deletion in the PA7 clade. RGP90 and RGP94 were also found colonized by phage DNA, suggesting the lecA area to be frequently challenged by phage attacks. In fact, a 35 kb-long DNA segment was found at RGP90 and matched a bacteriophage B3 element (Figure 6 and Supplementary Figure S9). RGP94 is bordered by a tRNA^{Ser} gene which appeared to have been targeted by a P52 bacteriophage in the EML548 genome (Supplementary Figure S10). It is noteworthy that the tRNA^{Cys} gene was found in a region of great instability harboring RGP92 which harbored a CDS encoding a fimbrial protein and a transcriptional protein of the MerR family in EML545 and EML548 genomes (2.9 kb).

Comparative analysis of lecB genomic regions in *P. aeruginosa* PAO1, PA7, and the EML strains revealed a higher conservation than the lecA ones. Only four RGP could be detected (Supplementary Figure S11 and Supplementary Table S8). The lecB gene was physically linked to RGP96. This RGP was delimited by the stop codon of lecB in PAO1, and PA7 and closely related genomes. Small insertions of 0.6 kb in PAO1 and 0.9 kb in PA7 and alike genomes were detected at this RGP but these inserts were unrelated. These changes in the terminator region could be affecting the lecB expression levels but the promoter region did not appear to vary from one strain to another. A tRNA^{Arg} gene favored the integration of foreign DNA at RGP58. No phase DNA, integrase CDS or adhesin-like CDS could be detected in the analyzed lecB area.
FIGURE 6 | ACT analyses of lecA genomic regions of P. aeruginosa strains PAO1, PA7, EML528, EML545 and EML548. (A) PAO1 (top) against PA7 (bottom); (B–D) PA7 (top) against, respectively, EML528 (B), EML545 (C), and EML548 (D) regions (bottom). Horizontal black lines indicate the genome sequences compared over a length of about 80 kb. Gray indicates conserved regions, and white the variable ones. Distances between variable regions are given in kb. Regions of genomic plasticity (RGP) are indicated in black-boxes. Selected CDS (above or below the ACT panels), and tRNA genes (in black) are shown. Orthologous tRNA genes between pairs of genomes are highlighted by a light-green line crossing the ACT panels.
**P. aeruginosa** Cells Aggregation Behavior According to LecA/LecB Types Combinations

Global incidence of lectin types on bacterial aggregation was investigated by analyzing the number of aggregates (>20 µm²) in a PBS solution of *P. aeruginosa* after exposure to various ligands. Calix[4]arenes glucosylated (Calix-Glc₄), mannosylated (Calix-Man₄), fucosylated (Calix-Fuc₄), and galactosylated (Calix-Gal₄) glyoclusters (Supplementary Figure S1) at 100 µM were added to a panel of PBS-washed *P. aeruginosa* cells representative of the observed LecA/LecB types combinations (Figure 7). The aggregation values were compared with those obtained from cells exposed to PBS or methyl glycosides (GlcOMe, ManOMe, FucOMe, and GalOMe; Supplementary Figure S1).

Considering LecA composition to be quite conserved among *P. aeruginosa*, it is its absence which should lead to differences in the aggregation assays involving galactosylated ligands. To verify this point, a PAO1ΔlecA mutant (Boukerb et al., 2014) was included in the assays. This mutant did not produce significant aggregates while being exposed to galactosylated glyoclusters (p > 0.05; Figure 7A). However, this calix[4]arene was also poorly efficient at inducing aggregation with wildtype PAO1 cells. Similarly, a PAO1ΔlecB mutant was analyzed to investigate changes after exposure to fucosylated and mannosylated glyoclusters. This PAO1ΔlecB mutant showed lower aggregation than those obtained with PAO1 exposed to Calix-Fuc₄ (p < 0.05; Figure 7A). These results are in line with an involvement of LecB in the aggregation phenotype. Surprisingly, higher aggregation was observed for this mutant compared to the wild-strain exposed to Calix-Man₄ (p < 0.05; Figure 7A).

Strains harboring various LecB types were tested in these aggregation assays with calix-Man₄, calix-Fuc₄, calix-Gal₄, and calix-Glc₄. LecB cluster I strains (PAO1 and bpoe1643) showed significantly higher numbers of aggregates when exposed to calix-Fuc₄ than to monovalent FucOMe (Figure 7 and Supplementary Figure S12). Cluster II strains were much more reactive toward glyoclusters and yielded significantly higher numbers of aggregates with Calix-Man₄, Calix-Fuc₄, and Calix-Glc₄ (Figure 7 and Supplementary Figure S12). However, numbers of aggregates were always much lower for strain PA14 (ST253). This behavior appeared specific of this strain because bpoe1392 which is also part of ST253 showed very strong aggregation responses toward the tested glyoclusters (Figure 7). LecB cluster III strains yielded significantly higher numbers of aggregates when exposed to Calix-Fuc₄ and Calix-Glc₄ than monovalent sugars (Figure 7 and Supplementary Figure S12). However, significant differences could not be obtained using Calix-Man₄ even though, in average, higher numbers were measured. Interestingly, for cluster II and III strains, glucosylated glyoclusters were found inducing aggregation, and sometimes at levels higher than those observed with fucosylated, mannosylated, or galactosylated glyoclusters (Supplementary Figure S12). All these datasets are in line with a role of lectins in *P. aeruginosa* aggregation properties. However, other adhesins or cellular products appeared to be involved.

**DISCUSSION**

Recent studies demonstrated the great potential of sugar solutions and glyoclusters in the prevention of *P. aeruginosa* infections (e.g., Chemani et al., 2009; Boukerb et al., 2014). However, these investigations have, so far, been done without considering the great diversity observed among this bacterial species, and the impact of naturally selected amino acid changes on the specificity and affinity of these lectins for their ligands. Here, the analysis of a significant collection of clinical and outdoor *P. aeruginosa* isolates revealed genetic and genomic changes at the *lecA* and *lecB* loci that could affect the efficacy of such glyco-therapies.

In this work, loss and IS disruption of *lecA* among *P. aeruginosa* genotypes were observed but a higher genetic drift and likelihood of recombination were computed for *lecB* than *lecA* CDS. PyMoI analyses of LecA types suggested low or no incidence of the detected amino acid changes on its functioning. More significant changes were detected among LecB types, and suggested modifications in the interactions with the ligands, and changes in the binding spectra. *P. aeruginosa* LecB types were divided into three clusters (I, II, and III). Cluster I variants (PAO1-related strains) were not likely to have sugar affinities different from those reported so far. Cluster II (PA14-related strains) variants showed a Ser23Ala modification which was previously shown to double affinities toward fucose (Adam et al., 2007). LecB cluster III variants (PA7-related strains) showed modifications likely affecting the LecB sugar binding site and its oligomerization. These latter changes led us to further characterize the structure, specificities, and affinities of a LecB from this cluster i.e., the one of strain PA7.

LecBPₐ₇ bound to the same oligosaccharide epitopes as LecBPₐ₇ on the tested glycan array but showed higher affinity for some organizations such as the ones of blood group O/H over those of A and B, and those of Lewis*⁺ (Le⁺) and Lewis*⁻ (Le⁻) epitopes. Titration microcalorimetry analyses of LecBPₐ₇ showed methyl-fucoside and Lewis*⁺ to have a stronger affinity (Kₐ ∼ 4 μM) than methyl-mannoside (Kₐ = 73 μM). Nevertheless, since LecBPₐ₇ binds as efficiently to mannose as LecBPₐ₇, but less efficiently to fucose, this confirms its lower specificity for fucosylated saccharides. When analyzing the thermodynamics of binding, the lower affinity of FucOMe was not related to a less efficient binding in the LecBPₐ₇ pocket. Indeed, the enthalpy contribution was higher than for LecBPₐ₇ but the entropy barrier was much stronger, pointing out for a different contribution of water molecules. Structural analysis of LecBPₐ₇ complexed with Lewis*⁺ tetrasaccharide validated this inference. This structure was compared with the one of the LecBPₐ₇/Lewis*⁺ complex (Perret et al., 2005) that shares the same overall organization. Fucose residues have the same orientation in both lectins but galactose (Gal) and GlcNAc of Lewis*⁺ are slightly reoriented. In LecBPₐ₇, Ser97, instead of Gly97 in LecBPₐ₇, can establish a hydrogen bond with one of the two structural water molecules of the binding site. The resulting disturbance of the water molecule network in this area correlates with the stronger entropy barrier.
measured for LecB<sub>PA7</sub> when binding to sugars than LecB<sub>PA01</sub>. Furthermore, the Ala24Gly modification in LecB<sub>PA7</sub> generated an additional hydrophobic contact close to the CH<sub>2</sub> group at C6 of the GlcNAc residue. These two major modifications likely contributed to the changes observed over fucose and Lewis<sup>a</sup> affinities.
Further analyses of the diversity of Lec types in *P. aeruginosa* and comparisons with the phylogenetic allocations of concatenated DNA sequences from the MLST scheme of the *P. aeruginosa* strains used in this study allowed inferring dominant LecA and LecB types in the collection. Nine combinations of LecA/LecB types were recorded in the PAO1 cluster, four in PA14 one, and three in PA7 one. The PA7 and alike llecB types were found to have a restricted distribution. Some peculiar combinations suggesting either genetic convergence or genetic recombination in the *P. aeruginosa* global population were identified. In fact, LecA types 2 and 4 were found restricted to the PAO1 phylogenetic group but types 1 and 3 were recruited by both the PAO1 and PA14 groups. Nevertheless, the llecA loci did not appear to have undergone significant changes by recombination in the *P. aeruginosa* population. However, the llecB loci were found to have been impacted by recombination events. In fact, LecA types 1 and 3 were restricted to cluster I. LecB types 5 and 6 were found detected among highly distant lineages. The above observations clearly demonstrated that genetic and genomic diversifications among *P. aeruginosa* will affect the lectin-ligand affinity profiles of a strain. One could thus question the use of glycomimetics in the prevention of *P. aeruginosa* infections. Still, several other adhesins play part in the oligosaccharide-mediated recognition and adhesion to host cells (e.g., Duque et al., 2013). In fact, *in silico* searches of PAO1, PA7 and the closely related genomes reported here can lead to a long list of CDSs likely involved in the expression/synthesis of adhesins e.g., znuA, cupA4, cupB6, cupE6, fimA, pilE, fliC, and fliD (data not shown). In order to clarify this situation, synthetized tetravalent glyco-clusters built over a calixarene backbone were used to evaluate their ability at favoring aggregation of *P. aeruginosa* cells. This assay detected a significant effect of Calix-Gal4 on strain PA7 (which is not having llecA) confirming the likely involvement of other adhesins/processes on the generation of these aggregates. It is noteworthy that, in these assays, Calix-Fuc4 was found the most efficient glycocluster for inducing aggregation on a large panel of *P. aeruginosa* strains. Furthermore, LecB cluster II (PA14 and related strains) and cluster III (PA7 and related strains) were much more reactive toward the tested glyco-clusters and yielded significantly higher numbers of aggregates with Calix-Fuc4 and Calix-Glc4 than cluster I strains (PAO1 and related strains). These data sets are in line with a role of lectins in *P. aeruginosa* aggregation properties. High concentration of Calix-Fuc4 could thus be efficient at preventing lung colonization not only of strain PAO1, as reported by Boukerb et al. (2014), but of most *P. aeruginosa* strains.

**CONCLUSION**

This work showed genomic and genetic changes at the llec loci suggesting ongoing adaptive processes among *P. aeruginosa* which can lead to their loss or changes in their affinities toward...
sugar ligands. The PA7 clade was found to have counter-selected the lecA locus while maintaining the lecB one. Significant differences with LecB<sub>PAO1</sub> were recorded and shown to have affected LecB<sub>PA7</sub> affinities for sugar ligands. These changes appeared related to the early divergence observed among the <i>P. aeruginosa</i> radiation between the PAO1/PA14 and PA7 lineages. Genetic exchanges among the PA7 clade at the lecA and lecB loci were not detected but analyses of their neighboring sequences showed several RGP with integration of exogenous genetic elements including the well-described PAGI-2 at RGP29 of strain EML528. This is indicative of genetic exchanges among the PA7 clade. All these changes did not appear to have significantly impacted the interactions with Calix-Fuc<sub>4</sub>. This glyco-cluster was found to significantly increase aggregation of <i>P. aeruginosa</i> cells from the three main clades (PAO1, PA14, and PA7 clades), confirming its great potential as a glycomimetic. High concentration of this glycomimetic could thus be efficient at preventing lung colonization by most <i>P. aeruginosa</i> strains.

**AUTHOR CONTRIBUTIONS**

Conceived and designed experiments: AB, SV, AI, and BC. Performed the experiments: AB, AD, SR, RT, AR, LR, SV, AV, and BC. Analyzed the data: AB, AI, SR, LC, SB, ADJ, AV, BC. Wrote the paper: AB, AI, BC. All authors read and approved the final manuscript.

**REFERENCES**

Adam, J., Pokorná, M., Sabin, C., Mitchell, E. P., Imberty, A., and Wimmerová, M. (2007). Engineering of PA-IIL lectin from <i>Pseudomonas aeruginosa</i> - unravelling the role of the specificity loop for sugar preference. <i>BMC Struct. Biol.</i> 7:36. doi: 10.1186/1472-6807-7-36

Battye, T. G., Kontogianni, L., Johnson, O., Powell, H. R., and Leslie, A. G. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFELM. <i>Acta Crystallogr. D Biol. Crystallogr.</i> 67, 271–281. doi: 10.1107/S0907444910048675

Bernardi, A., Jiménez-Barbero, J., Casnati, A., De Castro, C., Darbre, T., Fieschi, F., et al. (2013). Multivalent glycoconjugates as anti-pathogenic agents. <i>Chem. Soc. Rev.</i> 42, 4709–4727. doi: 10.1039/C2CS35408J

Boukerb, A. M., Marti, R., and Cournoyer, B. (2015). Genome sequences of three strains of the <i>Pseudomonas aeruginosa</i> PA7 clade. <i>Genome Announc.</i> 3: e01366-15. doi: 10.1128/genomeA.01366-15

Boukerb, A. M., Rouset, A., Galanos, N., Méar, J. B., Thépaut, M., Grandjean, T., et al. (2014). Antiadhesive properties of glycoclusters against <i>Pseudomonas aeruginosa</i> lung infection. <i>J. Med. Chem.</i> 57, 10275–10289. doi: 10.1021/jm500338p

Braid, M. D., Silhavy, J. L., Kitts, C. L., Cano, R. J., and Howe, M. M. (2004). Complete genomic sequence of bacteriophage B3, a Mu-like phage of <i>Pseudomonas aeruginosa</i>. <i>J. Bacteriol.</i> 186, 6560–6574. doi: 10.1128/JB.186.19.6560-6574.2004

Cecioni, S., Imberty, A., and Vidal, S. (2015). Glycomimetics versus multivalent glycoconjugates for the design of high affinity lectin ligands. <i>Chem. Rev.</i> 115, 523–561. doi: 10.1021/c500303t

Chemani, C., Imberty, A., de Bentzmann, S., Pierre, M., Wimmerová, M., Guery, B. P., et al. (2009). Role of LecA and LecB lectins in <i>Pseudomonas aeruginosa</i>-induced lung injury and effect of carbohydrate ligands. <i>Infect. Immun.</i> 77, 2065–2075. doi: 10.1128/IAI.01204-08

**FUNDING**

This work was partly funded by Anses project “pyo-eau” #2011/1/137 of the "Programme Environnement-Santé-Travail" (French Ministers in charge of ecological and environmental issues). We thank the CNRS, Université Lyon 1, VetAgro Sup, Cluster “Infectiology” of the Rhône-Alpes region, and Labex IMU (Intelligence des Mondes Urbains) (France) for having funded parts of this work. Parts of this work were also funded by the French Fond Unique Interministériel (managed by Oseo and DGCIS), FUI anti-pyo, Labex ARCANÉ (ANR-11-LABX-0003-01), COST Action CM-1102 MultiGlycoNano, and the Conseil Régional Rhône-Alpes. AB, AI, and BC acknowledge support from GDR Pseudomonas.

**ACKNOWLEDGMENTS**

We thank Sophie de Bentzmann and David Redelberger of UMR CNRS 7525, Aix Marseille Université, for having made available the PAO1<sub>lecA</sub> and PAO1<sub>lecB</sub> mutants. The glycan array resources were provided by the Consortium for Functional Glycomics (GM62116).

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.00811
Boukerb et al. Diversification of Pseudomonas aeruginosa lecA and lecB Regions

genome of the opportunistic pathogen updated. PLoS ONE 10:e0126468. doi: 10.1371/journal.pone.0126468

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D Biol. Crystallogr. 67, 235–242. doi: 10.1107/S0907444910045749

Winsor, G. L., Griffiths, E. J., Lo, R., Dhillon, B. K., Shay, J. A., and Brinkman, F. S. (2015). Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic Acids Res. 44, D646–D653. doi: 10.1093/nar/gkv1227

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Boukerb, Decor, Ribun, Tabaroni, Rousset, Commim, Buff, Doleans-Jordheim, Vidal, Varrot, Imberty and Cournoyer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.