Towards the identification of the common features of bacterial biofilm development

Summary. Microorganisms can live and proliferate as individual cells swimming freely in the environment, or they can grow as highly organized, multicellular communities encased in a self-produced polymeric matrix in close association with surfaces and interfaces. This microbial lifestyle is referred to as biofilms. The intense search over the last few years for factors involved in biofilm development has revealed that distantly related bacterial species recurrently make use of the same elements to produce biofilms. These common elements include a group of proteins containing GGDEF/EAL domains, surface proteins homologous to Bap of *Staphylococcus aureus*, and some types of exopolysaccharides, such as cellulose and the poly-1,6-N-acetylglucosamine. This review summarizes current knowledge about these three common elements and their role in biofilm development.

**Key words**: biofilms · PIA/PNAG · cellulose · c-di-GMP · GGDEF proteins · Bap protein

Introduction

If we queried microbiologists about landmark events in microbiology in the last decade, we would probably obtain a wide consensus focused on two subjects, the availability of complete bacterial genomes and the recognition that biofilms represent the natural way of bacterial life outside laboratory flasks. At the end of the nineteenth century, Robert Koch, Louis Pasteur, and others established the methodology to work with microbial pure cultures. Since then, microbiologists have carried out most of their studies with bacteria grown planktonically (swimming freely in liquid media). However, direct observation of bacteria in natural settings has shown that they usually grow adhered to surface–liquid or liquid–air interfaces and embedded in a self-produced extracellular polymeric matrix [10–12]. Examples of this bacterial lifestyle are abundant in daily life: the slimy material that covers flower vases, pipelines, submerged rocks, and even the surface of teeth (Fig. 1).

Biofilm formation occurs through sequential steps in which the initial attachment of planktonic bacteria to a solid surface is followed by their subsequent proliferation and accumulation in multilayer cell clusters, and the final formation of the bacterial community enclosed in a self-produced polymeric matrix. Once the structure has developed, some bacteria are released into the liquid medium, enabling the biofilm to spread over the surface [24,25,44,55] (Fig. 2).

The composition of the extracellular matrix is complex and variable among different bacterial species and even within the same species under different environmental conditions. Despite their heterogeneous composition, exopolysaccharides are an essential compound of the biofilm matrix, providing the framework into which microbial cells are inserted [8]. Among the many different exopolysaccharides that have been described, cellulose and β-1,6-linked N-acetylglucosamine are the most common components of...
the biofilm matrix of many different bacteria. The synthesis of exopolysaccharides incorporated into the extracellular matrix is highly regulated, and recent evidence has revealed that different bacteria use the same secondary signal, c-di-GMP, for this purpose. Since the levels of c-di-GMP in the cytoplasm depend on the activity of proteins containing a GGDEF/EAL domain, these proteins also represent a common factor in biofilm development. Besides exopolysaccharides, surface proteins also play an important role in biofilm formation. Most of the surface proteins involved in biofilm formation share several structural and functional features, and, consequently, the existence of a group of surface proteins has been proposed [34]. The first member of this group was described in a Staphylococcus aureus bovine mastitis isolate and was named Bap, for biofilm associated protein.

Bearing in mind that biofilm formation represents the normal lifestyle of bacteria in the environment, and that all microbes can make biofilms, it is reasonable to assume that there are a set of common principles for biofilm formation. This review summarizes evidence showing that cellulose,
β-1,6-linked N-acetylglucosamine, GGDEF-domain-containing proteins, and Bap-related proteins are common elements of the biofilm formation process among bacteria.

**Common exopolysaccharides in the biofilm matrix**

The efforts of many laboratories allowed the purification and characterization of the main exopolysaccharide compound of the biofilm matrix of *Staphylococcus epidermidis*—a linear homoglycan of β-1,6-linked N-acetylglucosamine, named PIA, for polysaccharide intercellular adhesin [4,9,36,37,60]. Later, a PIA-related exopolysaccharide, PNAG, whose chemical composition is based on that of poly-β-1,6-linked N-acetylglucosamine, was described in *S. aureus* [41,38]. The double assignation of PIA and PNAG to a similar polysaccharide in two closely related bacteria caused some confusion in the literature. Moreover, it has been shown that PIA and PNAG are structurally and immunologically identical [51], and that both are synthesized by the action of four homologous proteins (IcaA, IcaD, IcaB, and IcaC) encoded by genes organized in a single operon (*icaADBC*) [13,26].

During the screening of a collection of transposon mutants to identify genes required for biofilm formation in *Escherichia coli*, a locus (*pgaABCD*) exhibiting significant similarity to the *icaADBC* operon of *S. aureus* was identified [63]. PgaABCD proteins are involved in the synthesis of an unbranched β-1,6-linked N-acetylglucosamine identical to PIA/PNAG. A search for conserved protein domains predicted that PgaC is a N-glycosyltransferase homologous to IcaA (35% amino acid identity and 57% similarity). PgaB is a lipoprotein with putative polysaccharide N-deacetylase domains similar to those of IcaB. PgaA and PgaD have no functional homologies. As predicted, PGA exopolysaccharide serves as an adhesive that stabilizes biofilm formation under a variety of growth conditions, although it is not the only polysaccharide that is necessary for biofilm development in *E. coli*.

Another example of the relationship between β-1,6-linked N-acetylglucosamine and the biofilm formation process is provided by the periodontopathogen *Actinobacillus actinomycetemcomitans* [31,32]. Kaplan et al. [32] identified a protein, dispersin B, that is capable of degrading its own biofilm matrix as well as the biofilm matrix of *S. epidermidis*. Moreover, they also showed that the exopolysaccharide produced by *pgaABCD* cross-reacted with an antiserum raised against PIA/PNAG of *S. aureus*. Homologues of *icaADBC* and *pgaABCD* have also been described in the genomes of several other bacteria, including *Yersinia pestis*, *Yersinia enterocolytica*, *Bordetella pertussis*, *Bordetella bronchiseptica*, *Burkholderia cepacia*, *Pseudomonas fluorescens*, *Ralstonia solanacearum*, *Photobacterium profundum*, and *Chromobacterium violaceum* [22,32]. The presence of these genes strongly suggests that production of the biofilm matrix component β-1,6-linked N-acetylglucosamine exopolysaccharide is widely conserved among eubacteria.

Cellulose production is considered to be characteristic of the plant kingdom; however, the ability of some bacteria to produce extracellular cellulose has long been recognized. *Gluconacetobacter xylinus* (formerly called *Acetobacter xylinum*) has served as the model organism for elucidating the basic features of cellulose biosynthesis in bacteria [48–50]. The production of cellulose has also been described in *Agrobacterium tumefaciens* and *Rhizobium leguminosarum* bv. *Trifolii*, some species of cyanobacteria, and the gram-positive bacterium *Sarcina ventriculi* [40,42,43,48]. Initially, the proposed function for cellulose in these bacteria was not linked with biofilm formation. Instead, cellulose production was thought to be involved in the maintenance of *G. xylinum* in anoxic liquid environment and for the anchoring of *Rhizobium* and *A. tumefaciens* to plant tissue. Recent studies on the biofilm formation process of *Salmonella enterica* ser. Typhimurium (S. Typhimurium) and *E. coli* have revealed that these bacteria, as well as many other species of the Enterobacteriaceae family, such as *Citrobacter* spp., *Enterobacter* spp., and *Klebsiella* spp., produce cellulose as a crucial component of the bacterial extracellular matrix [53,64,65].

Studies carried out in a murine and a chicken model to determine the role of cellulose production on *Salmonella* virulence showed that cellulose does not play a significant role in the pathogenic process of *Salmonella*. Nevertheless, cellulose-deficient mutants proved to be more sensitive to chlorine treatment, suggesting that cellulose production and biofilm formation are important to bacterial survival on surface environments [53]. The connection between cellulose and biofilm formation has encouraged a reevaluation of the function of cellulose production by *A. tumefaciens*. In this bacterium, cellulose production is related to biofilm formation on plant roots and does significantly affect virulence [39].

Besides the finding that the same exopolysaccharides are used by diverse bacterial species to produce the biofilm matrix, it is particularly interesting that a single bacterial species, such as *E. coli*, has the capacity to produce several different extracellular polysaccharides, including β-1,6-linked N-acetylglucosamine, cellulose, and colanic acid. This situation raises many intriguing questions: Are these different exopolysaccharides simultaneously produced? What properties does each polysaccharide provide to bacteria?
rial multicellular behavior? What is the regulatory mechanism that guarantees the coordinated expression of different exopolysaccharides?

**GGDEF-domain-containing proteins and biofilm formation**

In a pioneering study that characterized the regulatory elements involved in cellulose production in *G. xylinum*, Benziman and colleagues [50] discovered that cellulose synthase was allosterically activated by c-di-GMP, and proposed a hypothetical model for the regulation of cellulose synthesis in this bacterium. In this model, they anticipated that the level of c-di-GMP in the cell is affected by the opposing action of two enzymes, diguanylate cyclase and phosphodiesterase (Fig. 3). The function of c-di-GMP in the bacterial physiology almost went unnoticed until a key paper by Tal et al. [56] related c-di-GMP metabolism with proteins containing a GGDEF and an EAL domain. These domains contain approximately 180 and 240 residues, respectively, and they very often include the conserved motifs Gly-Gly-Asp-Glu-Phe and Glu-Ala-Leu, which are used to name such domains. Bacterial genome sequencing has shown that GGDEF and EAL proteins are present in the majority of bacteria [19], starting from the ancient hyperthermophilic *Thermotogales* [30]. Thus, over 2200 proteins with GGDEF/EAL domains are currently contained in protein databases. Interestingly, these domains are not present in any protein encoded by the genomes of any Archaea or Eukarya, suggesting that c-di-GMP is a trait exclusive to the bacterial kingdom.

In the last two years, there has been an explosion of functional descriptions of GGDEF and EAL domain proteins in various bacteria. Most of these proteins have been related to exopolysaccharide production in biofilms [5,6,7,16,20,21,23,33,45,51,54,59] but there are also examples of GGDEF proteins involved in the differentiation process of *Caulobacter crescentus* [1], motility and sessility in *Escherichia*, *Salmonella* and *Pseudomonas aeruginosa* [28,52], photosynthesis gene expression in *Synechococcus elongatus* [58], and virulence in *S. Typhimurium* [27].

GGDEF and EAL domains are usually found in multidomain proteins linked to signal sensing and signal transduction domains. This location suggests that the GGDEF domain represents a novel, complex, bacterial signal-transduction network that is able to convert signals from different cellular compartments into the production of a secondary messenger, c-di-GMP [17,18,29]. The question is how the different signals can induce the appropriate response, if all of them converge in a central, freely diffusible cellular pool of c-di-GMP. An illustrative example showing that different GGDEF proteins of the same bacteria can influence the same metabolic pathway was found by our group in *S. Typhimurium* [20]. Eleven proteins containing the GGDEF domain had been annotated in the complete genome of *S. Typhimurium*. Apart from AdrA, a GGDEF protein required for cellulose synthesis, none of the remaining GGDEF proteins had been characterized [47]. A systematic complementation experiment, in which each of these GGDEF proteins was expressed from a multicopy plasmid in an AdrA mutant, revealed that four other GGDEF proteins (Stm1987, Stm4551, YegE, Stm3388) bestowed the capacity to produce cellulose to the *adrA* mutant [20]. The simplest interpretation of these results is that, as GGDEF proteins share the final product of their enzymatic activity, overexpression of one of these proteins can bypass the absence of other members of the family.

![Fig. 3. Structure of the cyclic dinucleotide c-di-GMP.](image-url)
The observation that many bacterial genomes have multiple open reading frames with putative GGDEF/EAL domains also raises several interesting questions. First, why is the number of putative GGDEF/EAL domain-containing proteins so variable among diverse bacterial species: 66/33 (GGDEF/EAL domain-containing proteins) in the Vibrio vulnificus genome, 33/21 proteins in P. aeruginosa, 19/17 proteins in E. coli, 4/3 in Bacillus subtilis, 2/0 in S. aureus, and 0/0 proteins in Chlamydia trachomatis, Mycoplasma pneumoniae, and Helicobacter pylori [19,46]? Second, if the perturbation of c-di-GMP levels caused by disruption of a single member of the family often leads to physiological changes, why can the activity of the remnant members of the family not compensate the deficiency of the one member of the family? Third, how do bacteria coordinate the expression and activity of all these proteins so that adequate levels of c-di-GMP are reached? Fourth, often a GGDEF and EAL domain occur in the same protein, whereby the GGDEF domain is located N-terminal to the EAL; if the two domains are responsible for opposing functions, is there a hierarchical or additive relationship between them? Fifth, why have most of these proteins gone unnoticed up to now? Sixth, how does c-di-GMP influence the activity of the target proteins? Concerning the last question, a very recent paper of Amikan and Galperin [2] described the identification of a new domain, PilZ, which is present in the sequences of bacterial cellulose synthases, an alginate biosynthesis protein, and many other proteins. Based on sequence analysis and on the phenotypes already described for mutants in proteins containing this domain, the authors proposed that PilZ is a part of the bacterial c-di-GMP binding protein.

It becomes obvious that, in the next few years, characterization of GGDEF/EAL proteins and the signal transduction mechanism mediated by c-di-GMP will be a very exciting area of research in microbiology.

**Bap-related proteins in biofilm formation**

A group of surface proteins sharing several structural and functional features is emerging as an important element of biofilm formation by diverse bacterial species [34]. The first member of this group was identified in a S. aureus mastitis isolate, strain V329, and was named Bap (biofilm associated protein). As common structural features, all Bap-related proteins: (i) are present on the bacterial surface; (ii) are of high molecular weight; (iii) contain a core domain of tandem repeats; (iv) confer on bacteria the capacity to form a biofilm; and (v) play a relevant role during bacterial infectious processes.

We identified Bap during screening of a library of mutants generated from S. aureus strain V329, and found that the protein is essential for biofilm formation [14]. Primary attachment, intercellular aggregation, and biofilm formation studies showed that Bap promotes both primary attachment to abiotic surfaces and intercellular adhesion. Since disruption of the bap gene provoked a decreased accumulation of PIA/PNAG, the main exopolsaccharide of the biofilm matrix of S. aureus, our first interpretation was that biofilm deficiency of Bap mutant strains was mainly caused by a decreased accumulation of PIA/PNAG. However, when the icuADBC operon was inactivated in a bap-positive strain, the bacteria retained the capacity to produce biofilm [15]. In agreement with this result, we have observed that all the staphylococcal isolates tested that harbor the bap gene are strong biofilm producers, despite the fact that most of them do not contain the icuADBC operon [62]. These findings indicate that Bap is able to mediate a mechanism of biofilm development that differs from the PIA/PNAG-dependent mechanism.

Bap of S. aureus is a 2276-amino-acid surface protein with a multidomain architecture characteristic of surface-associated proteins from gram-positive bacteria (Fig. 4). The N-terminal domain (819 amino acids) includes a signal sequence for extracellular secretion and, for the most part, is devoid of repetitions. The region encompassing amino acid residues 820–2148 is composed of 13 identical repeats, each of which is 86 amino acids long. The C terminus of Bap contains a typical cell-wall attachment region comprising an LPXTG motif, a hydrophobic transmembrane sequence, and positively charged amino acids. Bap contains four sites with >80% similarity to the consensus EF-hand motif. It has been shown that addition of millimolar amounts of calcium to the growth medium inhibits intercellular adhesion and biofilm formation, whereas calcium depletion enhances bacterial aggregation in liquid medium [3]. Gene bap has never been found in S. aureus human isolates, strongly suggesting that human and ruminant mastitis isolates are not clonally related, and that specific host-dependent pathogenic factors evolved independently in humans and ruminants. In spite of this, the presence of bap is not restricted to Staphylococcus aureus, and bap ortholog genes have been found in isolates of several coagulase-negative staphylococcal species, including S. epidermidis, S. chromogenes, S. xylosus, S. simulans, and S. hyicus [62].

Since the finding of Bap, different Bap-related surface proteins involved in biofilm formation by diverse bacterial species have been described. Among these proteins, our group has shown that Esp (enterococcal surface protein) of Enterococcus faecalis and the protein encoded by stm2689 of S. Typhimurium are involved in biofilm formation by these...
bacteria [35]. Esp is a 1873-amino-acid surface protein, and its N-terminal domain shows 33% homology with that of Bap (Fig. 4). The central part of Esp has a series of nearly identical 246-nucleotide tandem C-repeats. Sequence analysis comparison of the C-repeats of Bap and Esp shows 33% identity. An epidemiological study demonstrated a strong correlation between the presence of Esp and the ability of *E. faecalis* to form biofilms in vitro [61]. None of the esp-deficient isolates tested in that study were capable of forming biofilms and expression of Esp in an esp-negative strain restored biofilm formation by *E. faecalis* [57,61]. The functional relationship between Bap and Esp was unambiguously demonstrated when heterologous expression of Bap bestowed biofilm formation capacity to an *E. faecalis* esp-deficient and biofilm-deficient strain (our unpublished results).

*Stm2689* codes for the second largest protein of the *S. enterica* ser. Typhimurium (*S. Typhimurium*) genome. The protein contains 28 tandem imperfect repeats of 86–106 amino acids (amino acids 159–3003). Each repeat of *Stm2689* shows, on average, a 29% identity with a C-repeat of Bap (Fig. 4). Given the in silico homology exhibited by *Stm2689* and Bap, we completely deleted *stm2689* in a clinical *S. enteritidis* strain. The mutant strain lost the capacity to form a biofilm pellicle in Luria-Bertani (LB) medium at the air–broth interface, whereas overexpression of a chromosomal copy of *stm2689* increased the thickness and strength of the biofilm formed. The deficiency of *Stm2689* could be counteracted by overproduction of curli fimbriae but not cellulose. This result supports the idea that *Stm2689* plays a role complementary to that of fimbriae in connecting cells, either by strengthening fimbriae-mediated interactions or by allowing the interconnection of bacteria separated by relatively long distances. Based on these results, *stm2689* was renamed as *bapA* [35].

To promote biofilm development, Bap proteins might interact with themselves through homophilic mechanism, functioning as both receptor and ligand between two bacterial clusters. In support of this hypothesis, we have found that when BapA-overexpressing and BapA-minus cells were mixed together, the pellicle almost exclusively contained BapA-positive bacteria [35]. These results suggest that the presence of BapA on the surface of the cell is necessary to mediate bacterial recruitment into the biofilm pellicle, and that the presence of BapA on the surface of one bacterium might not compensate for the absence of the protein on the surface of a BapA-deficient bacterium.

**Conclusions**

Significant progress in the understanding of many key aspects of bacterial biofilm development has been made in the last few years. One consequence of these advances is the identification of elements persistently involved in biofilm formation in distantly related bacteria. Some of the common elements were described in this review, but the list is probably far from being complete and new common elements will no doubt be identified in the near future. Identification of these common elements, used by different bacteria to produce biofilms under different environmental conditions, will provide important clues about the general mechanism of bacterial multicellular behavior.

**Acknowledgements.** I express my gratitude to everybody in my lab and to the group of José R. Penadés for their invaluable work in the staphylococcal and *Salmonella* biofilm formation studies. I also thank Cristina Solano and Cristina Latasa for reviewing the manuscript. This work was supported by the BIO2002-04542-C02 and GEN2003-20234-C06-05 grants.
from the Comisión Interministerial de Ciencia y Tecnología (Madrid, Spain) and grant 17/2004 from the Departamento de Educación y Cultura del Gobierno de Navarra, Spain.

References


Hacia la identificación de las características comunes del desarrollo de biopelículas bacterianas

**Resumen.** Los microorganismos pueden vivir y proliferar como células individuales que nadan libremente en el medio o crecer en comunidades multicelulares muy bien organizadas dentro de una matriz que ellas mismas han sintetizado y asociadas a superficies o interfaces. Esta forma de vida microbiana recibe el nombre de biopelículas. La búsqueda intensa de los factores que intervienen en el desarrollo de la biopelícula llevada a cabo durante los últimos años ha revelado que especies bacterianas filogenéticamente alejadas recurren a los mismos elementos para producir la biopelícula. Entre los elementos comunes identificados hay proteínas que contienen los dominios GGDEF/EAL, proteínas de superficie que muestran homología con la proteína Bap de *Staphylococcus aureus*, y algunos exopolisacarídeos, como la celulosa y la poli-β-1,6-N-acetilgalactosamina. Esta revisión resume los conocimientos actuales sobre estos tres elementos y su función en la formación de la biopelícula. [Int Microbiol 2006; 9(1):21-28]

**Palabras clave:** biopelículas · PIA/PNAG · celulosa · c-di-GMP · proteínas GGDEF · proteína Bap

Para a identificação das características comuns do desenvolvimento de biofilmes bacterianos

**Resumo.** Os microorganismos podem viver e proliferar como células individuais que nadam livremente no meio ou crescem em comunidades multicelulares muito bem organizadas em uma matriz sintetizada pelos próprios microorganismos e associadas a superfícies ou interfaces. Esta forma de vida microbiana recebe o nome de biofilmes. A busca intensa de fatores implicados no desenvolvimento do biofilme durante os últimos anos revelou que espécies bacterianas filogeneticamente afastadas recorrem aos mesmos elementos para produzir o biofilme. Os elementos comuns identificados incluem proteínas que contêm os domínios GGDEF/EAL, proteínas de superfície que apresentam homologia com a proteína Bap de *Staphylococcus aureus* e exopolissacarídeos como celulose e poli-β-1,6-N-acetilgalactosamina. Esta revisão resume os conhecimentos atuais em relação a estes três elementos e sua função na formação do biofilme. [Int Microbiol 2006; 9(1):21-28]

**Palavras chave:** biofilme · PIA/PNAG · celulose · c-di-GMP · proteínas GGDEF · proteína Bap