Available online at www.sciencedirect.com
SCIENCE DIRECT®
Research in Microbiology 156 (2005) 289–297
Research in Microbiology
www.elsevier.com/locate/resmic

Mini-review

Sorting sortases: a nomenclature proposal for the various sortases of Gram-positive bacteria

Shaynoor Dramsi [a,*], Patrick Trieu-Cuot [a], Hélène Bierne [b]

[a] *Unité de Biologie des Bactéries Pathogènes à Gram-positif, URA CNRS 2172, 25, rue du Dr. Roux, 75724 Paris, France*
[b] *Unité des Interactions Bactéries-Cellules, Institut Pasteur, 28, rue du Dr. Roux, 75724 Paris, France*

Received 13 August 2004; accepted 8 October 2004

Available online 8 December 2004

Abstract

Bacterial surface proteins constitute a diverse group of molecules with important functions, such as adherence, invasion, signaling and interaction with the host immune system or the environment. In Gram-positive bacteria, many surface proteins are anchored to the cell wall envelope by an enzyme named sortase, which recognizes a conserved carboxylic sorting motif. The sequence of the prototype staphylococcal SrtA has been widely used to identify homologs in bacterial genomes, revealing a profusion of sortases in almost all Gram-positive bacteria, often with more than one sortase-like protein per genome [M.J. Pallen, A.C. Lam, M. Antonio, K. Dunbar, Trends Microbiol. 9 (2001) 97–102]. In light of increasing reports on the identification and/or characterization of paralogous sortase genes, a classification of sortases now appears necessary. This report provides an analysis of sixty-one sortases from complete Gram-positive genomes, and suggests the existence of four structural groups of sortases. We propose the classification of sortases into 4 classes designated A, B, C and D. This classification should help to discriminate between sortases in the future.

Keywords: Sortase; Cell-wall anchoring; LPXTG motif; Surface protein

1. Introduction

Sortase A (SrtA) is a membrane-bound thiol transpeptidase that cleaves LPXTG proteins between threonine and glycine of the LPXTG motif (for reviews see [24] and [28]). It catalyzes the formation of an amide bond between the carboxyl group of threonines and the amino group of cell-wall peptidic cross-bridges, thereby anchoring the protein to the peptidoglycan [32,34]. SrtA of *Staphylococcus aureus* is a 206-amino acid polypeptide harboring an N-terminal segment of hydrophobic amino acids that is believed to function as both a signal peptide for secretion and a stop transfer signal for membrane anchoring [22]. The three-dimensional structure of SrtA revealed a unique ß-barrel structure in which the catalytic site residue cysteine 184 is in close proximity with histidine 120, suggesting a model whereby sortase forms a thiolate–imidazolium ion pair for catalysis [15,33]. Recently, a conserved arginine residue has been shown to be required for efficient catalysis of sortase A [21].

The sequence of the prototype staphylococcal SrtA has been widely used to identify homologs in the genomes of many Gram-positive bacteria. Using PSI-BLAST analysis, Pallen et al. have identified sortases in almost all Gram-positive bacteria, with often more than one sortase-like protein per genome [27]. Three other reports mentioned the existence of different types of sortases based on sequence analysis [5,8,15].

With the plethora of studies reporting the identification and/or characterization of one or several paralogous copies of sortase genes [2,6,17,19,26,35], a classification of sortases now appears necessary. Based on detailed analysis of sixty-one sortases from complete Gram-positive genomes, we propose classifying sortases into four classes designated A, B, C and D.

* Corresponding author.
E-mail address: sdramsi@pasteur.fr (S. Dramsi).

0923-2508/$ – see front matter
doi:10.1016/j.resmic.2004.10.011

2. Phylogenetic analysis of sortases defines four clusters

The availability of complete genome sequences enabled us to identify sixty-one sortases originating from 22 Gram-positive bacterial species. In some cases, we noticed sequencing and/or annotation errors that were corrected before further analysis (for details, see legend of Fig. 1).

The sixty-one sortases were aligned using the Clustal X program [31] and subjected to phylogenetic analyses using the neighbor-joining method [30]. This analysis revealed four major clusters which defined the sortases: classes A, B, C and D (Fig. 1). Multiple sequence alignments revealed the presence of three conserved domains (D1, D2 and D3) (Fig. 2). As expected, the three amino acid residues thought to be involved in the formation of the catalytic site were conserved in domains D2 (H_{120}) and D3 (C_{184}, R_{197}) of all sixty-one sequences (numbering is according to the *S. aureus* SrtA sequence). Each class of sortases also possessed a specific pattern of conserved amino acids (supplementary material S1–S4). All sixty-one sortases possessed a signal peptide that was sometimes missing in the genome annotations. Remarkably, only sortases of class C displayed a C-terminal hydrophobic domain (Figs. 2 and 3).

3. Description of the four classes of sortases

3.1. The sortase A class

From last year to date, the role of sortase A in anchoring LPXTG-bearing proteins to the cell wall has been confirmed in many low GC% Gram-positive bacteria such as *Listeria monocytogenes* [5,12], *Streptococcus pyogenes* [2], *Streptococcus pneumoniae* [17], *Streptococcus suis* [26], *Streptococcus gordonii* [6], *Streptococcus mutans* [19], and *Streptococcus agalactiae* (Lalioui et al., submitted).

The sortase A class, which includes the prototype SrtA from *S. aureus*, is constituted of unique SrtA representatives from nearly all low GC% Gram-positive bacteria with the exception of *Enterococcus faecalis* strain V583 where two SrtAs have been found (EF3056 and EF2524) (see below). All SrtAs harbor the N-terminal signal peptide/transmembrane domain and, at the C-terminus, the catalytic TLXTC signature sequence, in which X is often a valine, an isoleucine or a threonine (Fig. 2); SrtA is a broad-range enzyme required for anchoring the majority or all of the LPXTG-containing proteins of a given bacteria [4,5,25,26]. The gene encoding SrtA is generally not clustered with any of its substrates.

In *E. faecalis* strain V583, two *srtA* genes have been found (coding for EF3056 and EF2524). However, a PCR analysis for the presence of these two genes in various *E. faecalis* strains (JH2.2, V583, UM4, NEM776, NEM786 and three clinical vancomycin-resistant strains) showed that all strains possessed a gene coding for EF3056 whereas only two vancomycin-resistant strains, V583 and UM4, harbored an EF2524 encoding gene (unpublished results). Indeed, Paulsen et al. mentioned that the region (EF2512-2542) is an integrated plasmid remnant that encodes a cell wall surface protein (EF2525) and a sortase (EF2524), which explains why EF2524 is not present in all *E. faecalis* strains and represents an example of lateral gene transfer [29].

3.2. The sortase B class

The sortase B class, which includes the prototype SrtB from *S. aureus*, is the smallest group with 7 members from a few low GC% Gram-positive bacilli and cocci (Fig. 1). The SrtB of *S. aureus* is required for anchoring IsdC, a surface protein with a NPQTN motif in place of the LPXTG motif [25]. The genes encoding SrtB and its target IsdC reside in the same operon involved in iron acquisition [23]. Sequence analysis revealed the presence of sortases B in a small number of Gram-positive bacteria, such as *L. monocytogenes*, *Bacillus anthracis*, *Bacillus cereus*, *S. pyogenes* and *Clostridium perfringens* (Fig. 1, cluster B; [5]). All sortases B class contain an N-terminal signal peptide/transmembrane domain, three amino acid segments (B1, B2 and B3) not present in SrtA enzymes and the catalytic TLXTC motif, in which X is usually a serine (Fig. 2). Secondly, as in *S. aureus*, the genes encoding SrtB and its putative targets are often part of the same locus. This has recently been shown in *L. monocytogenes*, in which two genes encoding the SrtB targets reside in the *srtB* operon [4]. The organization of the *srtB–isd* locus of *S. aureus* appears partially conserved in *Listeria innocua*, *B. anthracis* and *B. cereus*, as well as in the Gram-positive extremophile *Bacillus halodurans* (Fig. 4). Lastly, SrtB does not anchor LPXTG proteins and recognizes a novel type of sorting signal, which is NPQTN in *S. aureus* [25]. The other putative SrtB recognition motifs are: NAKTN (or NPKSS) in *L. monocytogenes*, NPQTG and NSKTA in *B. halodurans* and NPKTG and NSKTA in *B. anthracis* and *B. cereus* [4] (Fig. 4). No obvious substrates for the SrtB of *S. pyogenes* and *C. perfringens* could be found in the genome (see below). Although the genomes of 8 enterococcal–streptococcal–lactococcal species have been sequenced, only *S. pyogenes* (all 4 sequenced strains M1, MGAS315, MGAS8232 and SSI-1) appears to encode a SrtB. Moreover, a SrtB encoding gene is present in *S. aureus* genomes but absent from those of *S. epidermidis* ATCC 12228, which suggests a recent acquisition of this genetic system and a role in pathogenesis [16].

3.3. The sortase C class

The sortase C class is the largest group, with 24 sortases originating from high GC% (corynebacteria) and low GC% Gram-positive bacteria (bacilli, clostridia, enterococci, and streptococci) often present in several copies per genome. As shown in Figs. 2 and 3, this group of sortases is characterized by the presence of a C-terminal hydrophobic domain that could serve as a membrane anchor (type II membrane