

## ETUDE D'UNE ADHESINE IDENTIFIEE CHEZ RUMINOCOCCUS GNAVUS E1

***Radia Alatou***

Faculté des Sciences de la Nature et de la Vie,  
Département de microbiologie, Université Constantine 1, Algérie

***Gwenola Simon***

***Michel Fons***

BiosCiences / ISM<sup>2</sup> UMR CNRS-6263,  
Université Paul Cézanne Aix Marseille III, France

---

### Abstract

*Ruminococcus gnavus* E1 is a Gram positive strict anaerobic bacterium that was isolated from the dominant faecal microbiota of a healthy adult. A 6kb-long open reading fragment called *radA* was identified on the E1 chromosome, next to the genetic clusters involved in the biosynthesis of the RumA and RumC bacteriocins which are active against pathogenic *Clostridium perfringens*. *radA* shares a high sequence homology with genes of *Staphylococcus aureus*, *Bacillus cereus* and *C. perfringens* encoding adhesins of the MSCRAMMs family. RT-PCR experiments demonstrated that the *radA* gene was strongly expressed *in vivo*, when the E1 strain colonized the digestive tract of monoxenics animals, while little transcription occurred *in vitro*. These results suggest that RadA could play an important role in the colonization of the digestive ecosystem.

---

**Keywords:** Faecal microbiot, adhesin, adhesion, factor for colonization

---

### Resume

*Ruminococcus gnavus* E1, une bactérie à Gram positif, anaérobie stricte, isolée du microbiote dominant d'un homme sain. Un cadre ouvert de lecture (ORF) d'environ 6kb, nommé *radA*, a été identifié sur le chromosome de la souche E1, à proximité des clusters génétiques impliqués dans la biosynthèse des bactériocines RumA et RumC qui sont actives contre le pathogène *Clostridium perfringens*. *radA* présente de fortes homologies avec des gènes de *Staphylococcus aureus*, *Bacillus cereus* et *C. perfringens* codant pour des protéines d'adhésion de la famille des MSCRAMMs. Il a été montré par RT-PCR que le gène *radA* est fortement exprimé *in vivo*, quand

la souche E1 colonise le tube digestif d'animaux monoxéniques, et peu transcrit *in vitro*. Les résultats suggèrent que RadA pourrait jouer un rôle important dans la colonisation de l'écosystème digestif.

---

**Mots-clés:** Microbiote intestinale, adhésine, adhésion, facteur de colonisation

## Introduction

*Ruminococcus gnavus* est une espèce bactérienne à Gram positif, anaérobie stricte, qui appartient au cluster phylogénétique *Clostridium cocoides* regroupant différentes espèces du microbiote dominant de l'Homme (Fons *et al.*, 2000). *R. gnavus* E1 a été isolée du microbiote fécal d'un adulte sain et a suscité l'intérêt des chercheurs en raison de sa capacité à produire les bactériocines Ruminococcine A (RumA) et Ruminococcine C (RumC) actives contre le pathogène *Clostridium perfringens*. RumA est produite *in vitro*, sous la dépendance de trypsine (Dabard *et al.*, 2001). RumC est produite *in vivo*, également sous la dépendance de trypsine (Croft, 2008). RumA et RumC semblent jouer un rôle dans la capacité de la souche E1 à coloniser le tube digestif et dans la protection de l'hôte contre certains pathogènes intestinaux. La colonisation de l'hôte peut faire intervenir plusieurs mécanismes comme la compétition pour un substrat ou une niche écologique, mais il existe d'autres stratégies de colonisation tel que l'adhésion bactérienne. Ce phénomène est dans la très grande majorité des cas, une étape obligatoire pour la bactérie : c'est le début du processus de colonisation chez les commensales comme chez les pathogènes (Charachon, 2007). Il a été établi que la colonisation des muqueuses humaines par les bactéries à Gram positif était possible grâce à la présence de molécules d'adhésion spécifique nommée adhésines (Jett *et al.*, 2002). Ces quelques adhésines exprimées à la surface des bactéries à Gram positif sont classées dans un groupe à part comme étant une nouvelle famille de protéines d'adhésion nommée : Microbial Surface Components Recognizing Adhesive Matrix Molecules (MSCRAMM<sub>S</sub>) (Sillanpää *et al.*, 2008). Ces protéines d'adhésion sont retrouvées et caractérisées uniquement chez les bactéries à Gram positif (Zong *et al.*, 2005). Parmi les mieux caractérisées on peut citer la protéine modèle CNA de *S. aureus* impliquée dans la liaison au collagène (Ton-That *et al.*, 2000) ainsi que CNE de *Streptococcus equi* (Lannergård *et al.*, 2003). Les MSCRAMM<sub>S</sub> ont été également identifiées chez d'autres bactéries comme *Enterococcus faecium* (Nallapareddy Sreedhar *et al.*, 2003), *Streptococcus mutans* et *Bacillus anthracis* (Xu *et al.*, 2004). Les adhésines matures partagent une organisation en plusieurs domaines fonctionnels (figure 1) : une séquence signal à l'extrémité N-terminale, une région A riche en acides aminés hydrophobes et qui comprendrait un site spécifique de fixation aux constituants de la MEC (Xu

*et al.*, 2004). Elle est suivie de domaines répétés B dont le nombre diffère d'une adhésine à une autre. En position C-terminale, une ancre caractérisée par un motif LPXTG ou NPQTN qui est clivé par une sortase ce qui permet l'ancrage de la protéine au niveau du peptidoglycane (Zong *et al.*, 2005; Talay *et al.*, 2005)



**Figure 1:** Structure générale des adhésines. **S** : Séquence signal, **A** : Région A, **B1. B2. B3** : Domaines répétés, **W** : Ancre, **M** : partie membranaire, **C** : partie cytoplasmique

(Q73CV3) de *Bacillus cereus* et (Q93M90) *Clostridium perfringens*. L'analyse de la protéine putative déduite du gène *radA* montre une organisation en différents domaines identifiés chez les MSCRAMM<sub>S</sub> : un peptide signale constitué de 32 acides aminés suivie d'une région A avec 414 acides aminés hydrophobes, huit domaines Cna et à la fin une ancre qui comporte un motif **LPXTP** différents de ceux identifiés chez les autres adhésines (LPXTG), le LPXTP serait un nouveau motif d'ancrage. La protéine RadA pourrait jouer un rôle important dans la capacité de *R. gnavus* E1 à coloniser l'écosystème digestif. L'objectif du travail a été de réaliser des analyses bioinformatiques afin de déterminer la structure de la protéine RadA et contrôler les conditions d'expression du gène *radA*.

## Materiel Et Methode

La souche E1 a été isolée du microbiote fécale dominant d'un homme adulte sain (Ramer *et al.*, 1993) et identifié comme appartenant à l'espèce *Ruminococcus gnavus* E1 (Dabard *et al.*, 2001). E1 est cultivée en anaérobiose, dans une chambre de type 'Freter' en atmosphère contrôlée (85% N<sub>2</sub>, 10% H<sub>2</sub>, 5% CO<sub>2</sub>). Elles sont incubées à 37°C dans du milieu BHI supplémenté en extrait de levure et hémine (BHI-YH).

## Extraction des ARNm

### In vitro

Les ARNm ont été extraits à partir d'une culture de *R. gnavus* E1 sur milieu liquide. Le culot a été traité avec les solutions suivantes : TES (Tris 10 mM, EDTA 1 mM, pH8, NaCl 0,1 M), 20 µg/µL de lysozyme. Des billes de zirconium soumises 5 min d'agitation à l'amalgamateur à billes mini bead-beater<sup>8</sup>® au Fast Preps ont été utilisées pour faciliter la lyse cellulaire. Les ARN ont été purifiés à l'aide du RNeasy Mini Kit (Qiagen) selon le protocole adéquat (RNeasy Mini Protocol for isolation of total RNA from bacteria).

### **In vivo**

Les ARN *in vivo* ont été préparés à partir de contenu caecal de rats monoxéniques inoculés avec *R. gnavus*.

### **Techniques de Gnotoxenie:**

Les rats Fisher sont issus de l'élevage axénique de l'UEPSD (INRA Jouy-en-Josas) et sont élevés dans des isolateurs de type Trexler (La Calhène). Ils sont nourris *ad libitum* par une nourriture commerciale stérilisée par irradiation gamma (40 kGy) et boivent de l'eau stérilisée par autoclave (20 min, 120°C). Les animaux sont ensuite inoculés par voie orale, à l'aide d'une sonde intra-gastrique, avec 0,5 mL de suspension bactérienne. L'extraction est réalisée sur des échantillons de 200 mg congelés (Dabard *et al.*, 2001) Le matériel ainsi obtenu est purifié avec le kit RNeasy Mini Kit® (QIAGEN).

### **Extraction de l'ADN génomique à partir de culture en milieu liquide**

Des cultures de 50mL ont été utilisées. Le culot cellulaire est lavé avec 20µg/µL de lysozyme et 200µL de TES (Tris 10 mM, EDTA 1 mM, pH8, NaCl 0,1 M). La lyse est effectuée par l'ajout d'une solution de SDS à 20%. L'ajout de phénol, agent dénaturant des protéines, permet une séparation des phases. La phase aqueuse est récupérée puis traitée avec un mélange phénol/chloroforme/alcool isoamylique (25/24/1), puis un mélange chloroforme/alcool isoamylique (24/1). L'ADN est précipité à froid (-20°C) par ajout d'éthanol absolu. Le précipité est récupéré par centrifugation. Le culot d'ADN est ensuite lavé avec de l'éthanol à 70%. Le culot est repris dans du tampon TE (Tris 10 mM, EDTA 1 mM, pH8) contenant de la RNase (20µg/mL). La concentration et la pureté de l'ADN sont estimées par spectrophotométrie (BioPhotometer® Eppendorf, Hamburg, ALLEMAGNE) et sa qualité par électrophorèse sur gel d'agarose dans du TBE 0,5X (Tris 50 mM, acide borique 45 mM, EDTA 0,5 mM).

### **PCR**

Les PCR ont été réalisés sur l'ADN génomique purifié ou le produit RT. 2µM d'amorces spécifiques : C1L-C1R, C2L-C2R, C3L-C3R, C4L-C4R, C5L-C5R, C6L-C6R, C7L-C7R, C8L-C8R (tableau 1), 2.5mM de dNTP, 5U de l'Ampli Taq Gold<sup>R</sup> (Applied Biosystems) dans le tampon adéquat. Les réactions de PCR ont été effectuées dans un thermocycleur Master cycler personal (Eppendorf) dans les conditions suivantes : Une étape d'activation de l'Ampli Taq (10min à 94°C), une étape de dénaturation (30 sec à 94°C), une étape d'hybridation (30 sec) des amorces avec la matrice. Les températures d'hybridations différentes selon la nature des amorces, une étape d'élongation à 72°C dont le temps diffère selon la taille du fragment à

amplifier (1 min par 1kb à amplifier). Ces cycles ont été suivis d'une étape de poste élongation (10 min à 72°C).

**Tableau 1** : Amorces utilisées pour les PCR

Amorces	Séquences (5'-3')
C1L	5'-GGTTTTGAATGTCCCCAAA-3'.
C1R	5'-GCAGACAGATTCCCGAAAAA-3'
C2L	5'-GGGAAGAAGATGCAGAACCC-3'
C2R	5'-AGTAGGCGGCATCCAGTA-3'
C3L	5'-TGAAGAGCCAAGTGCAAAG-3'
C3R	5'-ATTTCCCCCAGACTGCTT-3'
C4L	5'-GGTGTCTGTGGAAGTGCTGA-3'
C4R	5'-AAGTTCCGGGGAACACAAAT-3'
C5L	5'-CATGAGGAACAGGCTCCAAT-3'
C5R	5'-TTTTCTGCATCCCAAAGTCC-3'
C6L	5'-GAGTTTGTAAAGAATGTGAGGATAAGAGAG-3'
C6R	5'-GTAATCCCTTTAAAATCAGAAGCCG-3'
C7L	5'-AGTTATGTGGACGCGAAAACCAA-3'
C7R	5'-TCTGTCATGCCGCCAAGTCTTA-3'
C8L	5'-AATCGGCTTCCATATGCGGT-3'
C8R	5'-CTGCCGTCATCGCTCTTAA

## RT-PCR

La RT-PCR est réalisée avec 30µg/µl de Random Primer. Le mélange est incubé à 70°C pendant 5 minutes, puis ajouté à un mélange qui contient : 100U de RNase inhibitor, avec son tampon adéquat (5X Buffer), et 2.5mM de dNTP ; le mélange réactionnel est incubé à 37°C pendant 5 minutes. La Transcriptase Réverse (10 000U) est ajoutée, puis le mélange est incubé pendant une heure à 42°C afin de synthétiser les ADNc. La PCR est ensuite effectuée sur l'ADNc. Les résultats de la PCR ainsi que la RT-PCR ont été analysés sur gel d'agarose. Les fragments d'ADN sont soumis à un courant continu 100V dans une cuve électrophorétique (Mupid-ex-Advance Eurogentec).

## Analyse bioinformatique

Recherche de cadres ouverts de lecture sur les séquences nucléotidiques a été réalisée à l'aide du programme ORF Finder: <http://www.ncbi.nlm.nih.gov/gorf/gorf.html>

- Les homologies des séquences peptidiques et nucléotidiques ont été réalisées avec les séquences présentes: <http://www.ncbi.nlm.nih.gov/BLAST/>
- Les amorces pour les différentes PCR ont été choisies grâce au programme : [http://biotools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotools.umassmed.edu/bioapps/primer3_www.cgi) ; et vérifiées par simulation grâce au logiciel AmplifX téléchargeable à l'adresse suivante : <http://crn2m.univ-mrs.fr/AmplifX>

- Les alignements nucléotidiques et peptidiques: <http://www.ebi.ac.uk/Tools/clustalw2/>,  
<http://www.ebi.ac.uk/Tools/emboss/align/index.html>
- L'identification des hélices transmembranaires a été réalisée avec le programme TMHMM Serverv.2.0: <http://www.cbs.dtu.dk/services/TMHMM/>
- La recherche des peptides signaux a été effectuée avec le programme SignalP Server 3.0 disponible à l'adresse suivante: <http://www.cbs.dtu.dk/services/SignalP/>
- La prédiction des terminateurs de transcription grâce au logiciel RNAfold (<http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi>). Ce logiciel permet de déterminer les structures secondaires possibles de l'ARN tels que les structures tigeboucles caractéristiques d'un stop de transcription rho-indépendant. Cette méthode est basée sur le calcul de l'énergie associée à la formation et à la stabilité thermodynamique des structures secondaire. Plus l'énergie calculé est faible plus la structure est possible et plus la probabilité d'être en présence d'un terminateur est élevée.
- La recherche des structures particulière: <http://coot.embl-heidelberg.de/SMART/>
- L'identification des hélices transmembranaires a été réalisée avec le programme TMHMM Serverv.2.0: <http://www.cbs.dtu.dk/services/TMHMM/>
- La recherche des peptides signaux a été effectuée avec le programme SignalP Server 3.0 disponible à l'adresse suivante : <http://www.cbs.dtu.dk/services/SignalP/>

## Resultats Et Discussion

### Vérification de la séquence:

Une panoplie d'oligonucléotides spécifiques qui recouvrent l'ensemble de la séquence de *radA* a été choisie pour amplifier des fragments chevauchants (tableau 1). Après amplification du gène *radA* avec le couple d'amorce C1L-C7R, la séquence a été vérifiée par séquençage des deux brins. Un ORF de 5217 nucléotides a été confirmé, correspondant à une séquence protéique déduite de 1739 acides aminés.

Après confirmation de la séquence, une analyse bioinformatique a été réalisée. En réalisant des recherches dans des banques de donnée (NCBI Blastp), 20% d'identités et 48% d'homologies ont été observées entre *RadA* et CNA 'figure 2', la protéine modèle d'adhésion au collagène de *S. aureus*. 41% d'homologies et 24% d'identités avec des protéines putatives d'adhésion au collagène de *Bacillus cereus* (Q73CV3) et de *Clostridium perfringens* (Q93M90) ont aussi été identifiés 'figure 3, 4'

CNA -----MNKNVLFKFMVFI ML LNIITPLFNKNEAFAARDISSTNVTDLTVSPS 46  
 RadA KEQTNERERERLLRKTWKRCTAGFLAVLLGCSLFLVGTGSAMAAPLEQSENKAETQAEDI 60  
 : . . . : \*\* : . . . \* : \* \* : . . .

CNA KIEDGGKTTVKMFTD-----DKNKGIQNGD-----TIKVAWPTSGTV 83  
 RadA TIQQGEMFDPASDFKGITVKDGEKISFVLSADKEGKLFADRPPTYDCIYQVQKPSGETY 120  
 . : . : \* \* . \*\* : \* : . \* : \* \* : . . .

CNA KIEGYSKTVSLTVKGEQ-----VG 102  
 RadA EITRKIIIVKEKGEKESPRKDKKQKKEQKNGEEDAEPDGAHLDTSAKKEEGVLFVSPVS 180  
 \* . . . \* . . .

CNA QAVITPDGATITFNDKVEKLSDVSGFAEFVQG-----RNLQTNTSDDKVAITITSG 154  
 RadA SMEQAREKASLIKGDRIQYPSDLGSYSTCYFVYVNDRIAYCLESNLQSPSSDYVAEIIYES 240  
 : : \* : : . : : \* : . : . . . \* : . . . : : . .

CNA NKSTNVTVHKSEAGT----- 169  
 RadA NLNLQKVLVYGYGGPGLDTGEYLKNYSNDVKYVVLTHLAASYCYGGAEVAFVGTQDGLKR 300  
 \* . . . : . \* .

CNA ---SSVYYKTGDMLP-----EDTTHVRWFLNINNEKR----- 199  
 RadA YGVMVEYINYLCGQEAPPSAAISLSSKETAFLGAVQKTKNITLNGDHRNYITLPLPEGV 360  
 . : \* \* \* : \* : : : \* : : : \* : . . .

CNA -----YVSKDITIKDQIQGG----- 214  
 RadA TYHDTAGKEQKGGSIQIYGGTTFYFTAETVHGTWNSGDLGGQVGAQWKTLVSTGSGNQ 420  
 : : : \* : . . : . \*

CNA -----QQLDLSTLNINVTGTHSNYYSGPNAITDFEKAPFGSKITVD 255  
 RadA DVGYGDFYEEPSAKVSFSIQWMDISWIEVIKEDAKSSVKLAGAVFGIYRDPACTDLILEM 480  
 \* : \* \* : : . : . . . . : . . . \*

CNA NTKNTIDVTPQGYGSLNSFSINYKTKITNEQQKEFVNNSQAWYQEHGKEEVNGKA---- 311  
 RadA PPTDEKATKAE LTKTQDTIYIKEITAPKGYKLNTTAYNVNLEVAETQTVTVKNEEQKQK 540  
 : . . : : : . \* \* . . . : . \* : \* . . .

CNA --FNHTVHNINANAGIEGTVKGELKVLKQDKDKAPIANVKFKLSKKDGSVVKDNQKEIE 369  
 RadA IVVHKQGEKLTGVSGEENLQFLYTNTAFAGAKYKIYAAEDIYSQDKQTKIYQAGDLAAE 600  
 : : : . . . \* : \*

CNA IKTDANGIANIKALPSGDYILKEIEAPAPYTFDKDKEYPFTMKDTONQGYFTTIENAKEI 429  
 RadA LETKEDGSCSSDMLYLGTYKVVEQQAPDSLITIGKTEEERTHMVTLSYAG--QTVEVVEEE 658

CNA ETKKDVSQAQKWEVGTQKVKPTIYFKLYKQDDNQNTTPVDKAEIKKLEDGTTKVTWSNLPE 489  
 RadA TQYENARPKVSVSEVLKQSSNDDAALKGAIFGLYANEDITGADGSVLVTRGTLIQKAESGG 718

CNA NDKNGKTIKYLKVEVNAQKDDTTPPEGYTKKEDGLVVTNTEKPIETTSISGEKVVDDKDNQ 549  
 RadA NGKALFTADIPIGFHYAVKEIQAPSLYFKGNDSEYFFYKNDTTYTYTFTHTFQNKVEVR 778  
 : : : : \* : :

CNA -----DGKRPEKVSVNLLANGEKV--ETVDV 573  
 RadA GEVHIKIDKDTQDSVSQGDGLNGAVYGLYAAEDIQHPNGKTGLLYKKDQLVVGQGTIEK 838  
 \* : \* : : \* : : \* \* :

CNA TSETNWKYFKDLPKYDEGKKIEYTVTEDHVKDYTTDINGTTITNKYTPGETSATVTKNW 633  
 RadA GVLNFEDLYLGKYYVQEI SPPPSNAYLLDQTKYPVELAYEGQDVEIVQKNVTVVETIKKQ 898  
 . . . : .

CNA DDNNNQDGKRPTIEIKVELYQDG-----KATGKTATLNE SNNWHTW 674  
 RadA AFQLIKISDDGSQTETELLEAGFKVYLIRELSKVKDGS LKPSNGTEYTPQDFIGYDFSK 958

CNA TGLDEKAKGQVQKVTVEELTKVKGYTTHVDNNDMGNLIVTNKYTPPETTSISGEKVVDDKD 734  
 RadA EKTASYENGEKIQTEEMFTDKGYLCSPELPGKYVCIESTIPENVEGIQPLVTIDED 1018  
 : . : : : . \*\* : . : . . . : . \* . \* : :

CNA NQDG-----KRPEKVSVNLLANGEKVKTLDVTS----- 762  
 RadA SREPQVWRVFNDRPMQFYFKI IKKDAQTELPILKNSAHYKIYDVEKKKVKMKVRYPKPE 1078  
 . : : \* : : : \* \* . . . . .

CNA -----ETN-WKYEFKDLPKYDEGKKIEYTV 786

RadA	TIDVFETNEEGYLLTPEPLKMGTYRIEEVKSPEINFVQTGFQALKNGEFLPLNEVTAEG 1138
	. : : . : . : . : . : . : . : . :
CNA	TEDHVKDYTTDINGTTITNKYTPGETSATVTKNWDNNNQDGKRPTIEKVELYQDGKATV 846
RadA	TYEKAPRESITIKVDSNTAHEMEEETGKYIVVVEVKNDKDEAVGSLTIQKTGEMLVGAEKIT 1198
	: . : * : * : * : . : . : . : . : . :
CNA	KTATLNESNNWHTWTGLDEKAKGQVQVYKTYVE----- 878
RadA	DQILTKLKNGLAKAVNQVSTLFTGEEVMEETEKGYAFSYEEQGLAGAEFSIYARETIYSPD 1258
	-----ELTKVKGYTHV 890
CNA	-----GQMDSEGRIIRFEKDALVGKIVTDEKKGKGLNNLPIGKFYIKETKTKTGSFVLDPKEQDF 1318
	-----
CNA	DNNDMGNLIVTNKYTPETTSISGEKVWDDKDNQDGKRPEKVSNNLLANGEKVKTLVDVTSE 950
RadA	EITYQGVAVDVTKEIKNQRQKVEIEVLKKSEATKEPLEGVSFGLYAGEDIVNAAGNV 1378
	: : . * : . : . :
CNA	TNWKYEFKNLPKYDEGKKIEYTVTEDHVKDYYTTDINGTT-----ITNKYTPGETSAT 1002
RadA	VVKKDELVAVEKTDKEGKLYSDTIPHGKYLYLELEGLPGYLPEEKBIBIDASYPDKLE 1438
	: : : : * : * : * : . : . :
CNA	VTKNWDNNNQDGKRPTIEKVELYQDGKATGKTATLNESNNWHTWTGLDE----- 1053
RadA	VISIQKEVENQPTKVEITKTDITGEKEIEGAKLQILDAGNVVEEWTSAKESHLIYALKP 1498
	* . : . : * : . : . :
	-----
CNA	-----GKYILHEEQAPIENGYVKAEDVEFTVEETGEIQKVSMDKDHDKVEITKTDITGEQIEGA 1558
	-----
CNA	-----KAKGQVQVYKTYVEELTKVNGYTHVDNNDMGNLIVTNKYTPKPKPI 1100
RadA	KLQVLDEEGNIIEEWSTKPYRIEYLQPKTYVLHEEAPEGLIAEDVEFTVEETGEI 1618
	. : : : * : * : . : . :
CNA	YPEKPKDKTPP----- 1111
RadA	QKVSMDDEVPMQQLVIKKTDAEDQTPLANVEFELKNETEEVVGKLTDDKDGVAETSELLP 1678
	: . : . : . :
	-----
CNA	-----TKPDHSNKVKPTPPDKPSKVDKDDQPKDNKTKPE 1145
RadA	IATYKDGKVPAPITYVLSETKPLDGYEKSTETYEVTFSYVDAKTKVIEMVKEIQNKLPQ 1738
	. * . * . :
	-----
CNA	NPLKELPKTG--MKIITSWITWVFIGILGLYLILRKRFS-- 1183
RadA	TPEKTEEVKTDGQTMLLLPIGIAVVAILGISIVLWRIRRTKR 1780

**Figure 2 :** *Alignement des séquences peptidiques de l'adhésine CNA et Rad.( \* ) : identifiées,(.) : homologues*

<i>B. cereus</i>	-----MNFLRK <b>S</b> FNQ <b>I</b> KK <b>LSSSFIVVLLVCMNFLIHL</b> PK <b>EA</b> AATELTKLGLDVSYYNA 55
<i>C. perf.</i>	-----MSKAKR <b>F</b> KL <b>I</b> T <b>T</b> IT <b>L</b> I <b>F</b> T <b>L</b> F <b>L</b> T <b>N</b> IK <b>V</b> FAVEITSTDA <b>ES</b> LY <b>N</b> YD <b>S</b> P <b>T</b> W <b>G</b> K <b>V</b> 50
<i>R. gnavus</i>	-----LRKT <b>W</b> KR <b>C</b> T <b>A</b> <b>G</b> <b>F</b> L <b>A</b> V <b>L</b> L <b>G</b> C <b>S</b> S <b>L</b> F <b>V</b> T <b>G</b> S <b>A</b> MA <b>A</b> PL <b>E</b> Q <b>S</b> ENKAETQAEDI 60
Bc	IIFG-----DHSATSADIEGAMAVQKNMNASSYTVV <b>AA</b> T <b>G</b> ANNLAGATVWDE <b>G</b> YPS 107
Cp	LPIG-----NHRYY---VPGDLTTCYCLNTG----- <b>AL</b> N <b>P</b> T <b>G</b> Q <b>D</b> Y <b>T</b> K <b>E</b> M <b>Q</b> V <b>D</b> A <b>G</b> I <b>E</b> T 94
Rg	TIQQGEMFDPASDFKGITVKDGEKISFVLSADKEGKLF <b>DR</b> <b>P</b> G <b>T</b> Y <b>D</b> CIY <b>Q</b> K <b>P</b> S <b>G</b> E <b>T</b> Y 120
BC	LLLG <b>Q</b> F <b>T</b> K <b>A</b> G-TGQV <b>I</b> Q <b>D</b> GT <b>V</b> AM <b>T</b> K <b>D</b> G <b>D</b> P <b>E</b> GAMKSS <b>Y</b> D <b>R</b> I <b>S</b> Y <b>K</b> E <b>Q</b> A <b>E</b> IDA <b>K</b> F <b>K</b> E <b>F</b> R <b>K</b> 165
Cp	ILYW <b>G</b> Y <b>P</b> A <b>K</b> D <b>G</b> -SDWGISADEYRYCT <b>Q</b> LAIWAY <b>Q</b> K <b>E</b> A <b>G</b> L <b>S</b> R <b>G</b> L <b>V</b> R <b>E</b> R <b>L</b> Q <b>S</b> G <b>T</b> V <b>L</b> P <b>S</b> L <b>K</b> L <b>K</b> 152
Rg	EITRK <b>I</b> IV <b>K</b> E <b>K</b> G <b>K</b> E <b>G</b> E <b>S</b> P <b>R</b> K <b>D</b> K <b>K</b> Q <b>K</b> E <b>Q</b> K <b>N</b> G <b>E</b> EA <b>D</b> E <b>P</b> D <b>G</b> A <b>H</b> L <b>D</b> T <b>S</b> A <b>L</b> K <b>E</b> E <b>E</b> G <b>V</b> L <b>F</b> S <b>V</b> V <b>P</b> S 180
BC	DINS <b>V</b> I <b>E</b> D <b>A</b> G <b>L</b> H <b>T</b> D <b>K</b> PK <b>P</b> G <b>M</b> T <b>F</b> G-----IGED <b>V</b> K <b>N</b> S <b>N</b> I <b>Y</b> V <b>S</b> S <b>G</b> L <b>K</b> G 207
Cp	PVID <b>F</b> L <b>V</b> D <b>K</b> A <b>H</b> N <b>K</b> E <b>M</b> P <b>T</b> F <b>F</b> E <b>V</b> S <b>P</b> N-----D <b>I</b> A <b>H</b> Q <b>E</b> G <b>D</b> Y <b>F</b> S <b>E</b> P <b>I</b> K <b>I</b> 194
Rg	S <b>M</b> E <b>Q</b> A <b>R</b> E <b>K</b> A <b>S</b> L <b>I</b> K <b>G</b> D <b>R</b> I <b>Q</b> Y <b>P</b> S <b>D</b> L <b>G</b> S <b>Y</b> S <b>T</b> C <b>Y</b> F <b>V</b> N <b>D</b> R <b>I</b> A <b>Y</b> C <b>L</b> E <b>S</b> N <b>L</b> Q <b>S</b> P <b>P</b> S <b>S</b> D <b>Y</b> V <b>A</b> E <b>I</b> Y <b>E</b> S 240
BC	-----Q <b>E</b> P <b>F</b> N <b>V</b> K <b>D</b> V <b>L</b> P <b>N</b> V <b>N</b> K <b>D</b> F <b>I</b> V <b>I</b> H <b>S</b> D <b>A</b> E <b>E</b> V <b>N</b> F <b>G</b> G <b>A</b> I <b>L</b> Y <b>D</b> T <b>T</b> D <b>K</b> G <b>F</b> T <b>L</b> 255
Cp	-----K <b>S</b> N <b>Y</b> T <b>L</b> S <b>G</b> V <b>K</b> V <b>T</b> -IK <b>S</b> A <b>S</b> N <b>P</b> E <b>L</b> T <b>K</b> D <b>I</b> V <b>I</b> K <b>D</b> M <b>D</b> G-----N <b>V</b> K <b>D</b> S <b>G</b> Y <b>K</b> A 235
Rg	N <b>L</b> N <b>L</b> Q <b>V</b> L <b>V</b> Y <b>G</b> Y <b>G</b> P <b>G</b> D <b>L</b> T <b>G</b> E <b>Y</b> L <b>K</b> N <b>Y</b> S <b>N</b> D <b>V</b> K <b>Y</b> V <b>L</b> T <b>H</b> L <b>A</b> A <b>S</b> Y <b>C</b> Y <b>G</b> G <b>A</b> E <b>V</b> A <b>F</b> V <b>G</b> T <b>Q</b> D <b>G</b> L <b>K</b> R 300



BC	VNTSQAYDPNSFFTELASKVIWVFPNAKKITTKGYGVVGS-----	295
Cp	NESFRVYIPS-NAETGDLKVSVKAKVDIPAMLGMYMTEPQ-----	273
Rg	YGVMEYINYLCGQEAPPAAISLSSTKETAFLEGAVQKTKNITLNGDHRNYITLPLPEGV	360
BC	VFAPNAVVETKGGSSINGQA-----FVGGHLHQRDGFVHNF	330
Cp	GIQDMAVSSLDTHSMD-----KDN	293
Rg	TYHDTAGKEQKGGSIQIYGTTFYFTAETKTVHGTWNSGDLGGQVGAQWKTLLVSTGSGNQ	420
BC	KFNWPWKWKPAAEKGNLQIKKVDENDENIVLKDAKFDVIDK-----DNNVVAT	378
Cp	KVSWTGLNG-----AVQVIKGGDDGKLLTGAKFVLKNA-----NDENVAE	333
Rg	DVGYGDFYEEPSAKVVSFIQWMDISWIEVIKEDAKSSVKLAGAVFGIYRDPACTDLILEM	480
BC	VTTNEKGIAEVKDLPGFD-YFVKEISAPEGYIKVDTPVKVTIDNTNIEFVMKNTKKVEN	437
Cp	ATSQD-CKAVFNDIKPAE-YTIEVEAPQGYLVTN-PVNVTVKPNKVSIAEMTDTQIK-388	
Rg	PPTDEKGAATKAELTKQDTIYIKEITAPKGYKLNTTAYNVNLEVAETQTVTVKNEEQK-538	
BC	GQFKLLKDS-----ESQLLPGAKFDVIDKDGKVVET-----	470
Cp	GKIQVLKVD-----ETNTPLQGAEFETIQ-DGKHIEET-----	420
Rg	GKIVVHKQGEKLTGVSGEENLQFLYNTAFAGAKYKIYAAEDIYSQDKQTKIYQAGDLA	598
BC	--IVTDDKGEALSKQLPVGSYTLKEVEAPKGYELSSSS---VSDVEVN-KVTVDVVN	523
Cp	--ITTGENGIATSSLEPFGNYLVKEIKAPSKYVLNGEE---HPVTISENGKTIEIHTN	474
Rg	AELETKEDGSCSSDMLYLGTYKVVEQQAPDSLITIGKTEEERTHMVTLASYAGQTVVEVEE	658
BC	KKIPEKVTGQFEIVKVD AEDKTKVLSDAEFVYKDGK-KVDTLRTDKTGKVVQSQKLEPG-	581
Cp	KLIKGLVA---VKKTDSEISDLNLEGAFTIYDNNKNI VATITTNKDGAESEPLNYG-	529
Rg	TQYENARPKVSVEVLKQSSNDDAALKGAIFGLYANEDITGADGSLVLTGKGLTIQKAESGG	718
BC	-----KYTLKETKAPQGYKLLKEEIEVVVEAN-----KVVQVQVENAKEL	621
Cp	-----TYTMQETKAPKGYLLSNKVWDININEND-----KTYTFDVSNDVIK	570
Rg	NGKALFTADIPIGFHYAVKEIQAPSLYFKGNDSYEFFYKNDTTYTYTFTHTFQNKVEVR	778
BC	GSLQVTKKDAESG-KVLEG-----AEFRLKNENGQ-----VVGETKTN-KD	661
Cp	GKLIQIVKVDSENEEKVPEG-----AGFDVIAVNVNGIKEGTVVHDVTTD-KN	616
Rg	GEVHIKKIDKDTQDSVSGDGDNLNGAVYGLYAAEDIQHPNGKTGLLYKKDQLVVGQGTIEK	838
BC	GVVKFENLVPGKYTLEETKAPEG-YKAVEVTVEVNVVAN-----EVVKQEVVNEK	710
Cp	GFAYTKDLRYGDYKFHETDTPKG-YWKSDEKEYSFNIAENG-----KTYVKYIKNSP	666
Rg	GVLNFEDLYLGYVYVQEI SP PPSNAYLLDQTKYVELAYEGQDVEIVQKNVTVVETIKKQ	898
BC	LTGQFEIVKVD AEDKAKVLSDAEFVYKD-----GKK-	742
Cp	IQAQVRVIVKVD SKD-GKPLKGVKQIRNA-----DTKK	698
Rg	AFQLIKISDDGQSTETELELEGAGFKVYLIRELSKVKDGLKPSNGTEYTPQDFIGYDFSK	958
BC	-----VETLRTDKTGKVIS-QKLEPGTYTLKETKAPQGYKLLK-----	779
Cp	LVEFTNFIGIIPMKTTTLETNKNLGVTPQNLAYGNYLLEVEPLEGYIKVNPPIPFKIDE	758
Rg	EKTASYENGEKIQTEEMFTDKKGYLCS-PELPYKGYVCIESTIPENVEGIQPFVLTIDE	1017
BC	-----EEIEVVVEANKVVEVQVENAKELGSLQVIKKDAESGKVLGAEFKLKNETGQVVG	834
Cp	N----SVLEIEIKDLGTIYTQKVSNDRITANMELKLDKETNKPLENIEFKVTALDGMFKG	814
Rg	DSREPQVWRVFNDRMPQFYFKI IKKDAQTELPILKNSAHYKIYDVEKKKYVKMKVRYPKP	1077
BC	ETKTNTKD---GVVKFENLVPGKYTLEETKAPEGYKAVEVTVEVNVVAN-EVVKQEVVNM	889
Cp	KTNWLNKSDDK-GLVSLKGLYGDYRVDVKTTLNVLNKEPIFFSVKENGKTIKLOMTN	872
Rg	ETIDVFETNEEGYLLTPEPLKMGTYRIEEVKS PENFVQTGFQALKNGEEFLPLNEVTAE	1137
BC	---EKVTGQFEIKVDANDKAKVLSD---AEFTVYKDGKVAELKTD ESGKVMSPKLP	942
Cp	---KIRGSVELFKFDK-DTNRPLEG---VKFDLLNGDKKVGTYTTDNTGKITVNNLEA	924
Rg	GTYEKAPRESITIKVDSNTAHEME EETGKYIVVVEVKNDEAVGSLTIQKTGEMLVGAEKI	1197

```

BC      GEYTVKETK-----APEGYKLSNKEWKVTIQNEKEVVKVEAENERILGSLQIIKTDDKDQ 997
p      GNYTWVEVE-----AIDHYNKVDKKYDFNIYKDGQLEKIDVANTVKTGELDFSKT-DVTT 978
Rg     TDQILTTLKLNGLAKAVNQVSTLFTGEEVMETEKGYAFSYEEQGLAGAEFSIYARETIYSP 1257

BC      AKRLSGAEFTLTKDAQGNVVKEGITTDKSGIVKVDGLVPGEYTLLEETKAPEGYELTKQVIH 1057
Cp     GDSIDGAKVKKITGLEPQNKHINIEFTSSKEGNKFTLPEGKYTFEETLAPEGYRINKEVGT 1038
Rg     DGQMDSEGNRIIRFEKDALVGKIVTDEKKGKTLNLPIGKIFYIKETKTGTSTFVLDPKEQD 1317

BC      VTVDGE-----KIVDVKVTNSKSLGQFEIVKVD AEDKAKVLSDAEFVYKDGK----- 1105
Cp     FEIKDG-----QITKANLKDERKQGLIFTKTDVTTGKVI EGAKIKITCTEGLS----- 1087
Rg     FEITYQGQEVAVDYVTKEIKNQRQKVEIEVLK KSEATKEPLEGVSFGLYAGE DIVNAAGN 1377

BC      ---KVELTRTDKTGKV---ISQKLEPGTYTLKETKAPQGYKLLKEEIEVVVEANK--- 1154
Cp     ---KGKVIDFTSSKDG---NKFTLDEGKYTFEETSAPNGYRINKEVGTFEIKDGE--- 1136
Rg     VVVVKDELVAVEKTDKEGKLYSDTIPHGKYYLRELEGLPGYLPYEEKIEIDASYTDPKL 1437

BC      -VVQVQVENAKELGSLQVIKDAESGKVLLEGAEFKLKNETGQVVGETKTTNKDGVVKFEN 1213
Cp     -ITKANLKERKQGDLEFTKTDVTDGR IIEGAKIKIICVEGLSKGKIEFTSFKDGNKFT 1195
Rg     EVISIQKEVENQPTKVEITKTDITGKEIEGAKLQILD AEGNVVEEWTSAKES-HLIYA 1495

BC      LVPGKYTLEETKAPEGYKALEVTVEVNVVANTVIKQEV LNEKVKEKIKQVEITKVDATD 1273
Cp     LDEGKYTFEEISAPNGYEINKEVGTFEIKDGEITKANLKDERTTGVL EFTKTDVATGEVL 1255
Rg     LKPGKYILHEEQAPIENGYVKAEDVEFTVEETGEIQVSMKDDHTKVEITKTDITGQEI 1555

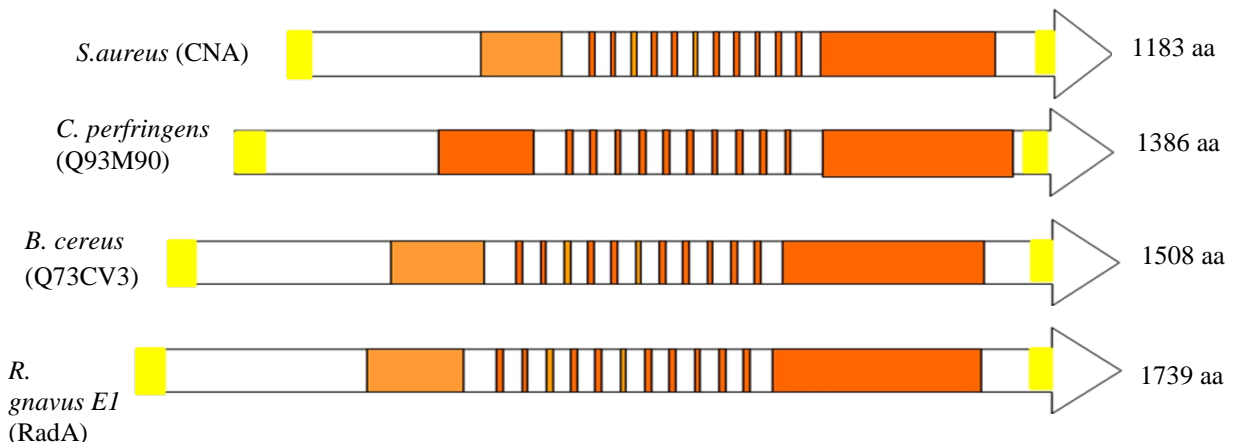
BC      TNKLAGAVFEILKDGTKIDTLTDDKNGKATSKELEPGDY IILKEVQAPEGYELSDKGI EF 1333
Cp     EGAKIKIECLEGLDQGVIEFTSSKEGNKFT---LAKGKYR ISETKAPEGYELTTETGEF 1312
Rg     EGAKLQVLDEE---GNIIEEWTSKPEYRIEYLQPGKTYV LHEEAAPGFLIAED-V EF 1610

BC      TISNQKIEVVKLQITNEKETS KGPENPGGETETPGGETETPGGETETPG EETEKPG EETE 1393
Cp     EITNQ-GDIITCNLTNKK----- 1329
Rg     TVEET-GEIQVSMKDEVPMG-----QLVIKKTDAEDQ TPLANVEFELKNKETEEVV 1661

BC      KPGEETEKPG EETGKPG EETEKPG EETGKPG EETEKPG EETGTSG EETEKPGGETGTPSE 1453
Cp     -----IEIVKTG----- 1336
Rg     GKLTDDKDG VATSELLPIATYKDGKPVAPITYVLSETKPLDGYEKSTETYEVTFSYVDAK 1721

BC      GMENVDK EKPTLPEKGGASHAQLPATGHD MNYL PFI GFALVLLGIRLRFMIKNS---- 1508
Cp     -----SRFDINSLIPLGILLVAGGIGGLFFTTKRKLS- 1368
Rg     TKVIEMVKEIQNKLPQTPEKTEEVKTG DQ TMLLLPIGIAVVAILGISIVLWRI RRTKR 1780
    
```

**Figure 3** : Alignement des séquences peptidiques des adhésines putatives de *Bacillus cereus* (Bc), *Clostridium perfringens* (Cp) et *R. gnavus* E1 (Rg). Les hélices transmembranaires sont surlignées en jaune, les identités sont représentées en rouge.



**Figure 4 :** Alignement schématique des quatre ORFs. Les parties en orange correspondent aux séquences présentant de fortes homologies. Les parties en jaune correspondent aux séquences codant pour des hélices transmembranaires

### Peptide signal:

L'analyse du domaine N-terminal des trois régions correspondant au produit du gène Q73CV3 de *B. cereus*, Q93M90 de *C. perfringens* et RadA de *R. gnavus E1* avec le logiciel TMHMM Server v. 2. 0

Des analyses statistiques ont révélé une très faible conservation de la structure primaire des PS. Seule la région qui entoure le site de clivage est conservée chez toutes les bactéries. Ces études ont permis d'établir la règle du «-1, -3» (Ramer *et al.*, 1993 ; Gunnar, 1984). Deux acides aminés sont conservés aux positions Ala-X-Ala. D'une façon plus générale, les précurseurs contenant soit Ala, Gly, Ser, Cys, ou Pro à la position -1 et Ala, Gly, Ser, Cys, Thr, Val, Ile, Leu ou Pro à la position -3, sont effectivement clivés par la SPase. Presque tous les résidus sont tolérés aux positions (-2) et (+1). En position (+2) les acides aminés sont généralement des résidus de petite taille.

La totalité de l'information nécessaire à l'activité de clivage de la SPase est contenue entre les résidus (-3) et (+2) du précurseur (Dev *et al.*, 1990). Ala est le résidu le plus fréquent, mais tous les autres sont tolérés, à l'exception de la Proline. La forme coude de la Pro interdit le clivage de PS, probablement par une insertion incorrecte dans le site catalytique. La règle du (-3,-1) stipule qu'une région (Ala-X-Ala) est conservée chez tous les PS<sub>S</sub> des bactéries à Gram positif (Gunnar, 1984). En effet, l'analyse bioinformatique avec le programme SignalP 3.0 Server du PS de RadA et de CNA confirme cette règle 'figure 5'



**Figure 5:** Mise en évidence du site de clivage chez RadA.

Le programme (Prédiction of transmembrane helices in proteins), révèle que les 30 résidus d’acides aminés observés seraient organisés en forme d’une hélice transmembranaire pouvant jouer le rôle du peptide signal. 31% d’identité ont été observés entre les trois régions. Cette hypothèse est confirmée par l’analyse réalisée avec le logiciel SignalP 3.0 Server qui révèle une répartition de charge similaire à celles des peptides signaux des bactéries à Gram positif.

Grâce au programme SMART (Simple Modular Architecture Research Tool), des domaines particuliers ont été découverts, nommés domaines Cna en référence à l’adhésine CNA de *S. aureus* (12). Ces domaines répétés dont le nombre varie d’une protéine à l’autre sont retrouvés chez toutes les adhésines de la famille des MSCRAMM<sub>5</sub> des bactéries à Gram positif. Chez RadA, huit domaines Cna ont été identifiés ‘figure 6’

**Séquences des régions particulières de *radA* selon le programme SMART :**

**Domaine cna B de la position 460 à 533:**

LAGAVFGIYRDPACTDLILEMPPTDEKGGATKAELTKTQDTIYIKEITAPKGYKLNTTY NVNLEVAETQTVTIVK

**Domaine cna B de la position 569 à 646:**

FAGAKYKIYAAEDIYSQDKQTKIYQAGDLAAELETKEGSCSSDMLYLGTYSKVVVE QQAPDSLITIGKTEEERTHMVTLT

**Domaine cna B de la position 817 à 885 :**

PNGKTGLLYKKDQLVVQGTIEKGVLFNFDLYLGKYYVQEISPPPSNAYLLDQTKYP VELAYEGQDVEIV

**Domaine cna B de la position 1241 à 1332 :**

LAGAEFSIYARETIYSPDGQMDSEGNRIIRFEKDALVGKIVTDEKGGKTLNLPIGKF YIKETKTGTSFVLDPKQDFEITYQGQEVAVDYV

**Domaine cna B de la position 1358 à 1443 :**

LEGVSFGLYAGEDIVNAAGNVVVKDELVAVEKTDKEGKLYSDTIPHGKYYLRE LEGLPGYLPYEEKIEIDASYPKLEVISIQ

**Domaine cna B de la position 1466 à 1536 :**

IEGAKLQILDAEGNVVEEWTSAKESHLYALKPGKYILHEEQAPIENGYVKAEDVEF  
TVEETGEIQKVSMK

**Domaine cna B de la position 1555 à 1624 :**

IEGAKLQVLDEEGNIIEEWTSTKEPYRIEYLPQPKTYVLHEEAPEGFLIAEDVEFTV  
EETGEIQKVSMK

**Domaine cna B de la position 1645 à 1725 :**

LANVEFELKNKETEEVVGKLTDDKDGVAATSELLPIATYKDGKPVAPITYVLSETKPL  
DGYEKSTETYEVTFYSYVDAKTKVI

**Séquence transmembranaire du coté N-ter de la position 21 à 43:**

TAGFLAVLLGCSSLFVGTGSAMA

**Séquence transmembranaire du coté C-terminal de la position :**

MLLLPIGIAVVAILGISIVLWRI

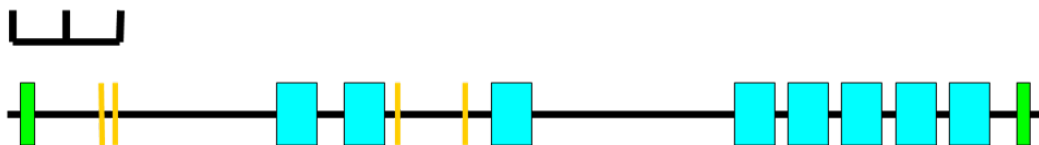
Parties répétées de la position 129 à 136: **KEKGKEGE**

Parties répétées de la position 140 à 149 : **KDKKQKKEQK**

Parties répétées de la position 652 à 658 : **VEVVEEE**

Parties répétées de la position 762 à 772 : **TTYTYTFHTF**

1 100 200 aa



**Figure 6 :** Représentation schématique de RadA et de ses domaines cna obtenus à partir du programme SMART.



L'analyse bioinformatique suggère très fortement que *radA* coderait pour une protéine d'adhésion de la famille des MSCRAMMs, qui pourrait jouer un rôle important dans la capacité de *R. gnavus* E1 à coloniser l'écosystème digestif. La masse moléculaire de la protéine RadA mature putative est de 198,9 KDa, son pHi est évalué à 4,7.

**C-terminal:**

Selon le programme TMHMM Server v. 2.0 (Prediction of transmembrane helices in proteins), la partie C-terminale se présente sous forme d'une hélice transmembranaire correspondant à la séquence qui jouerait un rôle dans l'ancrage de la protéine au niveau du peptidoglycane.

Chez les Gram positifs, toutes les ancres se caractérisent par la présence d'un motif très conservé «LPXTG», ce motif est clivé par la sortase, une enzyme qui permet l'ancrage de la protéine au niveau du peptidoglycane. Chez *S. aureus*, deux types de sortases ont été identifiées : la transpeptidase sortase A (SrtA) et la transpeptidase sortase B (SrtB) (Bentley *et al.*, 2007)

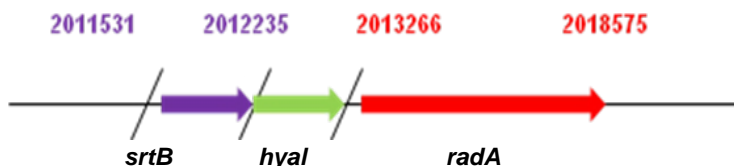
La différence entre les deux enzymes est que la SrtA reconnaît spécifiquement le motif LPXTG, et la SrtB, le motif NPQTN. Les deux sortases coupent après la thréonine, ce qui entraîne la formation d'une liaison amide avec le peptidoglycane. La collaboration avec le laboratoire de génomique comparative

(GENOSCOPE - Centre National de Séquençage - Evry) nous a permis d'avoir accès à la quasi-totalité du génome de *R. gnavus* E1 en septembre 2009, l'analyse a révélée la présence d'un gène codant pour une lyase 'hyal' (734pb) en aval de *radA*.

La lyase est l'une des principales protéines de surface identifiées chez le genre *Streptococcus*, sa présence a été également détectée chez le genre *Enterococcus*, *Staphylococcus* et *Streptomyces* (Rigden *et al.*, 2006). Elle a souvent été décrite comme un facteur de virulence (Berry *et al.*, 1994), son rôle principal est la dégradation des constituants de la MEC offrant une source d'énergie et de carbone aux bactéries. Elle dégrade principalement l'acide hyaluronique, largement réparti aux niveaux des tissus épithéliaux, le sulfate de chondroïtine présent dans les tissus conjonctifs. De ce fait, elle facilite la pénétration des bactéries dans les cellules hôtes (Rigden *et al.*, 2006). La Hyal est présente sous deux formes, la première forme qui est étroitement associée à la bactérie, et la deuxième est la forme libérée après action de la sortase (Ton-That *et al.*, 2000)

Un cadre ouvert de lecture codant pour une sortase 'srtB' (700pb) de *S. aureus* a été identifiée en aval de *radA* 'figure 7'. Afin de vérifier la présence d'une unité polycistronique, une analyse bioinformatique des séquences en utilisant le programme RNAfold a été réalisée.

Ce programme permet la détection des terminateurs de transcriptions. L'analyse révèle la présence de structure tige-boucle à la fin du gène lyase et du gène *srtB*. Il apparaît donc très probable que les trois gènes ne soient pas organisés en structure opéronique.



**Figure 7:** Représentation schématique de la disposition de *radA*, *srtB*, *hyal*

Chez RadA, nous avons retrouvé un motif **LPXTP** au lieu de LPXTG/NPQTN, sachant que la sortase coupe après le résidu T, cela laisserait penser que RadA disposerait peut être d'un nouveau motif d'ancrage 'figure 8'

KEQTNERRRERLLRKTWKKRCTAGFLAVLLGCCSLFVGTGSAMAAPLEQSENKAETQAEDITI  
 QQGEMFDPASDFKGITVKDGEKISFVLSADKEGKLFDAADRPGTYDCIYQVQKPSGETYEITRKI  
 IVKEKGKEGESPRKDKKQKKEQKNGEEDAEPDGAHLDTALKEEEGVLFVVPSSMEQAREK  
 ASLIKGDRIQYPSDLGSYSTCYFYVNDRIAICYLESNLQSPSSDYVAEIIYESNLNLQKVLYYGY  
 GPGDLTGEYLNKNSNDVKYVVLTHLAASYCYGGAEVAFVGTQDGLKRYGVMEYINYLCCG  
 QEAPPSAAISLSTKETAFLEGAVQTKNITLNGDHRNYITLPLPEGVYHDTAGKEQKGGSIQ  
 IYGGTTFYFTAECTVHGTWNSGDLGGQVGAQWKTLLVSTGSGNQDVGYGDFYEEPSAKVSF  
 SIQWMDISWIEVIKEDAKSSVLAGAVFGIYRDPACTDLILEMPPTDEKGATKAELTKTQDTIY  
 IKEITAPKGYKLNNTTAYNVNLEVAETQTVTVKNEEQK GKIVVHKQGEKLTGVS GEEGNLQFL  
 YTNFAFAGAKYKIYAAEDIYSQDKQTKIYQAGDLAAELETKEGSCSSDMLYLGTYKVVEQ  
 QAPDSL TIGKTEEERTHMVTL SYAGQTV E VVEEETQYENARPKVSVEVLKQSSNDDAALKGA  
 IFGLYANEDITGADGSLVTKGTLIQKAESGGNGKALFTADIPIGFHYAVKEIQAPSLYFKGND  
 SYEFFY EYKNDTTYTYTFHTHFQNK E V RGEVHIKKIDKDTQDSV SQDGD L N GAVYGLYAAE  
 DIQHPNGKTGLLYKKDQLV VQG TIEKGV LNFEDLYLGKYYVQEISPPSNAYLLDQTKYPVEL  
 AYE GQDVEIVQKNVT VVETIKKQAFQLIKISDDGSQTETELLE GAGFKVYLIRELSKVKDGS L  
 KPSNGTEYTPQDFIGYDFSKEKTASYENGEKIQT EEMFTDKKGYLCSPELPGYK YVCIESTIP  
 ENVEGIQPFLVTIDEDSREPQVWRVFNDRPMQFYFKIIKKDAQTELPIKNSAHYKIYDVEKKK  
 YVKMKVRYPKPETIDVFETNEEGYLLTPEPLKMGTYRIEEVKSPENFVQTGF EQALKN GEEFL  
 PLNEVTAEGTYEKAPRESITIKVDSNTAH E MEEETGKYIVVVEVKNDEAVGSLTIQKTGEMLV  
 GAEKITDQILTKLKNGLAKAVNQVSTLFTGEEVME TEKGYAFS YEEQGLAGAEFSIYARETIY  
 SPDGQMDSEGNRIIRFEKDALVGKIVTDEKGGKGT LNNLP IGFYIKETKTGTSTFVLDPKEQDFE  
 ITYQGGQEVAVDYVTKEIKNQ RQKVEIEVLKSEATKEPLEGV SFGLYAGEDIVNAAGNVVVK  
 KDELVAVEKTDKEGKLYSDTIPHGKY YLRELEGLPGYLPYEEKIEIDAS YTDPKLEVISIQKE  
 VENQPTKVEITKTDITGEKEIEGAKLQILD AEGNVVEEWTS AKESHLYALKPGKYILHEEQAPI  
 ENGYVKAEDVEFTVEETGEIQK VSMKDDHTKVEITKTDITGEQEIEGAKLQVLDEEGNIEEW  
 TSTKEPYRIEYLQPGKTYVLHEEAAPEGFLIAEDVEFTVEETGEIQK VSMKDEVPMGQLVIKKT  
 DAEDQTPLANVEFELKNKETE EVV GKLTTDKDGVATSELLPIATYKDGKPVAPITYVLSETKP  
 LDGYEKSTETYEVTFSYVDAKTKVIEMVKEIQNKK **LPQTPE**KTEEVKTDGQTM LLLPIGI AVV  
 AILGISIVLWRIRRTKR

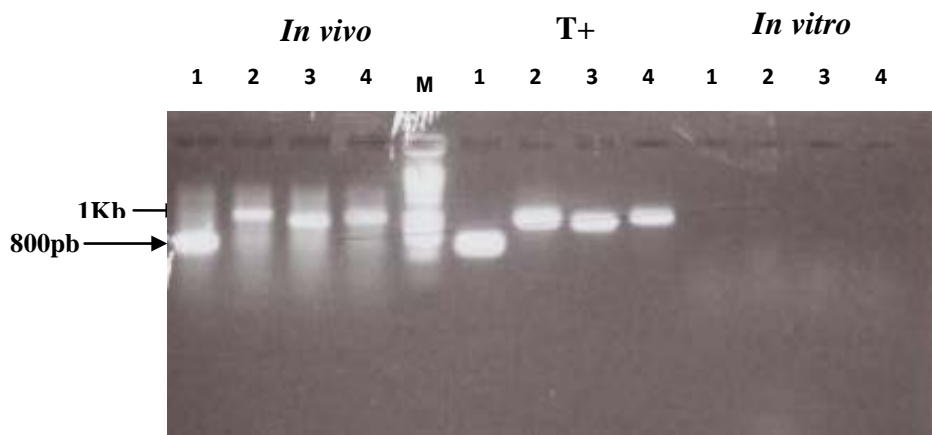
**Figure 8 :** Séquence peptidique de RadA

L'analyse bioinformatique suggère très fortement que *radA* coderait pour une protéine d'adhésion de la famille des MSCRAMM<sub>5</sub>, qui pourrait jouer un rôle important dans la capacité de *R. gnavus* E1 à coloniser l'écosystème digestif. La masse moléculaire de la protéine RadA mature putative est de 198,9 KDa, son pHi est évalué à 4,7.

### Détermination des conditions d'expression du gène *radA*

La RT a été réalisée avec des Random Primer (Invitrogen), suivie d'une PCR avec quatre couple d'amorces spécifiques de *radA*. Un fragment de 800pb a été observé après amplification avec C1L-C1R, et un fragment de 1Kb pour chaque couple d'amorce : (C3L-C3R) (C5L-C5R) (C7L-C7R) 'figure 9' Les résultats montrent que le gène *radA* est exprimé quand la bactérie se développe dans des conditions *in vivo*, en revanche il ne semble

pas être exprimé *in vitro* sur milieu BHY-YH lorsque E1 est en phase exponentielle



**Figure 9 :** Profil électrophorétique des produits RT-PCR. T+ : PCR sur l'ADN génomique de *R. gnavus* E1 ; 1(C1L-C1R), 2 (C3L-C3R), 3(C5L-C5R), 4(C7L-C7R).M : marqueur de taille moléculaire

### Conclusion

L'ensemble des résultats démontrent que le gène *radA* existe et qu'il est exprimé dans des conditions spécifiques du tube digestif (*in vivo*). La protéine RadA pourrait donc avoir un rôle très important dans l'intestin. Le fait que le gène *radA* se trouve à proximité des groupes de gènes impliqués dans la production de substances antimicrobienne (cluster *rumA*, *rumC*) qui ont un rôle dans la colonisation peut faire penser à une organisation en groupe de gènes impliqués dans une fonction spécifique. Ce genre de regroupement de gènes intervenant dans la virulence a été mis en évidence chez les bactéries pathogènes et appelé "îlots de pathogénécité". Le cluster *rum* associé à *radA* pourrait donc faire partie d'un "îlot de colonisation". La fonction des gènes environnants reste encore à étudier. Par analogie avec les adhésine de la famille des MSCRAMM<sub>5</sub> exprimée par des bactéries pathogènes à Gram positif, on peut penser que RadA serait une nouvelle protéine d'adhésion exprimée par une bactérie commensale et qui serait impliquée dans le pouvoir colonisateur de *R. gnavus* E1.

### References:

- Bentley, L. Matthew., Gaweska, Helena., M. Kielec, Joseph., McCafferty. G Dewey. Engineering the Substrat Specificity of *Staphylococcus aureus* Sortase A. J. Biol. Chem. 282, 6581-6571, 2007.
- Berry,A.M., Lock, R.A., Thomas, S.M., Rajan, D.P., Hansman, D and Paton, J.C. Cloning and nucleotide sequence of the *Streptococcus pneumoniae*



- hyaluronidase gene and purification of the enzyme from recombinant *Escherichia coli*, *Infect. Immun.* 62, 1101–1108, 1994.
- Charachon, S. Relation hôte-bactéries. MB7: Bactériologie, 2007.
- Crost, E. RumC peptides with antimicrobial activity, 2008.
- Dabard, J., Bridonneau, C., Philip., Anglade, P., Molle, D., Nadir, M., Ladire, M., Girardin, H., Marcille, F., Gomez, A. Ruminococcin A, a new lantibiotic produced by *Ruminococcus gnavus* strain isolated from human feces. *Appl Environ Microbiol.* 67, 4111-8, 2001.
- Dev, I.K ., Ray, P.H ., Novak, P. Minimum substrate sequence for signal peptidase I of *Escherichia coli*. *J. Biol. Chem.* 265, 20069 -72, 1990.
- Fons, M., Gomez, A., and Karjalainen, T. Mechanisms of colonization and colonization resistance of the digestive tract. *Microb. Ecol. Health Dis.* 12, 240–246, 2000.
- Gunnar, V.H. How signal sequences maintain cleavage specificity. *J. Mol. Biol.* 25, 173-243, 1984.
- Jett, BD, Gilmore, MS. Internalization of *Staphylococcus aureus* by human corneal epithelial cells: role of bacterial fibronectin-binding protein and host cell factors. *Infect Immun.* 70, 4697-700, 2002.
- Lannergård, J., Frykberg, L., and Guss, B. CNE, a collagen-binding protein of *Streptococcus equi*. *FEMS Microbiol. Lett.* 222, 69-74, 2003.
- Nallapareddy Sreedhar, R., Weinstock George, M., and Murray Barbara, E. Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Mol. Microbiol.* 47, 1733-1747, 2003.
- Ramer, F., Nicoli, J., Dabard, J., Corrig, T., Laddire, M., Gueugneau, A.M. and Raibaud, P. Trypsin-dependent production of an antibacterial substance by a human *Peptostreptococcus* strain in gnotobiotic rat *in vitro*. *Appl. Environ. Microbiol.* 59, 2876-83, 1993.
- Rigden, J. Daniel ., Littlejohn, E. James ., Joshi, V. Harshad ., Groot, L. Bert and Jedrzejewski, J. Mark J. Alternate Structural Conformations of *Streptococcus pneumoniae* Hyaluronan Lyase: Insights into Enzyme Flexibility and Underlying Molecular Mechanism of Action. *JBM.* 358, 1165-1178, 2006.
- Sillanpää, J., Nallapareddy, S.R., Prakash, V.P., Qin, X., Höök, M., Weinstock, G.M., Murray, B.E. Identification and phenotypic characterization of a second collagen adhesin, Scm, and genome-based identification and analysis of 13 other predicted MSCRAMMs, including four distinct pilus loci, in *Enterococcus faecium*. *Microbiol.* 154, 3199-211, 2008.
- Talay, S. R. Gram-positive Adhesins. *Contrib. Microbiol.* 12, 90-113, 2005.
- Ton-That, H., Mazmanian, S.K., Faull K.F. and Schneewind, O. Anchoring of surface proteins to the cell wall of *Staphylococcus aureus*. Sortase catalyzed

*in vitro* transpeptidation reaction using LPXTG peptide and NH(2)-Gly(3) substrates. J. Biol. Chem. 275, 9876–9881, 2000.

Zong, Y., Xu, Y., Liang X., Keene, D. R., Hook, A., Gurusiddappa, S., Hook, M., and Narayana Sthanam, V.L. A ‘Collagen Hug’ Model for *Staphylococcus aureus* CNA binding to collagen. EMBO J. 24, 4224–4236, 2005.

Xu, Y., Liang, X., Chen, Y., Koehler, T.M., Hook, M. (2004). Identification and Biochemical Characterization of Two Novel collagen Binding MSCRAMMS of *Bacillus anthracis*. J. Biol. Chem. 279, 51760-51768.