



Computational analysis suggests that virulence of *Chromobacterium violaceum* might be linked to biofilm formation and poly-NAG biosynthesis

Sidnei Becker, Cíntia Soares and Luismar Marques Porto

Laboratório de Tecnologias Integradas, Departamento de Engenharia Química e de Alimentos, Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil.

Abstract

Groups of genes that produce exopolysaccharide with a *N*-acetyl-D-glucosamine monomer are in the genome of several pathogenic bacteria. *Chromobacterium violaceum*, an opportunistic pathogen, has the operon hmsHFR-CV2940, whose proteins can synthesize such polysaccharide. In this work, multiple alignments among proteins from bacteria that synthesize such polysaccharide were used to verify the existence of amino acids that might be critical for pathogen activity. Three-dimensional models were generated for spatial visualization of these amino acid residues. The analysis carried out showed that the protein HmsR preserves the amino acids D135, D228, Q264 and R267, considered critical for the formation of biofilms and, furthermore, that these amino acids are close to each other. The protein HmsF of *C. violaceum* preserves the residues D86, D87, H156 and W115. It was also shown that these residues are also close to each other in their spatial arrangement. For the proteins HmsH and CV2940 there is evidence of conservation of the residues R104 and W94, respectively. Conservation and favorable spatial location of those critical amino acids that constitute the proteins of the operon indicates that they preserve the same enzymatic function in biofilm synthesis. This is an indicator that the operon hmsHFR-CV2940 is a possible target in *C. violaceum* pathogenicity.

Key words: biofilms, exopolysaccharide, *Chromobacterium violaceum* pathogenicity, comparative genomics.

Received: November 4, 2008; Accepted: May 25, 2009.

Introduction

Some microorganisms develop cooperative strategies in the formation of biofilms. Biofilm can be defined as interdependent communities of microorganisms, usually connected with a surface presenting high resistance to environmental stress. It consists of a complex symbiotic system of great importance to biotechnological and medical applications, since it is correlated with bacterial resistance to antibiotics and promotion of lethal infections.

The formation of biofilms in several microorganisms involves the presence of a complex matrix where the polymer poly(beta-1,6-*N*-acetyl-D-glucosamine), poly-NAG, can play an important role (Itoh *et al.*, 2005), possibly related to resistance to antibiotics and promotion of bacterial infections. Poly-NAG has been implicated in the formation of biofilms in several pathogen bacteria. The polymer, which is involved in cell adhesion in *Staphylococcus epidermidis* (Mack *et al.*, 1996), is also involved in abiotic

surface binding, and in intercellular adhesion formation of biofilm in *Escherichia coli* (Itoh *et al.*, 2005).

The hmsHFERS operon has been involved in the formation of biofilms *in vitro* in *Yersinia pestis*. This operon was related to the blockage of the digestive system of fleas, and it is associated with the transmission of *Y. pestis* to mammals (Jarrett *et al.*, 2004). The bacterium *Bordetella pertussis* has the operon bpsABCD related to a polymer that contributes to the formation of biofilms in the respiratory tract of rats (Sloan *et al.*, 2007). An understanding of the relationship between the operons of those bacteria in the formation of biofilms could potentially lead to the development of a vaccine applicable to different bacterial species.

Recently, the sequencing of the complete genome of the *Chromobacterium violaceum*, strain ATCC 12472 (Brazilian National Genome Project Consortium, 2003; www.brgene.lncc.br; www.ncbi.nlm.nih.gov: access number NC_005085) was carried out to promote domestic genome projects. The investment was justified due to the great biotechnological and pharmaceutical potential of this microorganism. *Chromobacterium violaceum* is considered a possible pathogen for humans and animals, with several reported cases of infection in humans and animals in tropical and subtropical areas, where it is normally found.

Its potential pathogenicity in humans was first described in 1927 in Malaysia (Sneath *et al.*, 1953). Although cases of infection with *C. violaceum* are rare, it has a mortality rate of more than 57%, as reported in a case of chronic granulomatous disease in children (Macher *et al.*, 1982). *C. violaceum* reveals some potentially pathogenic genes (Brazilian National Genome Project Consortium, 2003); however, their extent is not yet known or how these genes are organized in their genome.

This work aims to link, through comparative analysis, the genes *hmsH*, *hmsF*, *hmsR* and the ORF CV2940 of *C. violaceum* with the genes of organisms known and related to the formation of biofilm in an effort to shed light on the pathogenic potential of *C. violaceum*. The analysis of the *C. violaceum* genome reveals the presence of possible genes involved in the formation of biofilms, and the sequence of the group of genes *hmsHFR* and CV2940 are potentially functional for it. Initially, the relationship among the proteins is carried out by the alignment of amino acid sequences and it is complemented by structural modeling in order to identify the importance of the amino acid through spatial viewing.

Methods

Comparisons were carried out using the amino acid sequences of the organisms listed in Table 1 and obtained from GenBank. These organisms with their genomic regions were selected for this study because of their confirmed connection with biofilm formation or production of polysaccharides such as poly(beta-1,6-*N*-acetyl-D-glucosamine).

Table 1 - Bacteria with their genomic regions used in the comparative study.

Organism and GenBank access number	GenBank access number	Gene/ORF
<i>Chromobacterium violaceum</i> ATCC 12472	NP_902610	CV2940
	NP_902611	<i>hmsR</i>
	NP_902612	<i>hmsF</i>
	NP_902613	<i>hmsH</i>
<i>Escherichia coli</i> K12	P69435	<i>pgaA</i>
	P75906	<i>pgaB</i>
	P75905	<i>pgaC</i>
	P69433	<i>pgaD</i>
<i>Staphylococcus epidermidis</i>	AAZ78357	<i>icaA</i>
	AAZ78358	<i>icaD</i>
	AAZ78359	<i>icaB</i>
	AAZ78360	<i>icaC</i>
<i>Yersinia pestis</i> KIM6+	AAB66588	<i>hmsH</i>
	AAB66589	<i>hmsF</i>
	AAB66590	<i>hmsR</i>
	AAB66591	<i>hmsS</i>
<i>Bordetella pertussis</i> Tohama I	NP_880625	BP1941
	NP_880626	<i>hmsR</i>
	NP_880627	<i>hmsF</i>
	NP_880628	<i>hmsH</i>

An initial approach among groups of proteins (Table 1) was performed using the Clustal W server (Jeanmougin *et al.*, 1998) through multiple alignments in order to identify possible amino acids conserved among the proteins. The structural modeling was performed by 3D-JIGSAW (Bates *et al.*, 2001) and PHYRE servers (Kelley *et al.*, 2000) with the models with higher identity chosen as the template. The ProCheck program (Laskowski *et al.*, 1996) was used to check the quality of the three-dimensional model generated. The quality of the model was analyzed through the Ramachandran diagram (Ramachandran *et al.*, 1963). Analyses were carried out using the Verify3D program (Eisenberg *et al.*, 1997) in order to ensure a more solid and comprehensive quality of the model.

Results

In a multiple alignment comparison generated by the Clustal W server with other glycosyltransferase proteins, the critical amino acids related to formation of biofilms are preserved also in the HmsR protein of *C. violaceum* (Figure 1). Conserved amino acids that are critical for function of HmsR of *Y. pestis* (D176, D269, Q305 and R308) correspond to residues D135, D228, Q264 and R267 in *C. violaceum*, respectively.

The structural model generated (Figure 2) comprises the critical amino acids for the function of this protein in the formation of biofilms arranged in a three dimensional conformation with the possibility of fitting a substrate. The residue D228 is in exactly the same position as D191 (SpsA) with other conserved residues (D135, Q264 and R267) also positioned on the active site.

The multiple alignment among organisms in this study shows that the polypeptide HmsF of *C. violaceum* retains the amino acids that are crucial to the formation of extracellular matrix in biofilms (Figure 3). The D114 and D115 aspartate amino acids of the protein HmsF of *Y. pestis*, which have HmsF deacetylase activity functionality, are conserved on all related organisms in this study, and in *C. violaceum* correspond to aspartate D86 and D87. The important amino acids in the synthesis of extracellular matrix in *Y. pestis* (W143 and H184), which correspond in *C. violaceum* to W115 and H156, are also retained.

In the three-dimensional model of HmsF of *C. violaceum* obtained (Figure 4), the 3D-JIGSAW server made use of the protein SpPgdA as a template (PDB code: 2c1g) (Blair *et al.*, 2005), which is a *N*-acetyl-glucosamine deacetylase. The model generated for HmsF of *C. violaceum* presenting the spatial location of critical residues in the biofilm formation to *Y. pestis* shows that the residues D86, D87, W115 and H156 are closely located in space (Figure 4). These residues are situated around the active site.

In the multiple alignment of HmsH protein, it was observed that the residue arginine R113, which had a poor performance in the formation of biofilms in *Y. pestis* (For-

<i>Y. pestis</i>	121	I A I N D G S S D D	T A Q V L D A L L A	E D P R L R V I H L	A H N Q G K A I A L	R M G A A A A R S E	Y L V C I D S D A L
<i>E. coli</i>	108	I A V N D G S T D K	T R A I L D R M A A	Q I P H L R V I H L	A Q N Q G K A I A L	K T G A A A A K S E	Y L V C I D G D A L
<i>C. violaceum</i>	80	I A V N D G S R D G	T A A I L N Q M A Q	E H P R L R V I H Q	A H N Q G K A V G L	N T A L L L S R Y E	H L I C I D G D A L
<i>B. pertussis</i>	83	I A V N D G S R D N	T G A L L N E L A E	Q Y P R L L V V H Q	S R N E G K A I G L	N T A A Q L S D A E	F L L C I D G D S L
<i>S. epidermidis</i>	80	I I I N D G S S D N	T A E I I Y D F K K	N H D - F K F V D L	E V N R G K A N A L	N E G I K Q A S Y E	Y V M C I D A D T V
<i>Y. pestis</i>	181	L D K N A V F Y L V	A P L I A N P R T G	A V T G N P R I R T	R S T L I G R V Q V	G E F S S I I G L I	K R T Q R V Y G Q V
<i>E. coli</i>	168	L D R D A A A Y I V	E P M L Y N P R V G	A V T G N P R I R T	R S T L V G K I Q V	G E Y S S I I G L I	K R T Q R I Y G N V
<i>C. violaceum</i>	140	L D P H A A K W L M	R H F Q L G S R V G	A V T G N P R I R N	R T T L L G R L Q V	G E F T S I I G L I	K R A Q R T Y G R L
<i>B. pertussis</i>	143	P H P D S I T Y M L	T H F L H A G H V G	A V T G N P R I R N	R S T V L G R M Q V	G E F S S I V G L I	K R T Q Q L Y G K L
<i>S. epidermidis</i>	139	I D D D A P F Y M I	E D F K K N P K L G	A V T G N P R I R N	K S S I L G K I Q T	I E Y A S I I G C I	K R S Q S L A G A I
<i>Y. pestis</i>	241	F T V S G V V A A F	R R R A L A D V G Y	W S P D M I T E D I	D I S W K L Q L K H	W S V F F E P R G L	C W I L M P E T L R
<i>E. coli</i>	228	F T V S G V I A A F	R R S A L A E V G Y	W S D D M I T E D I	D I S W K L Q L N Q	W T I F Y E P R A L	C W I L M P E T L K
<i>C. violaceum</i>	200	F T V S G V I A A F	R K T A V V Q A G Y	W S P D M L T E D I	D I S W K I Q L Q H	W D V R F E P R A L	V W I L M P E T L K
<i>B. pertussis</i>	203	M T V S G V M T M F	R K Q A L H Q V G Y	W S P D M Q T E D I	D I S W K L Q L A G	W T L R Y E P R A L	T W I L M P E T F R
<i>S. epidermidis</i>	199	N T I S G V F T L F	K K S A L K D V G Y	W D T D M I T E D I	A V S W K L H L F D	Y E I K Y E P R A L	C W M L V P E T I G
<i>Y. pestis</i>	301	G L W H Q R I R W A	Q G G A E V F L K N	M F K L W R W R N R	R M W L L F L E Y S	L S I T W A F T Y L	F S I T I Y L L G L
<i>E. coli</i>	288	G L W H Q R I R W A	Q G G A E V F L K N	M T R L W R K E N F	R M W P L F F E Y C	L T T I W A F T C L	V G F I I Y A V Q L
<i>C. violaceum</i>	260	G L W H Q R I R W A	K G G T Q A L L R N	S E I F G S W R F R	R M W P V Y V E Y F	F S V V W S Y V M A	F I I I W W I V D V
<i>B. pertussis</i>	263	G L P H Q R Y R W A	R G G I E T A L K Y	A P R M L N P R Q I	L M W P I P L E F S	L S V L W A Y A M L	F L I C M A V L G Q
<i>S. epidermidis</i>	259	G L W H Q R V R W A	Q G G H E V L L R D	F W P T I K T K K L	S L Y I L M F E Q I	A S I T W V Y I V L	C Y L S F L V I T A

Figure 1 - Partial multiple alignment of HmsR from *C. violaceum* compared with HmsR of *Y. pestis*; PgaC of *E. coli*; IcaA of *S. epidermidis* and HmsR of *B. pertussis*. The critical residues aspartate (D), glutamine (Q) and arginine (R) are demarcated.

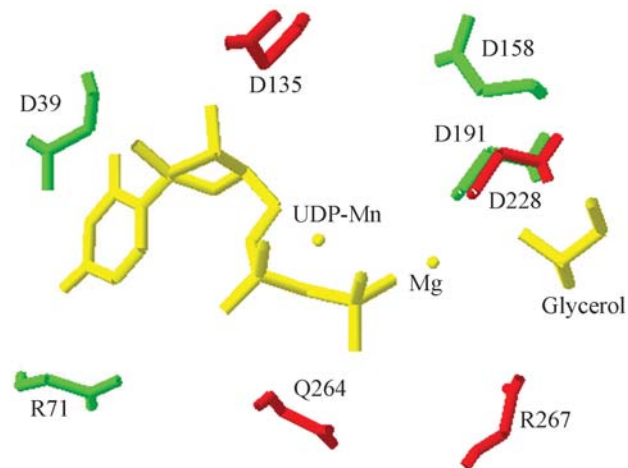


Figure 2 - Superimposed proteins closing up the active site of SpsA protein bound to Mn-UDP with the built model of HmsR protein *C. violaceum*. The residues of SpsA are green with HmsR residues in red and the UDP-Mn, Mg and Glycerol in yellow.

<i>Y. pestis</i>	89	V S V D Q I L A A R	N G G P T L P D K A	V L L T H D D G Y S	S F Y R R V Y P L L	K A Y K W S A V L A
<i>E. coli</i>	89	V S I A Q I R E A H	R G G K P L P E K A	V L L T H D D G Y Q	S F Y T R V F P I L	Q A F Q W P A V W A
<i>B. pertussis</i>	88	I S M Q Q I I D S R	A G V R P L P P R P	I L L T H D D G Y A	S T Y T K V F P L L	K K F N Y P A V V A
<i>C. violaceum</i>	61	V S L D D A M A A R	A G G P A L P D K S	V L L S H D D G Y R	S V Y T Q V Y P V L	K A F R A P A L I G
<i>S. epidermidis</i>	96	L T L K E F I K Y K	E K G K - F P K R S	V W I N F D D M D Q	T I Y D N A F P V L	K K Y H I P A T G F
<i>Y. pestis</i>	139	P V G T W I D T A T	D K K V D F G G - L	S T D R D R F A T W	K Q I T E M S K S G	L V E I G A H T Y A
<i>E. coli</i>	139	P V G S W I D T P A	D K Q V K F G D - E	L V D R E Y F A T W	Q Q V R E V A R S R	L V E L A S H T W N
<i>B. pertussis</i>	138	V V T S W I D A P A	G T K I R L S P K I	E V P H D F F M T W	A Q L R E M A Q S G	L V E L A S H B H N
<i>C. violaceum</i>	111	L V G S W L S P P E	G G S V N F A G - V	K V P R D D F L S W	G Q I R E M Q A S G	L V E V A N H T P D
<i>S. epidermidis</i>	145	L I T N H I G S T N	F H N L N L L S - -	-----K	K Q L D E M Y E T G	L W D F E S H T H D

Figure 3 - Multiple partial alignment of HmsF of *C. violaceum* compared with PgaB of *E. coli*; IcaB of *S. epidermidis*; HmsF of *Y. pestis* and HmsF of *B. pertussis*. The critical residue to extracellular matrix formation in *Y. pestis*: D (aspartate), W (tryptophan) and H (histidine) are demarcated.

man *et al.*, 2006), and that in *C. violaceum* is the residue R104, was conserved in almost all organisms studied, but it was not observed in *S. epidermidis*, a Gram-positive bacteria (not shown). The multiple alignment performed among bacteria of this study for ORF CV2940 shows that the critical residue tryptophan W80 of *Y. pestis* is also conserved in Gram-negative bacteria *E. coli* and *C. violaceum*, corresponding to the residue W94 in *C. violaceum*.

The analysis of the three-dimensional model generated for HmsR of *C. violaceum* by the Ramachandran diagram shows that the model holds 99.1% of the residues in allowed regions, demonstrating a good quality. The analysis of the environment of each amino acid residue carried out by the Verify3D program presented two regions (residues 175 to 187 and 230 to 237) in the HmsR with a negative 3D-1D score, indicating a probable incorrect folding in this region. However, the critical residues for HmsR in the formation of biofilm do not appear in this region. The struc-

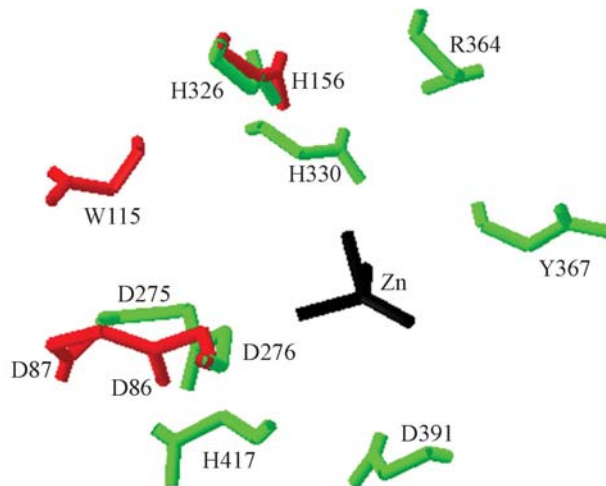


Figure 4 - Superimposed proteins *SpPgdA* and HmsF. The residues of *SpPgdA* are in green, the HmsF residues are in red and Zinc is in black.

tural model of HmsF protein generated also presents a good quality, with a Ramachandran plot showing 98.5% of the residues in allowed regions. The model also shows a good quality for HmsF when compared with the results obtained in the Verify3D program, providing a great number of amino acids in a positive 3D-1D score.

Discussion

The production of biofilms by several organisms suggests the conservation of proteins or a set of proteins with similar function for its synthesis.

The HmsR protein, an important glycosyltransferase in polysaccharide synthesis, maintains the critical residues for its function also in *C. violaceum*. Critical residues in the formation of biofilms in HmsR protein of *Y. pestis* (Forman *et al.*, 2006), which match in *C. violaceum* to aspartate D135 and D228, arginine R267 and glutamine Q264, are retained in all the bacteria shown in this study, indicating functional similarity for this protein. The spatial location of the critical residues for biofilm formation in an active site strongly emphasizes the enzymatic role that the HmsR protein of *C. violaceum* can perform in the synthesis of polysaccharide.

The HmsF protein of *C. violaceum* retains all the amino acids that are critical to the formation of biofilms. The generated model with the spatial location of critical residues in the formation of biofilms in *Y. pestis* shows that they are close to each other. The use of a deacetylase as template was important due to the role that this enzyme plays in the formation of biofilms. The deacetylation involves mainly aspartate (D) and histidine (H) residues (Blair *et al.*, 2005), and in this work these residues are located in a favorable spatial condition. Although actual evidence can be obtained from additional techniques such as substrate-enzyme reactions, it can be assumed that this model has a deacetylase catalytic region. The protein

HmsH preserves the residue arginine, which shows a moderate effect on the formation of biofilms in *Y. pestis* among all bacteria of this study, except for *S. epidermidis* which is a Gram-positive bacterium. In the three-dimensional model, the choice of the template of the OGT enzyme (PDB accession code 1w3b) was very useful, since this enzyme has an *N*-acetylglucosamine (GlcNAc) additive activity (Jinek *et al.*, 2004) that is the basic unit of the polymer poly(beta-1,6-*N*-acetyl-D-glucosamine).

Conservation and favorable spatial arrangement of those critical amino acids that constitute the proteins encoded by the elucidated operon indicate that they may express the necessary enzymatic function for resistant biofilm formation. Therefore the conclusion is that the operon *hmsHFR-CV2940* might be linked to *C. violaceum* pathogenicity.

References

- Bates PA, Kelley LA, Maccallum RM and Sternberg MJ (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins* 5:39-46.
- Blair DE, Schuttelkopf AW, Macrae JI and Van Aalten DM (2005) Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor. *Proc Natl Acad Sci USA* 102:15429-15434.
- Brazilian National Genome Project Consortium (2003) The complete genome sequence of *Chromobacterium violaceum* reveals remarkable and exploitable bacterial adaptability. *Proc Natl Acad Sci USA* 100:11660-11665.
- Eisenberg D, Luthy R and Bowie JU (1997) VERIFY3D: Assessment of protein models with three-dimensional profiles. *Meth Enzymol* 277:396-404.
- Forman S, Bobrov AG, Kirillina O, Craig SK, Abney J, Fetherston JD and Perry RD (2006) Identification of critical amino acid residues in the plague biofilm Hms proteins. *Microbiology* 152:3399-3410.
- Itoh Y, Wang X, Hinnebusch BJ, Preston JF and Romeo T (2005) Depolymerization of beta-1,6-*N*-acetyl-D-glucosamine disrupts the integrity of diverse bacterial biofilms. *J Bacteriol* 187:382-387.
- Jarrett CO, Deak E, Isherwood KE, Oyston PC, Fischer ER, Whitney AR, Kobayashi SD, Deleo FR and Hinnebusch BJ (2004) Transmission of *Yersinia pestis* from an infectious biofilm in the flea vector. *J Infect Dis* 190:783-792.
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG and Gibson TJ (1998) Multiple sequence alignment with Clustal X. *Trends Biochem Sci* 23:403-405.
- Jinek M, Rehwinkel J, Lazarus BD, Izaurrealde E, Hanover JA and Conti E (2004) The superhelical TPR-repeat domain of O-linked GlcNAc transferase exhibits structural similarities to importin alpha. *Nat Struct Mol Biol* 11:1001-1007.
- Kelley LA, Maccallum RM and Sternberg MJ (2000) Enhanced genome annotation using structural profiles in the program 3D-PSSM. *J Mol Biol* 299:499-520.
- Laskowski RA, Rullmann JA, Macarthur MW, Kaptein R and Thornton JM (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* 8:477-486.

- Macher AM, Casale TB and Fauci AS (1982) Chronic granulomatous disease of childhood and *Chromobacterium violaceum* infections in the southeastern United States. *Ann Intern Med* 97:51-55.
- Mack D, Fischer W, Krokotsch A, Leopold K, Hartmann R, Egge H and Laufs R (1996) The intercellular adhesin involved in biofilm accumulation of *Staphylococcus epidermidis* is a linear beta-1,6-linked glucosaminoglycan: Purification and structural analysis. *J Bacteriol* 178:175-183.
- Ramachandran GN, Ramakrishnan C and Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. *J Mol Biol* 7:95-99.
- Sloan GP, Love CF, Sukumar N, Mishra M and Deora R (2007) The *Bordetella* Bps polysaccharide is critical for biofilm development in the mouse respiratory tract. *J Bacteriol* 189:8270-8276.
- Sneath PH, Whelan JP, Bhagwan SR and Edwards D (1953) Fatal infection by *Chromobacterium violaceum*. *Lancet* 265:276-277.

Guest Editor: José Carlos Merino Mombach

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.