Journal of Proteomics 146 (2016) 25–33

Contents lists available at ScienceDirect

Journal of Proteomics

journal homepage: www.elsevier.com/locate/jprot

Deciphering the interactions between the *Bacillus cereus* linear plasmid, pBClin15, and its host by high-throughput comparative proteomics

Jean-Paul Madeira [a,b], Hélène Omer [a,b], Béatrice Alpha-Bazin [b], Jean Armengaud [b], Catherine Duport [a,*]

[a] SQPOV, UMR0408, Avignon Université, INRA, F-84914 Avignon, France
[b] CEA, DSV, IBiTec-S, SPI, Li2D, Laboratory "Innovative technologies for Detection and Diagnostics", Bagnols-sur-Cèze F-30200, France

ARTICLE INFO

Article history:
Received 19 October 2015
Received in revised form 22 April 2016
Accepted 15 June 2016
Available online 16 June 2016

Keywords:
Proteome
Bacillus cereus
Plasmid
Prophage
Shotgun proteomics

ABSTRACT

The pathogen, *Bacillus cereus*, is able to adapt its metabolism to various environmental conditions. The reference strain, *Bacillus cereus* ATCC 14579, harbors a linear plasmid, pBClin15, which displays a cryptic prophage behavior. Here, we studied the impact of pBClin15 on the aerobic respiratory metabolism of *B. cereus* by curing its host strain. We compared, by means of a high-throughput shotgun proteomic approach, both the cellular proteome and the exoproteome of *B. cereus* ATCC 14579 in the presence and absence of pBClin15 at the early, late and stationary growth phases. The results were visualized through a hierarchical cluster analysis of proteomic data. We found that pBClin15 contributes significantly to the metabolic efficiency of *B. cereus* by restricting the production of chromosome-encoded phage proteins in the extracellular milieu. The data also revealed intricate regulatory mechanisms between pBClin15 and its host. Finally, we show that pBClin15 provides benefit to its host to adapt to different ecologic niches.

Biological significance: Bacteria belonging to the *Bacillus cereus* group include *B. cereus*, a notorious food borne pathogen which causes gastroenteritis. The *B. cereus* type, strain ATCC 14579, harbors a linear plasmid, pBClin15, which displays cryptic prophage behavior. Here, we present data supporting the idea that pBClin15 may have a much greater role in *B. cereus* metabolism that has hitherto been suspected. Specifically, our comparative proteomic analyses reveal that pBClin15 manages *B. cereus* central metabolism to optimize energy and carbon utilization through the repression of several chromosome-encoded phage proteins. These results suggest that pBClin15 provides benefit to the host for surviving adverse environmental conditions.

1. Introduction

Members of the *Bacillus cereus sensu lato* group are found in diverse environments and include eight closely related species – *Bacillus cereus sensu stricto*, *Bacillus anthracis*, *Bacillus thuringiensis*, *Bacillus weihenstephanensis*, *Bacillus mycoides*, *Bacillus pseudomycoides*, *Bacillus cytotoxicus*, and *Bacillus toyonensis* [1,2]. The plasmids in this group display a strain-dependent distribution, with some strains containing no plasmid, whereas others have many. Some of these plasmids have a small genome size, only 2 kb, whereas others are very large, up to 600 kb [3]. Large plasmids are key components in defining the phenotypic traits associated with pathogenesis [4]. For example, emetic syndrome, which is associated with *B. cereus sensu stricto*, is caused by cereulide. The cereulide synthetase gene cluster, which encodes the enzymatic machinery required for the biosynthesis of cereulide, is located on a 208 kb megaplasmid [5]. As a mammalian pathogen, the ability of *B. anthracis* to cause anthrax originates from two large plasmids: pXO1 (181.6 kb) that encodes the tripartite lethal toxin complex and pXO2 (93.5 kb), which contains the biosynthetic genes for the poly-γ-D-glutamic acid capsule [6]. *B. thuringiensis* forms crystal-like parasporal inclusions during sporulation [7–9]. These inclusions contain δ-endotoxins with insecticidal properties. Most of the genes (*cry*) encoding these toxins are located on large plasmids [10]. Whereas the role of large plasmids in pathogenesis is well defined, the function of the other plasmids in the group is relatively unknown.

The *B. cereus* type, strain ATCC 14579, harbors a 15,274 bp linear plasmid, pBClin15, which displays cryptic prophage behavior [11–13]. The genetic structure of pBClin15 is reminiscent of temperate phages belonging to the *Tectiviridae* family. This family includes the bacteriophages, Bam35, GIL01 and GIL16 from *B. thuriengensis* and AP50 from *B. anthracis* [13–18]. The pBClin15 genome contains 28 open reading frames (ORFs) [17]. Some of these ORFs encode proteins involved in phage genome replication and regulation, virion structure and DNA packaging [17,19], but the function of many is so far undocumented. One of these ORFs, namely ORF2, encodes a protein (Bcp0002) that displays binding properties towards DNA fragments of both chromosomal and plasmid origin. The effect of Bcp0002 binding is currently unknown [20].

* Corresponding author at: UMR SQPOV, INRA PACA, 228, Route de l'Aérodrome, CS 40509, Domaine Saint Paul-Site Agroparc, 84914 Avignon Cedex 9, France.
E-mail address: catherine.duport@univ-avignon.fr (C. Duport).

http://dx.doi.org/10.1016/j.jprot.2016.06.022
1874-3919/

With the emergence of "global systems biology" tools (mainly DNA microarrays), cryptic prophages have gradually gained attention, as part of host physiology, especially under stress conditions [21]. Our goal here was to understand the importance of the effect of pBClin15 on *B. cereus* physiology under normal growth conditions. For this, bacteria were grown in pH-regulated batch cultures under fully oxic conditions. We designed a high-throughput shotgun proteomic approach to decipher the interaction of pBClin15 with its host. Hierarchical clustering analysis of the proteomic data identified the altered biological processes, according to the metabolic data, and the empirical Bayes approach identified the main regulators responsible for the changes in the phenotypes. Strikingly, we determined that pBClin15 has a beneficial impact on *B. cereus* physiology. This probably helps these bacteria to cope with adverse environments.

2. Materials and methods

2.1. Bacterial strains and growth conditions

The B. cereus strain used in this study was the reference strain, ATCC 14579 [12]. This strain was cured of pBClin15 as described by Voros et al. [21], resulting in the ΔpBClin15 strain. *B. cereus* strains (with and without pBClin15) were cultured in batches (three independent cultivations per strain) at $pO_2 = 100\%$ [22] in minimal MOD medium supplemented with 30 mM glucose as the carbon source [23]. Each batch culture was inoculated with an overnight subculture at an initial optical density at 600 nm (OD_{600nm}) of 0.02. Aerobic batch cultures were performed at 37 °C in a 3-liter bioreactor (BioFlo®/CelliGen®115, New Brunswick), with a working volume of 2 l. The pH was maintained at a controlled value of 7.2 by automatic addition of 5 M KOH. *B. cereus* growth was monitored spectrophotometrically at 600 nm.

2.2. Cellular and exoproteome extraction and metabolite assays

The cells were harvested by centrifugation at the early exponential (EE), late exponential (LE) and stationary (S) growth phases for each of the bioreactor cultures and stored frozen at −80 °C until analysis. Soluble cellular proteins were extracted as previously described [24]. The extracellular proteins were obtained by trichloroacetic acid precipitation as described in [22]. The protein concentration was determined by the Bradford protein assay after TCA precipitation. Enzymatic test kits from Roche (Meylan, France) were used to measure glucose, lactate, ethanol, formate, acetate, and succinate concentrations in the supernatants.

2.3. Proteolysis and shotgun tandem mass spectrometry

Protein samples were loaded onto NuPAGE 4–12% Bis-Tris gels (Invitrogen) for a short (about 3 mm) electrophoretic migration in denaturing conditions. For each protein sample, the whole content was extracted as a single polyacrylamide band. The bands were subjected to proteolysis with sequencing grade trypsin (Roche) using 0.01% ProteaseMAX surfactant (Promega) as previously described [25,26]. The resulting peptides were analyzed by nanoLC-MS/MS using an LTQ-Orbitrap XL hybrid mass spectrometer (ThermoFisher) coupled to an Ultimate 3000 nRSLC system (Dionex, ThermoFisher). The experimental set-up and conditions were as described [22,27]. Briefly, peptide digests were desalted on-line on a reverse-phase precolumn (Acclaim PepMap 100 C18, 5 μm bead size, 100-Å pore size, 300 μm i.d. × 5 mm, Dionex ThermoFisher). The peptides from the extracellular digests were then resolved on a nanoscale C18 PepMap100™ capillary column (3 μm bead size, 100-Å pore size, 75 μm i.d., 15 cm length, LC Packings, ThermoFisher) using a 90-min gradient. The peptides from cellular digests were resolved with the same column but 50 cm long using a 180-min gradient. In both cases, the gradient was from 4 to 40% solvent B (0.01% HCOOH, 100% CH_3CN) with solvent A being 0.01% HCOOH, 100% H_2O. Full-scan mass spectra were measured from *m/z* 300 to 1800 in data-dependent mode using the TOP3 strategy.

2.4. MS/MS spectra assignments to peptide sequences and protein validation

An in-house polypeptide sequence database was made of the sequences of all previously annotated proteins encoded by the *B. cereus* ATCC 14579 chromosome (NC_004722) and the plasmid pBClin15 (NC_004721), and 44 proteins identified by a previous proteogenomic study (see Supplementary Table 1 in Ref. [28]). This database, used to assign peptide sequences to MS/MS spectra, comprises 5299 polypeptide sequences totaling 1,464,675 amino acids. The MASCOT Daemon search engine (version 2.3.02; Matrix Science) was used for searching tryptic peptides with the following parameters: full-trypsin specificity, a mass tolerance of 5 ppm on the parent ion and 0.5 Da on the MS/MS, carboxyamidomethylated Cys (+57.0215) as a fixed modification and oxidized methionine (+15.9949) as a variable modification. The number of tolerated missed cleavages was set at 2. All peptide matches with a score below a *p*-value of 0.05 were filtered by the IRMa 1.28.0 parser [29]. A protein was considered validated when at least two different peptides were detected when considering all the samples. The false-positive rate for protein identification was estimated using the appropriate decoy database as below 0.1% with these parameters. The mass spectrometry proteomics data have been deposited in the ProteomeXchange Consortium (http://proteomecentral.proteomeexchange.org) via the

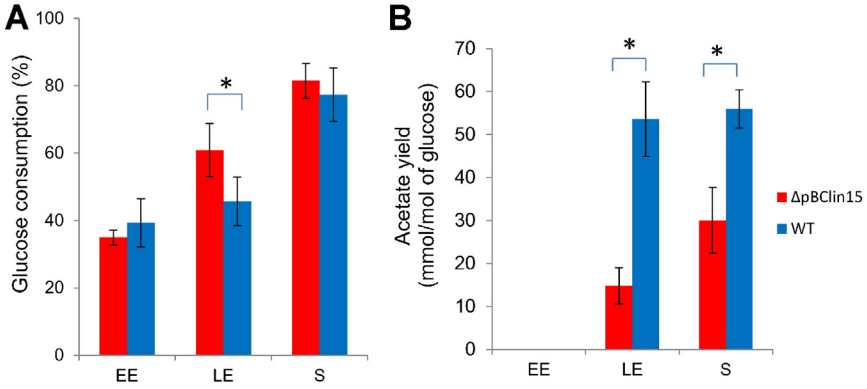

Fig. 1. Growth characteristics of ΔpBClin15 and its parental *B. cereus* ATCC 14579 strain. Panel A. Kinetics of glucose consumption in ΔpBClin15 and WT cells. Consumption of glucose was measured from the samples isolated at the EE, LE and S growth phases. Panel B. Kinetics of acetate production. Acetate yield was measured from the samples isolated at the EE, LE and S growth phases. Error bars represent the standard deviation from 3 independent measures. Significant differences (*p*-value < 0.05) are indicated with asterisks.