



HAL
open science

Escherichia Coli producteurs de colibactine et croissance tumorale, du mécanisme à la prévention.

Antony Cougnoux

► To cite this version:

Antony Cougnoux. Escherichia Coli producteurs de colibactine et croissance tumorale, du mécanisme à la prévention.. Médecine humaine et pathologie. Université d'Auvergne - Clermont-Ferrand I, 2013. Français. NNT : 2013CLF1MM01 . tel-01874964

HAL Id: tel-01874964

<https://theses.hal.science/tel-01874964>

Submitted on 15 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECOLE DOCTORALE DES SCIENCES DE LA VIE, SANTÉ,
AGRONOMIE, ENVIRONNEMENT
N° d'ordre :

Thèse

Présentée à l'Université d'Auvergne
Pour l'obtention du grade de

DOCTEUR D'UNIVERSITÉ

Spécialité : Microbiologie

Présentée et soutenue publiquement

par

Antony Cougnoux

janvier 2013

Escherichia coli producteurs de colibactine et croissance tumorale :
du mécanisme à la prévention.

Président :

Rapporteurs :

Pr. Erick DENAMUR, UFR Médecine Paris 7

Pr. Francis MÉGRAUD, UFR Médecine Bordeaux 2

Examineur :

Directeur de thèse : **Pr Richard BONNET**, UFR Médecine, Clermont 1

SOMMAIRE

I. Introduction bibliographique.....	8
1.L'écosystème intestinal.....	12
1.A. L'épithélium intestinal.....	12
1.A.a. Généralités.....	12
1.A.b. Les cellules différenciées de l'épithélium.....	13
Les Entérocytes.....	13
Les cellules sécrétrices.....	13
Les cellules M.....	14
1.A.c. Intestin grêle versus côlon.....	14
1.A.d. Renouvellement de l'épithélium.....	15
1.B. Le microbiote intestinal.....	16
1.B.a. Généralités.....	16
1.B.b. La phylogénie du microbiote intestinal.....	17
1.B.c. La dynamique du microbiote intestinal.....	17
1.B.d. Les fonctions du microbiote intestinal.....	19
1.B.e. Microbiote intestinal et pathologies associées.....	20
1.C. Le système immunitaire intestinal.....	21
1.C.a. Les cellules M.....	21
1.C.b. Le « homing » des lymphocytes.....	22
1.C.c. La tolérance du microbiote.....	22
2. Processus de maintenance de l'intégrité du génome.....	24
2.A. Le cycle cellulaire.....	24
2.A.a. Les principales étapes.....	24
2.A.b. La régulation du cycle cellulaire.....	24
2.A.c. Les points de contrôle.....	25
2.B. Les systèmes de réparation de l'ADN.....	26
2.B.a. Les systèmes BER, NER et MMR.....	26
2.B.b. Les systèmes par recombinaison.....	27
2.C. La mort cellulaire par apoptose.....	28
2.C.a. Généralités.....	28
2.C.b. La voie extrinsèque.....	28
2.C.c. La voie intrinsèque ou mitochondriale.....	29
2.C.d. Les caractéristiques des cellules apoptotiques.....	29
2.D. La sénescence cellulaire.....	29
2.D.a. Les sénescences répliquatives et induites par le stress.....	29
2.D.b. Les voies de signalisation cellulaire de la sénescence cellulaire.....	30
La voie p53/p21Cip.....	30
La voie p16ink4A/Rb.....	31
Les autres voies de signalisation.....	31

2.D.c. La SUMOylation des protéines et sénescence.....	32
2.D.d. Les caractéristiques des cellules sénescentes.....	34
3 Le cancer colorectal.....	35
3.A. Généralités.....	35
3.B. Épidémiologie du cancer colorectal.....	36
3.B.a. Facteurs de risque du cancer colorectal.....	37
Les facteurs de risque endogènes.....	37
Les pathologies génétiques.....	37
Les pathologies coliques prédisposantes.....	38
Les facteurs exogènes.....	39
3.B.b. Développement du cancer colorectal.....	39
La carcinogenèse colorectale.....	40
Les voies de signalisation impliquées dans le cancer colorectal.....	41
Le gène suppresseur de tumeur p53.....	41
La voie Wnt/APC/ β -caténine.....	41
La voie TGF- β /SMAD.....	42
La voie RAS/MAPK.....	42
La voie PI3K/Akt.....	43
3.B.c. L'instabilité génomique.....	44
L'instabilité chromosomique.....	44
L'instabilité des microsatellites.....	44
3.B.d. Les altérations épigénétiques.....	46
3.B.e. Les oncomirs.....	47
3.B.f. Carcinogenèse et inflammation.....	48
Les étiologies de l'inflammation associée au cancer	49
NF- κ B et carcinogenèse.....	50
Il-6 et carcinogenèse.....	50
Carcinogenèse et chimiokines.....	51
COX-2 et carcinogenèse.....	51
Récepteurs Toll-like et carcinogenèse.....	51
TNF- α et carcinogenèse.....	52
Carcinogenèse et stress oxydant.....	52
3.B.g. Modèles de carcinogenèse.....	53
Le modèle de Vogelstein : la séquence adénome cancer.....	53
Autres modèles.....	53
4. Microbiote et cancer colorectal.....	55
4.A. Le microbiote du cancer colorectal.....	55
4.A.a. Microbiote intestinal, alimentation, et carcinogenèse.....	56
4.A.b. Microbiote intestinal, inflammation et carcinogenèse.....	57
4.B. Streptococcus gallolyticus gallolyticus et carcinogenèse.....	58
4.C. Enterococcus faecalis et carcinogenèse.....	59

4.D. Bacteroïdes fragilis et carcinogène.....	59
4.E. Escherichia coli et carcinogène.....	60
4.E.a. Cytotoxic Necrotizing factor	60
Support génétique et structures.....	61
Mécanisme d'action dans la cellule cible.....	61
4.E.b. Cytolethal distending toxin.....	62
Support génétique et structure.....	62
Mécanisme d'action dans la cellule cible.....	62
4.E.c. Cell cycle inhibiting factor.....	63
Support génétique et structure.....	63
Mécanismes d'action dans la cellule cible.....	64
4.E.d. Colibactine.....	64
Support génétique	64
Structure.....	65
Mécanisme d'action de la colibactine.....	66
Épidémiologie des bactéries productrices de colibactine.....	67
II- Présentation des résultats expérimentaux.....	68
1. Les Escherichia coli producteurs de colibactine induisent la sénescence cellulaire et stimulent la prolifération cellulaire.....	69
2. Exploration structurale et fonctionnelle de la protéine ClbP.....	100
3. Inhibition de l'activité génotoxique et protumorale de bactéries productrices de colibactine.....	104
III- Conclusions et perspectives.....	131
1. E. coli producteur de colibactine, sénescence prématurée et carcinogène colorectale.....	132
2. E. coli producteur de colibactine, sénescence prématurée et métastases	134
3. E. coli producteur de colibactine, sénescence prématurée et angiogénèse.....	135
4. E. coli producteur de colibactine, sénescence prématurée et signalisation cellulaire.....	136
5. E. coli producteurs de colibactine, sénescence et dysbiose intestinale.....	137
6. ClbP et la synthèse de la colibactine.....	138
7. Inhibition de ClbP et prévention des effets délétères de la colibactine.....	139
IV Références bibliographiques.....	141
V Annexe.....	165

Liste des abréviations

AA	: Acide Aminé
ADN	: Acide DésoxyriboNucléique
AIEC	: « Adherent and invasive <i>E. coli</i> »
AIP1	: « Apoptosis signal-regulating kinase 1- Interacting Protein 1 »
AIRC	: Agence Internationale de Recherche sur le Cancer
AKT	: « Protein Kinase B »
Ala	: Alanine
AMP	: Adénosine MonoPhosphate
AP-1	: « Activator Protein 1 »
APAF-1	: « Apoptotic Protease Activating Actor 1 »
Asn	: Asparagine
ARN	: Acide RiboNucléique
ATM	: « Ataxia telangiectasia mutated »
ATR	: « Ataxia telangiectasia Rad3-related »
β-Gal	: beta-Galactosidase
BAC	: « Bacterial Artificial Chromosome »
BAD	: « Bcl-2-associated death promoter »
BAX	: « Bcl-2-associated X protein »
Bcl-2	: « B-cell lymphoma 2 »
BER	: « Base Exision Repair »
BFT	: « <i>Bacteroides Fragilis</i> Toxin » ou Fragylinin
BID	: « BH3 Interacting-Domain death agonist »
BMP	: « Bone Morphogenetic Protein »
CAD	: « Caspase-Activated DNase »
CDC25	: « Cell Division Cycle protein 25 »
CDC42	: « Cell Division Control protein 42 homolog »
CDB	: cassure double brin
CDK	: « Cyclin Dependant Kinase »
CDT	: « Cytolethal Distending Toxin »
CHK	: « Checkpoint kinase »
Cif	: « Cycle Inhibiting Factor »
CNF	: « Cytotoxic Necrotizing Factor »
Cib	: Colibactine
CoA	: Coenzyme A
COX-2	: CyclOXYgénase 2
CSB	: Cassure Simple Brin
CU	: Colite Ulcéreuse
Cys	: Cystéine
Da	: Dalton
DAB	: acide 2-4 DiAminoButyrique
DAPI	: 4',6'-diamidino-2-phénylindole
DAXX	: « Death-associated protein 6 »
DMSO	: DiMethyl SulfOxide
DNA-PK	: « DNA protein kinase »
DR5	: « Tumor necrosis factor receptor superfamily, member 10b »
Dsh	: Dishevelled
<i>E. coli</i>	: <i>Escherichia coli</i>
ECAD	: <i>E. coli</i> à Adhésion Diffuse
ECAI	: <i>E. coli</i> Adhérents et Invasif
ECEAgg	: <i>E. coli</i> EntéroAggregatif
ECEH	: <i>E. coli</i> EnteroHémorragique
ECAP	: <i>E. coli</i> EnterPathogène
EGF	: « Epithelial Growth Factor »
EGFR	: « Epithelial Growth Factor Receptor »
ELK1	: « E twenty-six-like transcription factor 1 »
ERK	: « Extracellular signal-Regulated Kinase »
ExPEC	: « Extra-intestinal Pathogenic <i>E. coli</i> »

EZH2	: « Enhancer of Zeste Homolog 2 »
FADD	: « Fas-Associated Death Domain »
FAK	: « Focal Adhesion Kinase »
FAP	: « Familial Adenomatous Polyposis »
FGF	: « Fibroblast Growth Factor »
G1	: « Gap 1 »
G2	: « Gap 2 »
Gab1	: « Grb2-associated adaptor protein 1 »
GADD45	: « Growth Arrest and DNA Damage »
Gly	: Glycine
GM-CSF	: « Granulocyte-Macrophage Colony Stimulating Factor »
Grb2	: « Growth factor receptor-bound protein 2 »
GSK-3β	: « Glycogen Synthase Kinase 3 beta »
HIF1-α	: « Hypoxia-Inducible Factor-1 α »
HNPCC	: « Hereditary Non-Polyposis Colorectal Carcinoma »
HR	: « Homologous Recombination »
ICAD	: « Inhibitor of caspase-activated DNase »
IGF	: « Insulin-like Growth Factor »
IGF2R	: « Insulin-like Growth Factor 2 Receptor »
IHH	: « Indian Hedgehog »
IL-	: InterLeukine
IL23R	: Récepteur à l'InterLeukine 23
JNK	: « c-Jun N-terminal kinases »
kb	: Kilo paires de base
kDa	: Kilo Dalton
LEE	: « Locus of Enterocyte Effacement »
Leu	: Leucine
LEF	: Lymphoid enhancer-binding factor
LLO	: LystérolLysine O
LPS	: LipoPolySaccharide
Lys	: Lysine
MAPK	: « Mitogen-Activated Protein Kinase »
MDM2	: « Murine Double Minute 2 »
MEKs	: « Mitogen-Activated Protein Kinase Kinase »
MLH1	: « MutL homolog »
MMP	: Métalloprotéase
MMR	: « MisMatch Repair »
MRE11	: « Meiotic Recombination 11 Homolog »
MSI	: MicroSatellite Instable
MSS	: MicroSatellite Stable
mTOR	: « mammalian Target Of Rapamycin »
NER	: « Nucleotide Exision Repair »
NF-κB	: « Nuclear Factor kappa B »
NHEJ	: « Non-Homologous End Joining »
NOXA	: « Phorbol-12-myristate-13-acetate-induced protein 1 »
NRP	: « Non-Ribosomal Peptide »
NRPS	: « Non-Ribosomal Peptide Synthase »
PAK	: « p21-activated kinase »
PCR	: « Polymerase Chain Reaction »
PDB	: « Protein Data Bank »
PGE2	: prostaglandine E2
PI3K	: phospho-inositide-3-kinase
PIP2	: Phosphatidyl inositol diphosphate
PIP3	: Phosphatidyl inositol triphosphate
PK	: « PolyKetide »
PKC	: Protein Kinase C
PKS	: « PolyKetide Synthase »
PML	: « ProMyelocytic Leukemia »
PMT	: « <i>Pasteurella Multocida</i> Toxin »
PNK	: polynucléotide kinase
PPARγ	: « Peroxisome Proliferator-Activated Receptor gamma »

PSAS	: Phénotype Sécrétoire Associé à la Sénescence
PTEN	: « Phosphatase and tensin homolog »
PUMA	: « p53 upregulated modulator of apoptosis »
RAF	: Rapidly Accelerated Fibrosarcoma
RAS	: « Rat Sarcoma oncogene homolog »
Rb	: Retinoblastoma
RISC	: « RNA induced silencing complex »
ROS	: radicaux oxygénés
SDS-PAGE	: « Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis »
SENP	: « SUMO/Sentrin specific Peptidase »
SHH	: « Sonic Hedgehog »
SIS	: « Stress-Induced Senescence »
SMAD	: « Mothers Against Decapentaplegic homolog »
SO	: « Spermin Oxydase »
SPJ	: Syndrome de Polypose Juvenile
spp.	: sous-espèce
SST3	: système de sécrétion de type 3
STAT3	: « Signal Transducer and Activator of Transcription 3 »
SUMO	: « Small Ubiquitin-like MOdifier »
TBX	: « T-box gene »
TCF	: « Transcription factor »
TGF-β	: « Transforming Growth Factor beta »
Thr	: Thréonine
TLRs	: « Toll-Like receptors »
TNF-α	: « Tumor Necrosis Factor alpha »
UV	: UltraViolet
X-gal	: 5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
XPC	: « Xeroderma Pigmentosum, complementation group C »
ψ	: Acide aminé hydrophobe

Listes des figures

- Figure 1 : structure de la paroi de l'intestin humain en coupe transversale.
- Figure 2 : l'organisation de l'épithélium intestinal et colique au niveau des cryptes et des villosités.
- Figure 3 : la diversité du microbiote intestinal humain.
- Figure 4 : la diversité du microbiote intestinal humain en condition physiologique ou pathologique.
- Figure 5 : le cycle cellulaire eucaryote et de sa régulation par les cyclines et les kinases dépendantes des cyclines.
- Figure 6 : les systèmes de réparation de l'ADN.
- Figure 7 : le système de réparation des cassures double brin de l'ADN par recombinaison non homologue.
- Figure 8 : signalisation cellulaire associée à l'apoptose.
- Figure 9 : sénescence répllicative.
- Figure 10 : la signalisation associée à la sénescence cellulaire.
- Figure 11 : La machinerie de SUMOylation.
- Figure 12 : caractéristiques des cellules sénescents.
- Figure 13 : les modifications acquises au cours de la tumorigenèse.
- Figure 14 : les facteurs de risques et les différentes formes de cancer colorectal chez l'Homme.
- Figure 15 : la voie de signalisation WNT/APC/ β -caténine.
- Figure 16 : la voie de signalisation TGF β /SMAD.
- Figure 17 : la signalisation cellulaire associée à RAS.
- Figure 18 : la voie de signalisation PI3K/AKT.
- Figure 19 : les modèles de progression du cancer colorectal.
- Figure 20 : synthèse et mécanisme d'action des microARN.
- Figure 21 : la voie de signalisation COX-2.
- Figure 22 : les différentes séquences adénome / carcinome.
- Figure 23 : implication de *Streptococcus gallolyticus* spp. *gallolyticus* dans le cancer colotrectal.
- Figure 24 : implication de *Enterococcus faecalis* dans le cancer colorectal.
- Figure 25 : implication de *Bacteroides fragilis* toxigène dans le cancer colorectal.
- Figure 26 : mécanismes d'action de la cyclomoduline CNF.
- Figure 27 : mécanisme d'action de la cyclomoduline CDT.
- Figure 28 : le mécanisme d'action de la cyclomoduline Cif.
- Figure 29 : organisation génomique de l'îlot *pks*.
- Figure 30 : la cyclosporine A, un exemple de composé issu d'une mégasynthèse de type NRPS.
- Figure 31 : organisation des mégasynthèses d'une chaîne d'assemblage.
- Figure 32 : le mécanisme d'action de la cyclomoduline colibactine.
- Figure 33 : la voie de signalisation HGF/c-Met.
- Figure 34 : effet de la dose infectieuse sur la croissance tumorale dans un modèle de xénogreffes de cellules HCT116.
- Figure 35 : action de *E. coli* producteurs de colibactine sur les lignées cellulaires d'origine prostatique PC3 et LNCaP.
- Figure 36 : action de la colibactine sur l'invasion cellulaire.
- Figure 37 : tests d'adhésion sur cellules HCT116 sénescents ou normales.

I. Introduction bibliographique

Le cancer colorectal est le troisième cancer le plus fréquent chez l'homme et la femme [1]. Sa prévalence est en progression constante depuis 1975 et représentait 1,2 million de cas en 2008. Avec environ 609.000 décès imputés par an, c'est également la quatrième cause de décès par cancer. Les formes familiales du cancer colorectal ont fait l'objet de nombreuses études et les gènes associés sont bien connus [2–4]. Cependant, les cancers colorectaux sporadiques représentent 95% des cas. Aussi, on considère que 75 à 80% des causes de cancer colorectal sont liées à des facteurs environnementaux, alors que les facteurs génétiques familiaux représentent seulement 15 à 20% de la causalité [5,6].

Le microbiote intestinal, qui regroupe l'ensemble des microorganismes présents dans l'intestin, est un élément essentiel de l'environnement colique et les interactions entre microbiote et cancer colorectal font l'objet d'un intérêt croissant [7,8]. Bien que la plupart des interactions entre le microbiote et l'intestin sont symbiotiques ou commensales, certaines bactéries du microbiote, appelées pathobiontes, ont le potentiel pour favoriser le développement de maladies et notamment le cancer colorectal [9,10]. Cependant, le délai entre l'initiation du cancer et les premiers symptômes de la maladie, la complexité de la flore bactérienne et les variations entre souches bactériennes entravent les tentatives visant à identifier un lien entre les bactéries du côlon et le cancer colorectal. Aussi, les données cliniques restent limitées et l'effet cancérigène de pathobiontes a été étudié pour seulement quelques bactéries, principalement *Bacteroides fragilis*, *Enterococcus faecalis* et très récemment *Escherichia coli* [11–14].

Découverte en 1885 par Théodore Escherich dans des selles de nourrissons, *E. coli* fait partie du microbiote intestinal. Il colonise le tractus gastro-intestinal des nourrissons dans les heures qui suivent la naissance, puis établit une relation symbiotique qui perdure pendant des décennies dans la couche muqueuse du côlon [15]. Compétiteur très performant sur ce site surpeuplé, *E. coli* est l'anaérobie facultatif le plus abondant du microbiote intestinal humain. Suite à l'acquisition d'attributs de virulence, *E. coli* peut se comporter comme un pathobionte [15,16]. Ainsi, le génome de *E. coli* comprend le core-génome constitué d'environ 2000 gènes essentiels aux fonctions cellulaires de base et un ensemble génique flexible. Cet ensemble variable d'une souche à une autre est composé de 2068 à 3379 gènes par souche, issus d'un pool d'environ 15800 gènes codant des fonctions optionnelles souvent impliquées dans la pathogenèse [17].

Les souches de *E. coli* pathogènes sont responsables d'infections intestinales ou extra-intestinales. Elles sont classées en pathotypes selon les manifestations cliniques engendrées, les facteurs de virulence hébergés et la nature des interactions avec les cellules hôtes [15]. Parmi les agents pathogènes intestinaux responsables de diarrhées, six catégories sont bien définies : les *E. coli* entéropathogènes, les *E. coli* entérohémorragiques, les *E. coli* entérotoxigènes, les *E. coli* entéroagréatifs, les *E. coli* entéroinvasifs et les *E. coli* à adhésion diffuse. D'autres pathotypes intestinaux ont été décrits, notamment les *E. coli* adhérents et invasifs associés à la maladie de Crohn, qui se répliquent dans les cellules épithéliales, les macrophages et induisent la sécrétion de cytokines pro-inflammatoires [15,18,19]. Les souches de *E. coli* à l'origine de pathologies extra-intestinales (ExPEC) sont encore mal définies [16,20,21]. Elles ont acquis la capacité à dépasser les défenses immunitaires de leur hôte et à se propager dans l'organisme. Elles sont responsables de méningites, d'abcès, de péritonites, de septicémies et surtout d'infections du tractus urinaire [22,23]. Trois pathotypes d'ExPEC sont couramment individualisés : les *E. coli* uropathogènes responsables d'infections urinaires, les *E. coli* associées aux méningites néonatales et des *E. coli* responsables d'infections systémiques en élevage aviaire [16,22].

Parmi les facteurs de virulence les plus importants de *E. coli*, on retrouve des adhésines, des invasines, des sidérophores, le flagelle, les antigènes O, la capsule et des toxines [21–24]. Ces facteurs contribuent à la colonisation et à la pathogénèse en bloquant ou détournant des processus cellulaires normaux, comme l'assemblage du cytosquelette ou l'apoptose [15,22]. Récemment, le groupe du Pr. Eric Oswald a défini un nouveau type d'effecteurs bactériens activant ou inhibant le cycle cellulaire eucaryote. Ces toxines, appelées cyclomodulines, sont identifiées dans un nombre croissant de bactéries [7,25]. Leur activité sur le cycle cellulaire des cellules hôtes suggère un rôle dans l'initiation et/ou la progression de cancers. Quatre cyclomodulines ont été identifiées chez *E. coli* ; Cytotoxic Necrotizing Factor (CNF), Cytolethal Distending Toxin (CDT), Cell cycle inhibiting factor (Cif) et une toxine récemment décrite, la colibactine [26].

La structure de la colibactine est inconnue à ce jour. Cependant, l'îlot génomique responsable de sa production, appelé *pks* code des synthétases de peptides non ribosomiaux (« Non-Ribosomal Peptide Synthases » ou NRPS), des synthétases de polykétides (« PolyKetide Synthases » ou PKS) et des enzymes hybrides PKS-NRPS. Ces enzymes sont connues pour être codées par des blocs de gènes et pour constituer des chaînes métaboliques

synthétisant des composés chimiques de type polykétide-peptide non ribosomal (« PolyKetide-Non-Ribosomal Peptide » ou PK-NRP), souvent toxiques et parfois utilisés en thérapeutique [27]. Les souches de *E. coli* porteuses de l'îlot *pks* induisent des cassures double brin de l'ADN *in vitro* et *in vivo* [14,26], des mutations et une instabilité chromosomique [14]; des lésions fréquemment observées dans le cancer colorectal [14,26]. De plus, Arthur et collaborateurs ont récemment montré que les souches de *E. coli* porteuses de l'îlot *pks* sont fréquemment observées au contact des tumeurs malignes coliques, et que ces souches avaient une activité pro-tumorale dans un modèle murin prédisposé [13].

L'objectif de ce travail était de mieux comprendre les relations entre les bactéries porteuses de l'îlot *pks* et le cancer colorectal ainsi que d'explorer sur le plan structural et fonctionnel la protéine ClbP afin de développer des inhibiteurs permettant de bloquer l'activité génotoxique et pro-tumorale des bactéries synthétisant la colibactine. Les résultats seront présentés sous forme d'articles précédés d'une revue bibliographique portant sur l'écosystème intestinal, les processus cellulaires et moléculaires contribuant au maintien de l'intégrité du génome, le cancer colorectal et avec le microbiote intestinal.

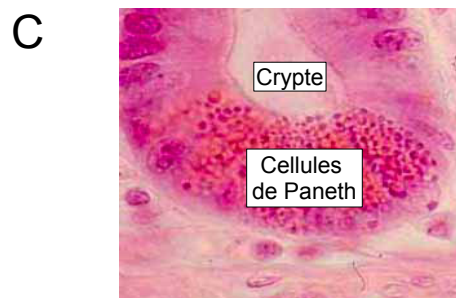
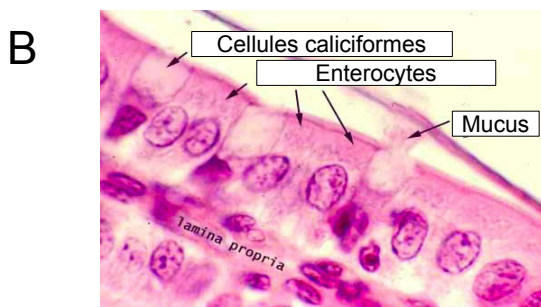
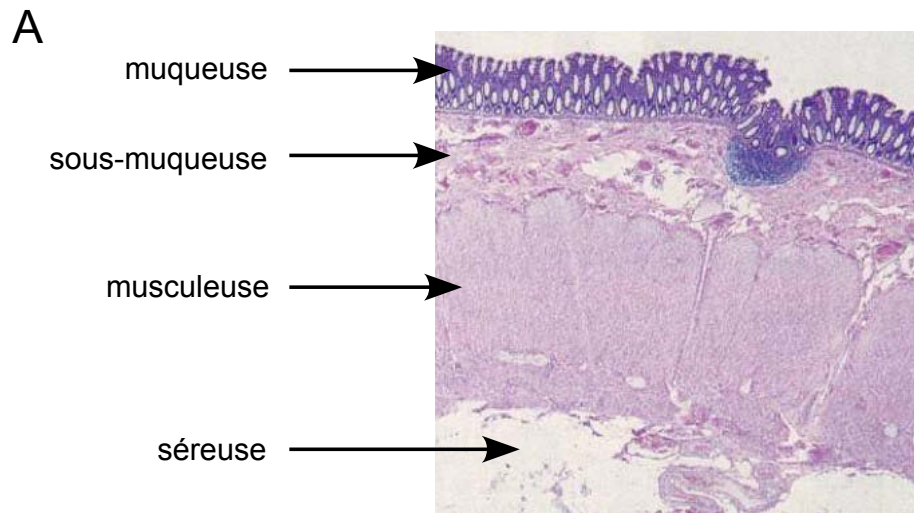


Figure 1 : structure de la paroi de l'intestin humain en coupe transversale. Vue globale de la paroi intestinale (A). L'épithélium de la muqueuse comprend différents types cellulaires (cellules caliciformes, les enterocytes, les cellules de Paneth situées au fond des cryptes (B et C). Les images proviennent de : <http://faculty.une.edu/com>

1. *L'écosystème intestinal*

Le tractus intestinal forme une interface entre le monde extérieur et le monde intérieur de 200 à 300 m² de superficie. Sa fonction physiologique principale est la digestion et l'assimilation des aliments tout en protégeant notre organisme contre des agents nocifs et des microorganismes pathogènes. C'est un écosystème complexe composé par l'alliance stable entre l'épithélium intestinal, le système immunitaire associé et la flore microbienne résidente. Des altérations génétiques ou fonctionnelles de l'un des trois composants de l'écosystème peuvent perturber cette alliance et, par conséquent, favoriser l'installation de pathologies [28].

1.A. L'épithélium intestinal

1.A.a. Généralités

L'architecture du tractus intestinal repose sur une organisation en cinq strates concentriques (**Figure 1**) :

- la muqueuse composée de l'épithélium, du chorion (ou lamina propria) sous-jacent, et de la *muscularis-mucosae* qui est une couche mince de tissu musculaire lisse,
- la sous-muqueuse, constituée de tissu conjonctif, du plexus nerveux de Meissner, de vaisseaux sanguins et lymphatiques,
- la musculature constituée de 2 couches de tissu musculaire lisse circulaire interne et longitudinal, séparées par le plexus nerveux d'Auerbach,
- et la séreuse, une tunique conjonctive externe.

Dans cette organisation, la muqueuse joue un rôle essentiel grâce à l'épithélium intestinal qui forme une barrière entre le milieu extérieur et le milieu intérieur. Il assure l'absorption des nutriments et participe à la réponse immunitaire. Le chorion agit en étroite collaboration avec l'épithélium. Il est riche en vaisseaux, qui assurent la récupération des nutriments liés à la fonction d'absorption. Il contient un tissu lymphoïde diffus et des follicules lymphoïdes pouvant être structurés en amas appelés plaque de Peyer. Dans la suite de cet exposé, nous nous intéresserons principalement à l'épithélium, qui est à l'origine des cancers colorectaux.

1.A.b. Les cellules différenciées de l'épithélium

Les fonctions de l'épithélium intestinal sont liées à la polarité structurale et fonctionnelle de cellules épithéliales différenciées que sont les entérocytes, les cellules sécrétrices, et les cellules M (« Microfold cells »).

➤ Les Entérocytes

Les entérocytes représentent 80% des cellules au niveau des villosités intestinales. Ils assurent une fonction de barrière mécanique grâce aux jonctions serrées qui les relient. Ces cellules épithéliales possèdent deux fonctions antinomiques : limiter l'entrée de grosses molécules ou de molécules indésirables tout en absorbant les petites molécules de nutriments grâce aux multiples enzymes et transporteurs membranaires. Elles présentent à leur surface apicale des microvillosités appelées bordure en brosse, qui augmentent la surface d'absorption. Les entérocytes produisent une série de cytokines pro-inflammatoires telles que l'interleukine 6 (IL-6) [29] , l'interleukine 8 (IL-8) [30] et le TNF- α (Tumor Necrosis Factor alpha) [31]. La sécrétion des cytokines est différente selon la présence de bactéries pathogènes ou commensales [30,32]. Ainsi les entérocytes signalent la présence de pathogènes invasifs aux cellules immunitaires et inflammatoires présentes dans la muqueuse intestinale *via* ces cytokines [33]. Les entérocytes synthétisent également des peptides antimicrobiens appelés β -défensines [34]. Ces réponses aux agressions microbiennes font partie du système immunitaire inné [35,36].

➤ Les cellules sécrétrices

Les cellules sécrétrices se composent des cellules entéroendocrines, des cellules caliciformes et des cellules de Paneth. Les cellules entéroendocrines font partie des types cellulaires qui prennent « la mesure » du contenu de la lumière intestinale avec le système nerveux et le système immunitaire. En réponse aux signaux intestinaux, elles sont capables de produire plus d'une vingtaine d'hormones différentes comme le glucagon, la gastrine, la β -endorphine ou encore la sécrétine [37]. Ces hormones sont déversées dans le sang pour agir sur des organes distants de l'intestin (effet endocrine). Elles sont aussi libérées dans la muqueuse intestinale pour agir sur l'intestin (effet paracrine).

Les cellules caliciformes sécrètent les glycoprotéines constitutives du mucus. Ce dernier

tapisse l'épithélium intestinal et joue un rôle cytoprotecteur vis-à-vis de l'acidité gastrique et des agressions bactériennes (entérotoxines). Il y a 2 types de mucus dans l'intestin : un gel insoluble, qui adhère fortement aux cellules et une couche visqueuse, qui est soluble dans l'eau et couvre le gel [38]. Ce mucus est composé de mucines, qui sont des glycoprotéines [39] constituant un large répertoire de sites de fixation pour les adhésines bactériennes. Le génome de l'hôte contrôle ainsi l'adhérence des bactéries [38]. Les cellules caliciformes sécrètent également des peptides qui stimulent des processus de réparation de l'épithélium intestinal [40].

Les cellules de Paneth produisent des molécules antimicrobiennes telles que les peptides de type α -défensine, le lysozyme ou la phospholipase A2 sécrétée [41]. Associées au mucus, elles contribuent à maintenir les cryptes stériles. Elles influencent également le microenvironnement de la crypte en stimulant la sécrétion d'ions chlore par les cellules épithéliales ou en modulant la fréquence des mitoses des cellules indifférenciées [41]. Les cellules de Paneth régulent aussi l'angiogenèse intestinale [42].

➤ **Les cellules M**

Les cellules M situées au regard des organes lymphoïdes (ex. plaques de Peyer), assurent un échantillonnage des antigènes endo-luminaux par endocytose puis les transfèrent aux cellules présentatrice d'antigènes, les cellules dendritiques et/ou macrophages situés au niveau du pôle basal, afin d'initier une réponse immunitaire ou un état de tolérance (cf. ci-dessous). Localisés entre les entérocytes, se trouvent aussi des lymphocytes T spécifiques de l'épithélium muqueux : les lymphocytes intra-épithéliaux, qui modulent la réponse immunitaire locale en sécrétant des cytokines [43].

1.A.c. Intestin grêle versus côlon

Contrairement à l'intestin grêle dont l'impératif est l'absorption des nutriments (il contient une majorité d'entérocytes, peu de cellules caliciformes à mucus, des villosités intestinales et des valvules conniventes pour augmenter la surface absorptive), la muqueuse colique est constituée principalement d'un épithélium de revêtement composé de cellules caliciformes sans villosités intestinales et de nombreuses cryptes. Il y a peu d'entérocytes et pas de cellule de Paneth. Le chorion est par contre riche en tissu lymphoïde (lymphocytes diffus et follicules lymphoïdes). La principale fonction du côlon est d'absorber l'eau présente dans les aliments.

1.A.d. Renouvellement de l'épithélium

Le maintien des fonctions de l'épithélium intestinal est associé à un renouvellement constant et rapide à partir de cellules progénitrices situées à la base des cryptes [44]. Au cours du renouvellement cellulaire, les cellules progénitrices se divisent, se différencient en entérocytes, cellules caliciformes ou cellules de Paneth. Elles montent progressivement à la surface des cryptes au fur et à mesure de leur différenciation selon un gradient de différenciation intra-cryptique jusqu'au sommet de l'épithélium où elles sont éliminées par apoptose. Ce processus assure un renouvellement en cinq jours de l'épithélium. Le maintien de l'épithélium intestinal est donc un processus dynamique qui nécessite un contrôle des processus de prolifération, de différenciation et d'apoptose dans l'espace et dans le temps. Ce processus complexe de renouvellement rapide de l'épithélium peut également expliquer la vulnérabilité de l'épithélium digestif au regard du risque de cancer.

La régulation du renouvellement de l'épithélium intestinal repose sur l'action coordonnée de plusieurs voies de signalisation cellulaire (**Figure 2**). Les voies Wnt/APC/ β -caténine et NOTCH sont impliquées dans la formation du tube digestif au cours du développement embryonnaire [45,46]. Dans l'organe mature, elles sont activées au niveau des cryptes selon un gradient décroissant partant de la base des cryptes [47,48]. Ainsi, elles assurent la prolifération des cellules dans la partie inférieure des cryptes du côlon. La superfamille du « Tumor Growth Factor beta » (TGF- β) comprend plusieurs membres dont les « Bone Morphogenic Proteins » (BMP), les activines et les TGF- β , qui sont des morphogènes importants au cours de l'embryogenèse. Dans l'organe mature, comme pour les voies WNT/APC/ β -caténine et NOTCH, le rôle de la voie TGF- β /BMP perdure. Par contre, elle s'exprime selon un gradient croissant en partant de la base des cryptes. L'expression est donc maximale au sommet des cryptes, dans les cellules différenciées, où elle contribue au blocage de la prolifération cellulaire [49] et contrôle l'apoptose nécessaire à l'élimination des cellules [50]. La voie Hedgehog émerge également comme un élément régulateur important du renouvellement épithélial. Tout comme les voies précédentes, elle joue un rôle important dans l'embryogenèse. Elle est également impliquée dans la prolifération et la survie cellulaire chez l'adulte [51]. L'activation de la voie Hedgehog dans l'intestin repose sur 2 protéines appelées SHH (« Sonic Hedgehog ») et IHH (« Indian Hedgehog »). Chez les mammifères adultes, SHH et IHH sont exprimées dans l'épithélium mais avec un profil d'expression différent. Le gène *SHH* est exprimé à la base des cryptes à très bas niveau (la protéine SHH est indétectable) et pour le côlon, SHH n'est observable que dans certaines cryptes [52]. Pour IHH, tant le transcrit que la protéine sont

exprimés au niveau des cellules différenciées du côlon au sommet des cryptes.

Au total, ces 5 voies de signalisation co-existent dans la crypte avec les voies WNT/APC/ β -caténine, NOTCH et SHH activées à la base des cryptes et associées à la prolifération des cellules alors que les voies TGF- β /BMP et IHH sont actives au sommet de la structure et sont impliquées dans la différenciation des cellules et leur élimination par apoptose. Des interactions croisées assurent la coordination de leurs actions ; la voie du TGF- β /BMP inhibe la voie WNT/APC/ β -caténine et inversement [45]. La voie IHH est capable de bloquer la voie WNT/APC/ β -caténine, mais active la voie TGF- β [53]. Une dérégulation de ces voies de signalisation est observée dans la carcinogenèse colorectale.

1.B. Le microbiote intestinal

1.B.a. Généralités

Notre organisme est composé d'un amalgame de cellules d'*Homo sapiens* et de microorganismes. Dès notre naissance nous sommes colonisés par des bactéries [54]. Puis, c'est au cours des premiers mois de notre vie que le microbiote s'installe [55]. Il représente l'ensemble des microorganismes peuplant notre organisme : bactéries, champignons, archae, virus et protistes. Cependant le terme microbiote est fréquemment utilisé pour définir la seule composante bactérienne de la flore. L'essentiel de notre microbiote se répartit tout du long de notre tractus digestif. Il est estimé à plus de 100 mille milliards de bactéries, soit dix fois plus que nos propres cellules humaines. Il atteint une densité maximale dans notre côlon distal avec 10^{11} bactéries pour un gramme de contenu [56,57]. Ainsi, l'estimation actuelle du nombre de gènes dans le génome humain est évaluée aux alentours de 23 000 gènes [58], tandis que pour le métagénome intestinal, l'ensemble combiné des génomes de notre microbiote intestinal, lui, s'établit à plus de neuf milliards [59].

Aussi bien sur le plan de l'abondance structurelle que génétique, le microbiote intestinal, anciennement appelé « flore intestinale », peut être considéré comme un organe à part entière tant son impact est important sur notre vie. En effet, sans lui, nous ne pourrions pas digérer certains composants de notre alimentation [60,61], notre système immunitaire serait immature [62] et la paroi de notre intestin serait faiblement développée. A mi-chemin entre le mutualisme et le symbiotisme, nous ne pourrions vivre l'un sans l'autre [61,63,64].

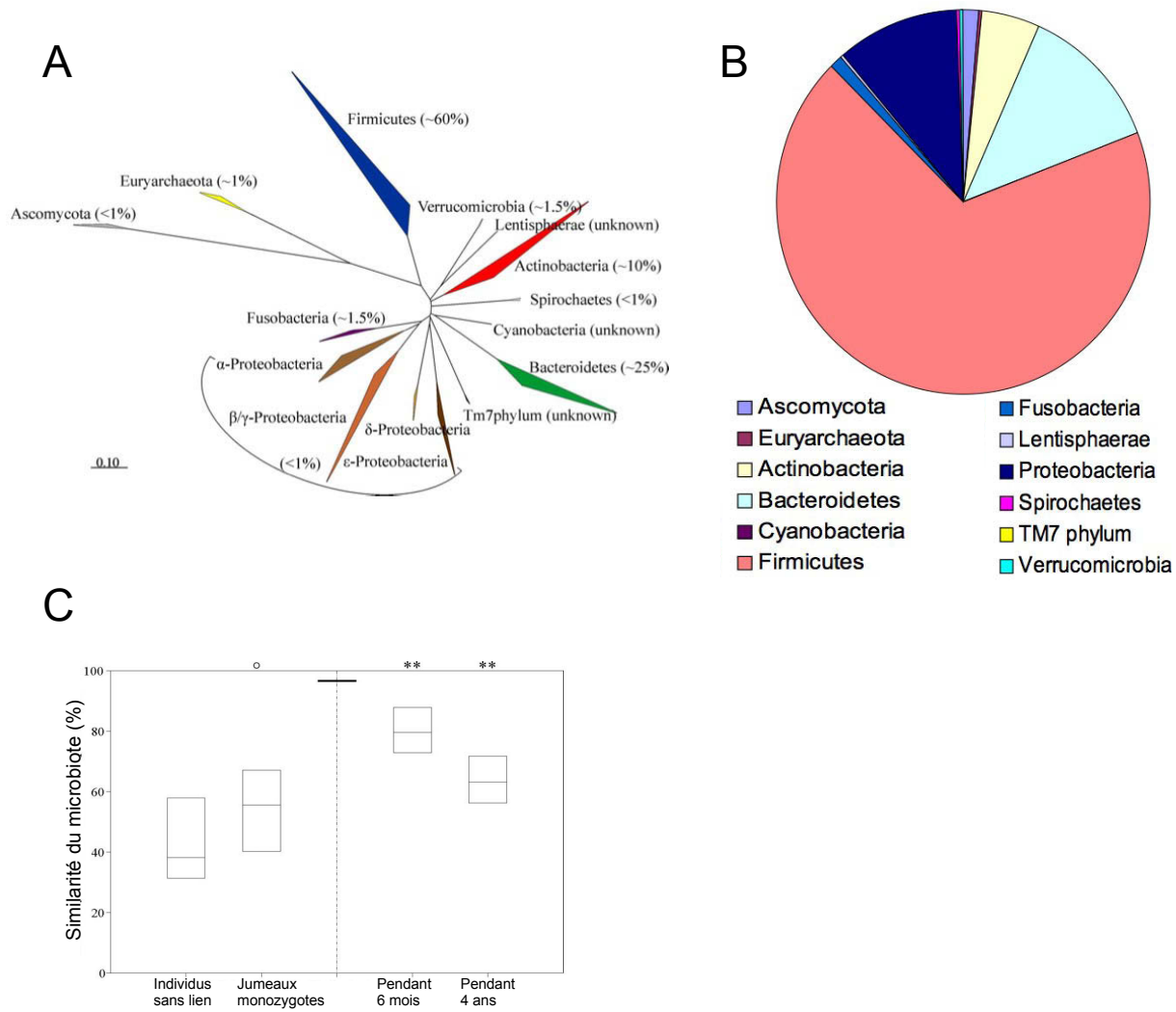


Figure 3 : la diversité du microbiote intestinal humain. Le microbiote intestinal comprend plusieurs milliers d'espèces bactériennes différentes réparties en 13 *phyla*, principalement composés des Firmicutes et des Bacteroidetes (A et B). La composition bactérienne de ce microbiote est variable d'un individu à un autre, mais est stable dans le temps chez l'adulte (C). D'après Rajilic-Stojanovic, [102].

1.B.b. La phylogénie du microbiote intestinal

Les études de la composante bactérienne du microbiote intestinal humain basées sur l'inventaire moléculaire du gène codant pour la sous-unité 16S de l'ARN ribosomal ont montré que plus de 70% des séquences étaient issues de bactéries non cultivées [65,66]. Près des deux tiers sont spécifiques à chaque individu. Cependant, plus de 95% des séquences sont assignées seulement aux Firmicutes, Bacteroidetes, Actinobacteria et Proteobacteria, et les deux premiers *phyla* se partagent la grande majorité de l'écosystème (**Figure 3**) [65,66].

Le phylum des Firmicutes (bactéries à Gram positif possédant des génomes faibles en G+C) représente en général 50 à 75% des espèces détectées. Plus de 95% des espèces des Firmicutes appartiennent à la classe des Clostridii [66]. Elles se répartissent en deux groupes Clostridiales ; le groupe Clostridiale XIV dit « *Clostridium coccoïdes* », qui comprend des espèces des genres *Eubacterium*, *Butyrovibrio*, *Roseburia*, *Dorea* et *Lachnospira*, et le groupe Clostridiales XV dit « *Clostridium leptum* », avec notamment les espèces *Faecalibacterium prausnitzii*, *Ruminococcus albus* et *Ruminococcus flavefaciens* [66]. Les études récentes utilisent un classement des Clostridii par familles, où les Lachnospiraceae, Clostridiaceae et Ruminococaceae dominent le phylum des Firmicutes.

Les Bacteroidetes représentent 10 à 40% du microbiote [65–67] et sont représentés par les espèces des genres *Bacteroides*, *Prevotella* et *Porphyromonas*. Le *phylum* des *Bacteroides* est très variable d'un individu à l'autre en termes d'abondance et de distribution des espèces [66]. Cependant, 30-35% des séquences assignées au *phylum* Bacteroidetes sont représentées par *Bacteroides vulgatus*. Même si on dénombre moins de *Bacteroides* que de Firmicutes, leur activité métabolique serait très importante et à l'origine des ressources énergétiques du tractus intestinal en assurant la production d'acides gras à chaîne courte [68].

Les *phylum* *Actinobacteria* et *Proteobacteria* sont moins systématiquement détectés et représentent en proportion de séquences, moins de 2% des bactéries totales [66,69]. Les *Actinobacteria* sont représentés par les bifidobactéries et les bactéries du groupe *Collinsella–Atopobium*, le *phylum* *Proteobacteria* par l'espèce *Escherichia coli* le plus souvent [66].

1.B.c. La dynamique du microbiote intestinal

A la naissance, l'implantation du microbiote s'effectue rapidement. La population «

source » est composée principalement de bactéries anaérobies facultatives comme des entérobactéries, des bifidobactéries et des lactobacilles [70]. Dès la diversification alimentaire, les deux principaux *phyla*, *Bacteroidetes* et *Firmicutes*, surpassent en nombre, ainsi qu'en diversité, les *Actinobacteria* et les *Proteobacteria* implantés précédemment. Cette mise en place s'accompagne de changements métaboliques. En effet, alors que les capacités fermentaires du microbiote conduisent à une production de lactate et d'acétate pendant les premiers mois de la vie, les concentrations de butyrate et de propionate deviennent dominantes et stables dès la deuxième année de la vie.

Les populations bactériennes dominantes restent alors relativement stables chez l'adulte sain [71]. L'exposition aux bactéries maternelles, l'environnement pendant les phases initiales de la vie et le patrimoine génétique de l'hôte définissent les caractéristiques microbiennes singulières propres à un individu. Cependant, malgré la grande variabilité d'espèces microbiennes interindividuelles, la représentation phénotypique des gènes incorporés dans le génome bactérien varie peu d'un individu à l'autre du fait d'une certaine redondance fonctionnelle des gènes [72]. Ainsi, un core génome du microbiote intestinal commun à tous les individus a été identifié, le microbiome central [73]. Cet ensemble de gènes code pour des fonctions métaboliques, dont le rôle est essentiel dans le maintien de la relation entre l'hôte et son microbiote. Chaque individu possède également un microbiome variable, qui apporte un ensemble de gènes différents selon les individus [74]. Il est à noter que des membres d'une même famille ont, le plus souvent, une flore bactérienne variable plus proche que des individus sans lien [75,76]. Cependant, le rôle exact du microbiome central et du microbiome variable dans la physiologie humaine est encore mal connu. Un effort de séquençage et des études fonctionnelles sont donc nécessaires afin de mieux comprendre leurs rôles précis [77].

Les effets du vieillissement sur le microbiote sont mal caractérisés. Cependant, les bifidobactéries déclinent chez les personnes âgées au profit des entérobactéries et des clostridii [78,79]. Parallèlement, la diversité des bifidobactéries décroît et se limite à deux espèces : *Bifidobacterium longum* et *Bifidobacterium adolescentis*. Cette chute de la richesse en espèce de bifidobactéries peut avoir des conséquences sur la santé des personnes âgées car elles sont impliquées dans le métabolisme (notamment du lactose) du microbiote et les interactions avec le système immunitaire en stimulant la production d'IgA. De plus, une baisse des *Bacteroides* a été également montrée et contribue à des changements significatifs du ratio Firmicutes/Bacteroidetes [80]. La chute des *Bacteroides* possédant des facultés à dégrader les

polyosides et à produire des acides gras à chaîne courte (AGCC), leur diminution peut avoir une influence sur la digestion et la capture d'énergie [81].

1.B.d. Les fonctions du microbiote intestinal

Le microbiote intestinal exerce de nombreuses fonctions physiologiques dont la plupart sont nécessaires et bénéfiques au maintien de la santé de l'hôte. Les bactéries commensales participent aux défenses (« barrière intestinale ») contre des bactéries pathogènes. Elles occupent des sites d'adhésion et entrent ainsi en compétition avec les pathogènes pour l'accès aux récepteurs de l'hôte. Elles consomment les nutriments disponibles limitant la source de nutriments pour les pathogènes. De plus, elles sécrètent des molécules antimicrobiennes telles que les bactériocines [82] au niveau de la surface épithéliale et agissent aussi sur le milieu (par exemple, l'acidification).

Un autre rôle de ce microbiote est la digestion de nutriments. La flore intestinale peut métaboliser de nombreuses substances d'origine exogène et endogène. La dégradation et la fermentation de ces substrats conduisent ensuite à la production de divers métabolites comme des acides gras à chaîne courte, des vitamines B (B1, B2, B6 et B12), de la vitamine K, etc... Ces productions fournissent aux bactéries l'énergie nécessaire à leur croissance, mais sont aussi, pour la plupart, absorbées et utilisées par l'organisme hôte. Le microbiote intestinal contribue à l'absorption par l'hôte de glucides et de lipides [61,83] et régule le stockage des graisses [84,85]. La présence du microbiote aboutit à une augmentation du stockage de triglycérides dans les adipocytes ainsi qu'à une activité lipoprotéine lipase plus élevée [86]. Le microbiote intervient également dans la détoxification de métabolites toxiques, de xénobiotiques et facilite le transit intestinal. Enfin, les bifidobactéries et les lactobacilles, ainsi que certains entérocoques, streptocoques et bactéroïdes, se distinguent par leurs effets bénéfiques sur la santé de l'hôte, telles que l'amélioration de la maturation, de l'intégrité de l'intestin et la modulation de la fonction immunitaire [32,87].

Les bactéries contribuent au développement de la muqueuse en facilitant la prolifération et la différenciation des cellules épithéliales. En absence de flore intestinale, la vitesse de renouvellement des cellules par les cellules de la crypte est ainsi réduite d'environ 20% [88]. Les réseaux de vaisseaux sanguins des villosités intestinales de souris adultes initialement axéniques sont deux fois moins denses par rapport aux souris conventionnelles [42]. Le

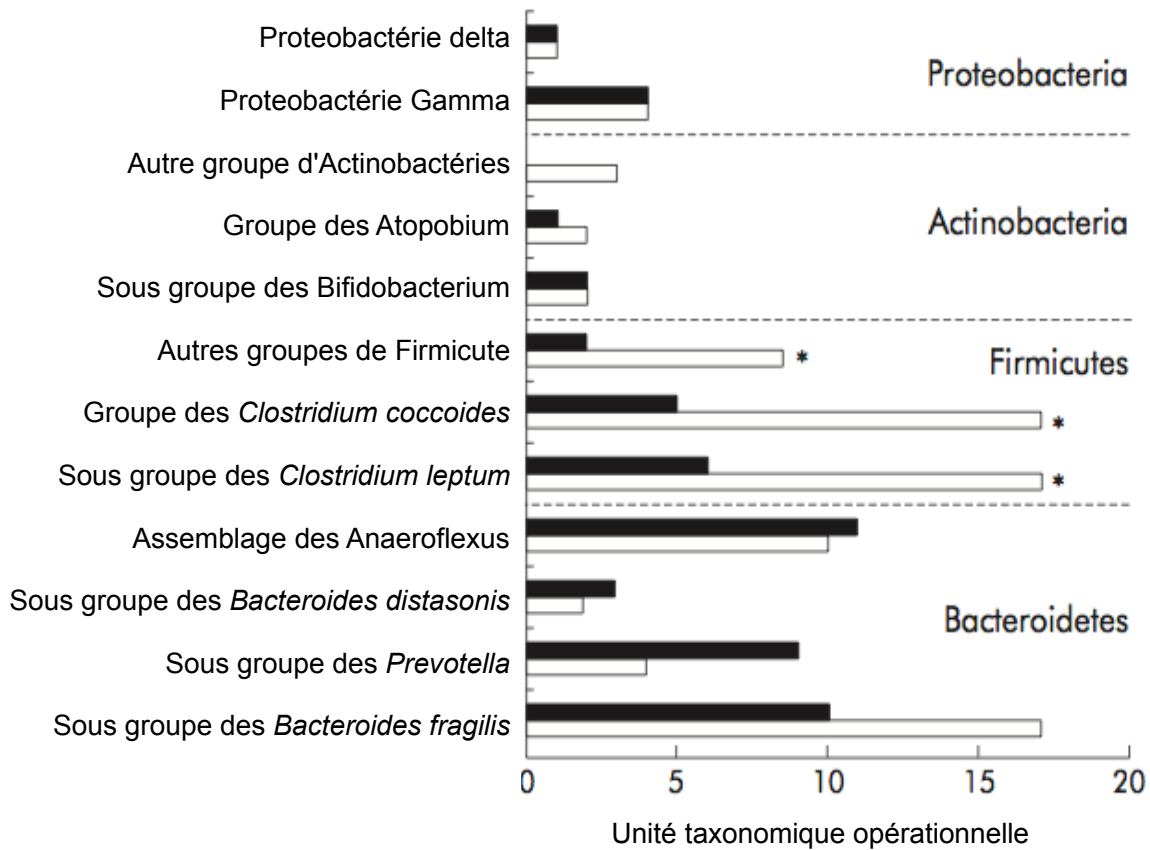


Figure 4 : la diversité du microbiote intestinal humain en condition physiologique ou pathologique. Modification de la composante bactérienne du microbiote intestinal entre des patients sains (diagramme blanc) et atteints de la maladie de Crohn (diagramme noir). Dans la maladie de Crohn une diminution significative des OTU des Firmicutes est observée par rapport aux patients sains. D'après Manichanh et collaborateurs, [63].

microbiote intestinal joue donc un rôle important dans l'angiogenèse intestinale.

Le microbiote est un acteur essentiel dans le développement et la maturation du système immunitaire. Chez les animaux axéniques, le système immunitaire est peu développé, déficient [89–91], et la tolérance du microbiote est affectée [92–94]. Les anomalies observées ne se limitent pas à l'épithélium intestinal puisque le développement de la rate et des ganglions lymphatiques est également altéré [95].

1.B.e. Microbiote intestinal et pathologies associées

Comme l'ensemble de nos organes, le microbiote intestinal s'adapte aux facteurs environnementaux comme l'alimentation et aux facteurs endogènes comme un système immunitaire déficient. Une dysbiose peut alors s'installer et influencer l'état de santé de l'hôte en entraînant des désordres intestinaux mais aussi des maladies métaboliques telles que l'obésité [85,96], des maladies inflammatoires chroniques de l'intestin comme la maladie de Crohn [67,97–99], des allergies [100–102], voire le cancer colo-rectal [7,8,103]. Les dysbioses associées à la maladie de Crohn ou l'obésité ont été les plus étudiées (**Figure 4**) [67,97]. Chez les patients atteints de la maladie de Crohn, le groupe « *Clostridium leptum* » et notamment l'espèce anti-inflammatoire *Faecalibacterium prausnitzii* sont fortement réduits, aussi bien en diversité qu'en abondance alors que l'on observe une colonisation anormalement élevée par des *E. coli* adhésifs invasifs pro-inflammatoires [67].

Les maladies humaines non infectieuses caractérisées par une inflammation intestinale exacerbée, telles que les allergies ou les maladies inflammatoires chroniques de l'intestin (MICI) (maladie de Crohn, rectocolite hémorragique), pourraient conduire à des activations pro-inflammatoires non contrôlées avec un impact majeur sur la composition du microbiote intestinal (**Figure 4**) [100,104]. L'inflammation rompt l'homéostasie du microbiote intestinal en favorisant les bactéries pathogènes au détriment des microorganismes symbiotiques protecteurs [105,106]. En maîtrisant l'effet létal de la réponse inflammatoire, les pathobiontes comme les bactéries entéropathogènes se servent de l'inflammation pour éliminer l'effet barrière du microbiote, éventuellement rompre la barrière physique de l'épithélium et assouvir leurs besoins métaboliques [107]. Des modifications spécifiques du microbiote ont été associées au développement de certaines maladies chez l'Homme [98,108,109], bien qu'il soit difficile de savoir s'il s'agissait de la cause ou de la conséquence. Les MICI sont caractérisées par une

réduction marquée de la diversité bactérienne dans les groupes IV et XVa de *Clostridium*, une diminution des *Bacteroidetes*, et une augmentation correspondante de *Proteobacteria* et *Bacillus* [80,109]. De façon analogue l'inflammation intestinale a été généralement liée à une augmentation marquée des Entérobactéries, notamment les *Escherichia coli* adhérents et invasifs [110–114] et une diminution d'autant des membres de la flore résidente [115–117].

1.C. Le système immunitaire intestinal

Le système immunitaire intestinal est le premier organe immunitaire de l'organisme. Le système digestif produit 60% des immunoglobulines corporelles totales et contient la majorité des lymphocytes du corps humain, évaluée à 10^6 lymphocytes par gramme de tissu [118]. Le système immunitaire associé à l'intestin, souvent nommé GALT pour « gut associated lymphoid tissue », comprend des structures organisées (plaques de Peyer et ganglions mésentériques), des structures plus diffuses (follicules lymphoïdes disséminés) et des cellules immunocompétentes dispersées dans la muqueuse. Le chorion (ou lamina propria) est un site effecteur contenant un grand nombre de cellules immunitaires activées et matures : des cellules présentatrices d'antigènes (cellules dendritiques et macrophages), des lymphocytes T dont une majorité de cellules T de type helper (CD3+, CD4+) et des cellules « mémoire », des lymphocytes cytotoxiques (CD3+, CD8+), des lymphocytes B et des plasmocytes dont 90% synthétisent les IgA [119]. Ces dernières sont excrétées dans la lumière intestinale sous une forme dimérique associée à une protéine, la pièce sécrétoire, qui leur confère une résistance à l'action des enzymes protéolytiques intestinales présentes dans l'intestin [119]. La sécrétion de ces anticorps empêche la colonisation de l'épithélium par des pathogènes et interdit la pénétration de matières étrangères nuisibles [38].

1.C.a. Les cellules M

Les cellules M de l'épithélium intestinal sont considérées comme les portes du système immunitaire muqueux. Elles captent par endocytose les antigènes endo-luminaux et les adressent par transcytose à leur membrane basolaterale [120]. Les cellules M sont situées sur une structure spécifique à la muqueuse : le dôme, qui n'est que faiblement recouvert par le mucus. Les antigènes sont donc plus facilement accessibles [121]. Ayant traversé l'épithélium, ils sont captés par les cellules présentatrices des antigènes situées sous les cellules M. Ces cellules peuvent alors rejoindre les plaques de Peyer, qui se trouvent à la base des dômes.

C'est le siège préférentiel de la présentation des antigènes endo-luminaux aux lymphocytes pour initier la réponse immunitaire acquise.

1.C.b. Le « homing » des lymphocytes

Après activation, différenciation et prolifération dans les plaques de Peyer, les lymphocytes T et B migrent vers le compartiment systémique *via* les ganglions mésentériques et la lymphe pour rejoindre le courant sanguin. Puis, ils atteignent les muqueuses où ces cellules seront fonctionnelles. Les réponses immunes initiées au niveau d'une plaque de Peyer peuvent donc être disséminées à l'ensemble des muqueuses. Le retour sur le site d'initiation intestinal s'appelle le « homing ». Les plasmocytes, dérivés des cellules B sécrètent des IgA spécifiques qui vont neutraliser les antigènes luminaux. Les cellules T cytotoxiques vont neutraliser et tuer les pathogènes ou les cellules reconnues [122,123].

1.C.c. La tolérance du microbiote

La muqueuse intestinale est au contact des agents microbiens présents dans la lumière intestinale. Le système immunitaire intestinal doit reconnaître et éliminer de nombreux micro-organismes pathogènes et molécules toxiques présents sur la muqueuse intestinale. En même temps, il ne doit pas déclencher de réponses immunitaires inflammatoires contre des antigènes alimentaires et les bactéries commensales. Cette cohabitation entre le microbiote et les cellules de l'hôte est la tolérance immune de l'organisme, qui se définit par la suppression des réponses immunes humorales et cellulaires spécifiques [124].

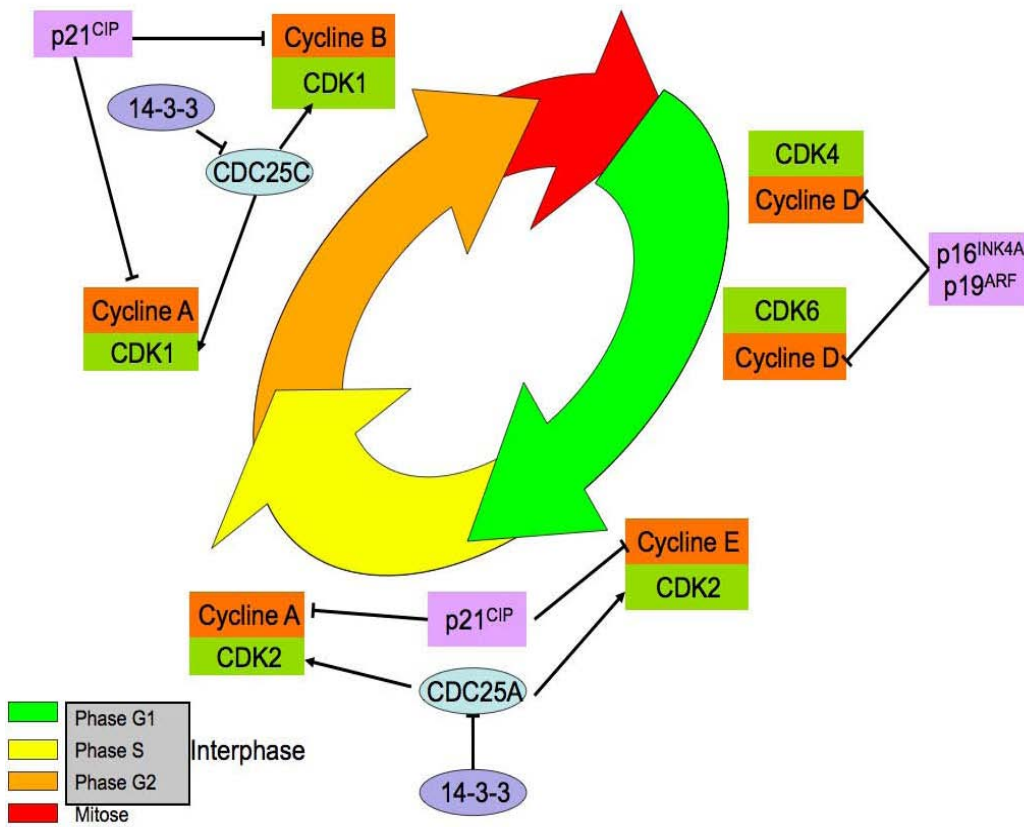
Le système immunitaire de l'intestin est capable d'élaborer deux types de réponses contradictoires, défense ou tolérance [119]. Il faut donc un système immunitaire dynamique et permissif qui puisse ajuster la réponse immunitaire en fonction du degré de danger : les bactéries commensales sont tolérées par une «inflammation constitutive physiologique de bas grade» alors que les pathogènes induisent une réaction inflammatoire intense destinée à provoquer leur élimination. Le premier des acteurs de ce système immunitaire est l'épithélium intestinal qui évalue la proximité et la densité des micro-organismes résidents [125]. Les entérocytes sondent activement les molécules bactériennes par des transporteurs situés sur leur pôle apical tels hPepT1 ou OCTN2, et peuvent également évaluer la densité de certaines bactéries. La perception des micro-organismes se fait par l'intermédiaire de récepteurs

spécifiques appelés PRR (« *pattern recognition receptors* ») qui reconnaissent spécifiquement des ligands macromoléculaires bactériens appelés MAMP (« *microbial-associated molecular patterns* »). Les PRR incluent 2 types de récepteurs : les « Nod-like receptors » (NLR) et les « Toll-like receptors » (TLR). Les NLR gardent le compartiment intracellulaire et sont alimentés en MAMP *via* des transporteurs situés sur le pôle apical ou à la faveur de l'entrée de pathogènes dans la cellule. Les TLR sont trans-membranaires et balayent l'espace extracellulaire et intra-cellulaire. L'activation de ces PRR induit l'activation des voies de signalisation pro-inflammatoires comme celles passant par le facteur NF- κ B [119].

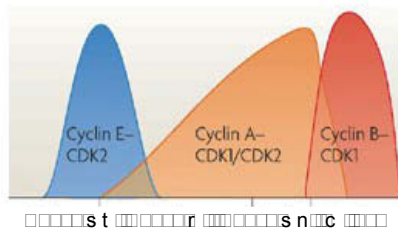
Les antigènes bactériens du compartiment luminal activent les PRR membranaires et NLR cytoplasmiques permettant ainsi aux entérocytes de distinguer les microorganismes commensaux des pathogènes [44]. Leur réponse aux MAMPs commensaux induit un signal de tolérance immunitaire *via* le TGF- β , la *thymic stromal lymphopoietin* (TSLP) ou l'IL-10 produit par les lymphocytes T régulateurs sous-épithéliaux [125]. Ces médiateurs anti-inflammatoires concourent à établir cette réponse immunitaire physiologique de bas grade envers le microbiote intestinal. Parallèlement, en réponse à ces antigènes commensaux, l'épithélium intestinal exprime de multiples effecteurs anti-microbiens comme les peptides antimicrobiens (defensines), des espèces réactives oxygénées (ROS pour « reactive oxygen species ») et le mucus intestinal [100]. L'action concertée de ces composants protecteurs aide au maintien d'un « no man's land » bactérien au delà de l'épithélium de surface.

Si les microorganismes commensaux n'expriment pas de facteurs de virulence et évoluent le plus souvent à distance des entérocytes protégés par le mucus, les microorganismes pathogènes possèdent en revanche un arsenal de facteurs d'adhésion et d'invasion qui leur permet d'entrer en contact avec les entérocytes voire même d'envahir l'épithélium. Cela provoque une activation intense des NLR et des TLR qui engendrent la biosynthèse d'effecteurs antimicrobiens, de cytokines et chimiokines inflammatoires (« tumor necrosis factor », interféron- γ (IFN- γ), IL-6, IL-8 et IL-23), de prostaglandines pour provoquer une réaction inflammatoire majeure [126]. Cette réaction inflammatoire intense permet le recrutement des cellules immunitaires au niveau des foyers d'invasion, pour concourir à l'élimination des bactéries pathogènes.

□



□



r s n e h d (a h u h d n A e d L y i n) 4

□

□

2. Processus de maintenance de l'intégrité du génome

Si le microbiote, l'épithélium et le système immunitaire contribuent à l'homéostasie intestinale, d'autres éléments permettent au niveau cellulaire de maintenir l'intégrité du génome de l'hôte. L'objectif de cette section est de décrire les principaux processus altérés lors de la carcinogenèse.

2.A. Le cycle cellulaire

2.A.a. Les principales étapes

Le cycle de vie des cellules eucaryotes est un ensemble d'événements conservés et ordonnés, qui régissent la croissance et la division cellulaire. Lorsqu'elles ne se divisent pas, les cellules sont dites quiescentes ou en phase G_0 . Elles représentent une partie importante des cellules du corps humain. Le passage de ces cellules dans une phase de prolifération nécessite des signaux mitogènes qui déclenchent le cycle cellulaire. Ce dernier comprend une période de croissance cellulaire appelée interphase, suivie de la mitose (**Figure 5**). L'interphase comprend trois phases : (i) la phase G_1 (« Gap 1») qui prépare la synthèse de l'ADN et assure le passage d'un point de non-retour rendant le processus de division indépendant des signaux mitogènes, (ii) la phase S (« Synthesis») de réplication de l'ADN et (iii) la phase G_2 (« Gap 2») de préparation à la mitose. La phase M de mitose correspond à la ségrégation des chromosomes répliqués et à la répartition des organites dans les cellules filles. Elle est classiquement subdivisée en 4 phases appelées prophase, métaphase, anaphase et télophase. Ces processus sont étroitement régulés par des mécanismes qui permettent ou limitent la progression du cycle cellulaire afin d'assurer l'intégrité des cellules filles [127].

2.A.b. La régulation du cycle cellulaire

La régulation du cycle cellulaire et sa progression sont principalement contrôlées par des sérines/thréonines kinases dépendantes des cyclines ou CDK (« cyclin-dependent kinase») (**Figure 5**). Le rôle des CDK est de phosphoryler, au niveau de sites consensus, des protéines clés de la progression du cycle cellulaire [127]. Le niveau de production des CDK est stable au

cours du cycle cellulaire mais leur activité est contrôlée par de nombreux effecteurs. Les premiers d'entre eux sont des cofacteurs protéiques appelés cyclines dont l'expression varie au cours du cycle cellulaire pour réguler l'avancement de ce dernier (**Figure 5**).

Indépendamment des cyclines, l'activité des CDK est également régulée par des phosphorylations activatrices ou inhibitrices de résidus thréonine et tyrosine. Elles seront actives ou non selon le profil des phosphorylations portées. Le complexe CDK7-cycline H, appelé aussi CAK (« Cdk-Activating Kinase ») active les principaux complexes CDK-cycline par phosphorylation. L'activation passe également par la déphosphorylation d'autres sites grâce à trois isoformes (a, b et c) des phosphatases Cdc25, qui jouent un rôle essentiel dans les transitions G_1/S et G_2/M . L'activité des CDK peut être bloquée par des protéines inhibitrices du cycle cellulaire, appelés inhibiteurs de CDK (« Cycline-dependent Kinase Inhibitor » ou CKI), qui se lient aux CDK ou aux complexes CDK-cycline. Deux familles de CKI ont été décrites, la famille INK4 comprenant $p15^{INK4B}$, $p16^{INK4A}$ et $p18^{INK4C}$, et la famille Cip/Kip qui est composée des protéines $p21^{Cip1}$, $p27^{Kip1}$ et $p57^{Kip2}$.

Enfin, la localisation intracellulaire des protéines effectrices du cycle cellulaire contribue aussi à sa régulation. Le groupe de protéines 14-3-3 régule le trafic intracellulaire de différentes protéines dont notamment la phosphatase Cdc25c. Elles maintiennent cette dernière dans le cytoplasme pendant l'interphase afin d'éviter l'activation de CDK1.

2.A.c. Les points de contrôle

Les effecteurs de la régulation du cycle cellulaire sont mis en jeu au niveau de points de contrôle qui peuvent stopper l'avancée du cycle cellulaire pour permettre l'achèvement d'une étape ou réparer des erreurs ou altérations survenues lors du cycle cellulaire. Ces points de contrôle assurent la surveillance de paramètres essentiels à la poursuite du cycle cellulaire et les trois principaux sont :

- le DDCP (« DNA Damage check point ») qui contrôle l'intégrité du génome. Activé lors des transitions G_1/S et G_2/M , il bloque l'avancée du cycle pendant que la machinerie cellulaire recherche d'éventuelles erreurs ou cassures de l'ADN.
- le RCP (« replication check point ») qui vérifie l'achèvement de la réplication avant l'entrée en mitose (transition S/M),

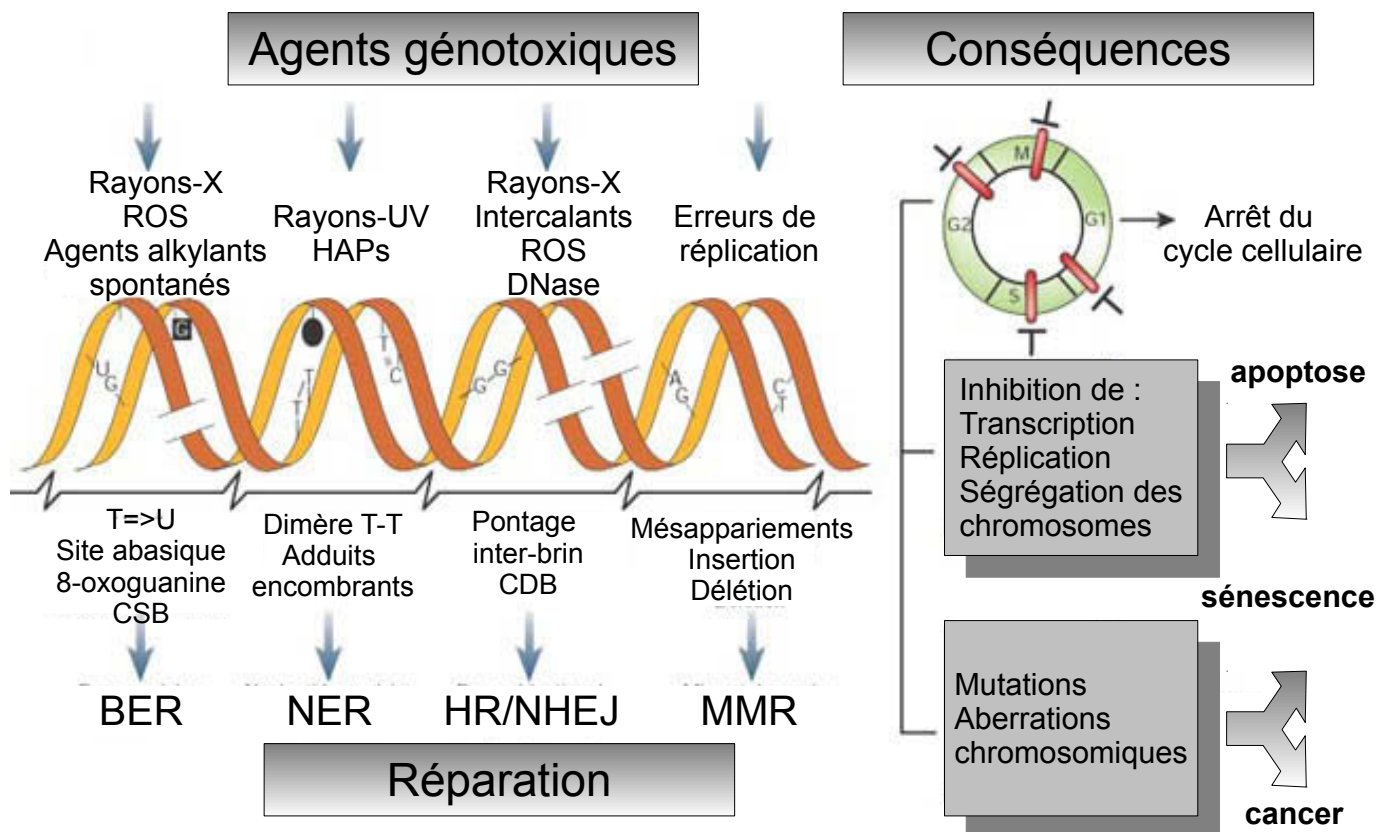


Figure 6 : les systèmes de réparation de l'ADN. Les agents génotoxiques peuvent induire différents types de lésions à l'ADN. Les systèmes de réparations de l'ADN mis en jeu sont différents suivant le type de lésion : le BER (« base excision repair »), le NER (nucleotide-excision repair), le HR/NHEJ (« Homologous recombination / non-homologous dna end joining ») et le MMR (« mismatch repair »). CDB : cassure double brin ; CSB : cassure simple brin ; HAP : hydrocarbure aromatique polycyclique ; ROS : radicaux oxygénés ; T-T : dimère de thymine ; T=>U : conversion thymine en uracil ; UV : ultra-violet. D'après Hoeijmakers, [222].

- le MCP (« mitotic checkpoint ») qui contrôle le bon positionnement des chromosomes sur le fuseau mitotique avant la ségrégation des chromatides (transition métaphase-anaphase).

2.B. Les systèmes de réparation de l'ADN

Les cellules sont en permanence au contact d'agents altérant l'ADN qui peuvent affecter l'intégrité du patrimoine génétique. Des erreurs survenant lors de la réplication peuvent également affecter ce dernier. Afin d'éviter l'apparition et la transmission de ces erreurs ou lésions, les cellules disposent de plusieurs systèmes de réparation adaptés à la nature des dommages. Une altération de ces mécanismes de réparation favorise la carcinogenèse en créant un état permissif à l'apparition de mutations.

2.B.a. Les systèmes BER, NER et MMR

Les systèmes de réparation BER (« Base Excision Repair »), NER (« Nucleotide Excision Repair ») et MMR (« Mismatch Repair ») reposent sur une même logique de fonctionnement ; la reconnaissance des lésions (bases altérées pour BER, ponts intra-brins pour NER et mésappariement pour MMR), l'excision et le remplacement des nucléotides altérés ou mésappariés (**Figure 6**) [128,129].

Le système BER met en jeu une ADN glycosylase qui se fixe au niveau de la base modifiée et clive la liaison N-glycosidique. Cette hydrolyse entraîne l'apparition d'un site apurique-apyrimidique (AP), qui est ensuite clivé par une AP-endonuclease I. L'ADN polymérase β ajoute ensuite un nucléotide au niveau de l'extrémité hydroxyle libre. Enfin, la liaison phosphodiester entre le nucléotide incorporé et le suivant est rétablie par une AP ligase.

Le système NER effectue les réparations via l'excision sur le brin endommagé d'un segment de plusieurs nucléotides encadrant la lésion. Il fait intervenir un ensemble de facteurs protéiques. Tout d'abord, des facteurs de reconnaissance qui se lient à la région endommagée, puis des hélicases afin d'ouvrir la molécule d'ADN de part et d'autre. Ensuite vont intervenir des endonucléases simple-brin pour couper en 5' et en 3' de la lésion. Après excision, une ADN polymérase synthétise le brin éliminé en utilisant le brin intact comme matrice. La réaction est terminée par l'action d'une ADN ligase qui restaure le squelette phosphodiester.

Système de recombinaison non homologue (NHEJ)

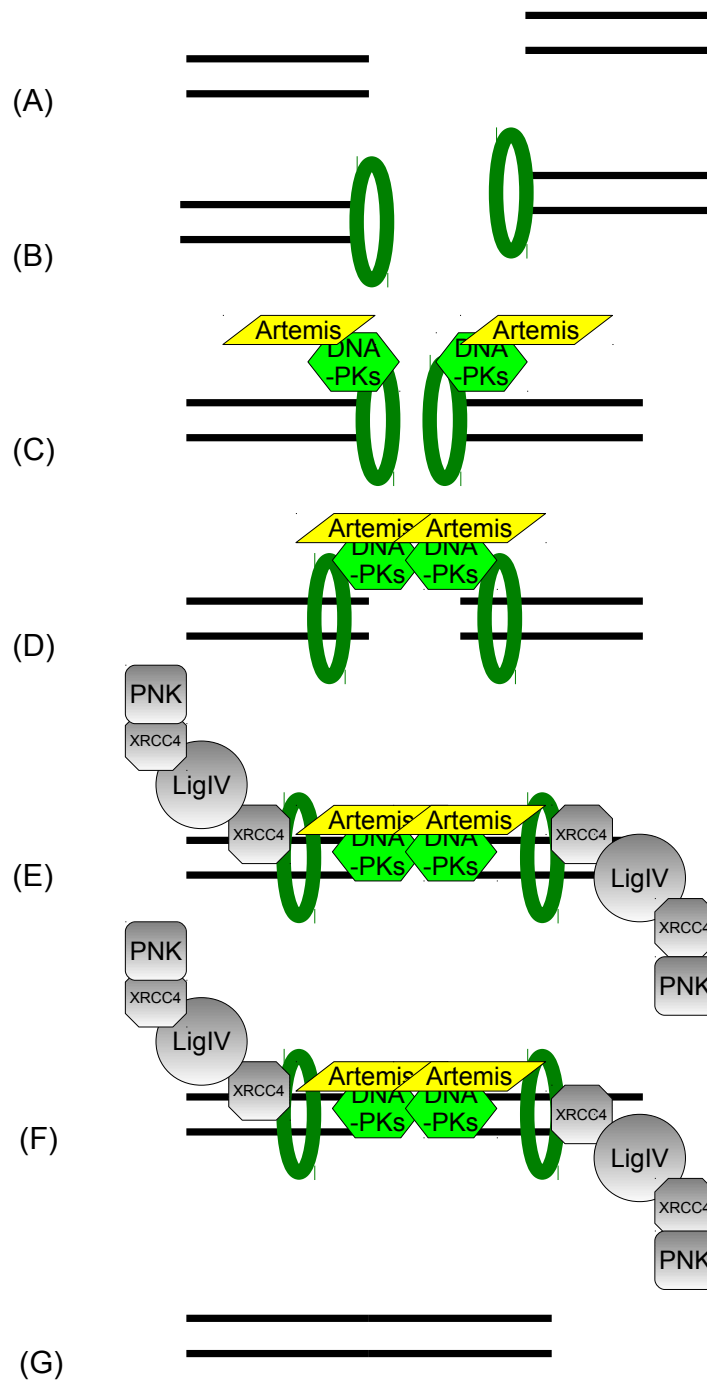


Figure 7 : le système de réparation des cassures double brin de l'ADN par recombinaison non homologue. L'apparition de cassures double brin (A) conduit au recrutement des hétérodimères Ku (cercle vert) au niveau des lésions de l'ADN (B) sont alors recruté la DNA protéine kinase (DNA-PKs) et Artemis (C). Le complexe DNA-PK/Artemis permet le pontage entre les sites endommagés (D). Suite à ce pontage, l'ADN ligase IV (LigIV) et son facteur associé XRCC4 sont recrutés (E). XRCC4 interagit avec la polynucléotide kinase (PNK) (F), qui permet à l'ADN ligase IV la ligation des deux extrémités de la molécule d'ADN par l'ADN ligase IV (G). D'après Hefferin et Tomkinson, [133]

Le système MMR (*MisMatch Repair*) reconnaît les mésappariements de bases générés par les phénomènes de tautomérisation et par la réplication. Tout d'abord, les mésappariements sont reconnus par le complexe MSH2-MSH6-MLH1. Cela permet de recruter différentes enzymes qui vont cliver en 5' et en 3' le brin porteur du mésappariement, puis de le dégrader. Il y a ensuite reconstruction de la double hélice par l'ADN polymérase δ qui utilise le brin intact comme matrice. Enfin, la ligation des extrémités par une ADN ligase rétablit l'intégrité de la molécule d'ADN.

2.B.b. Les systèmes par recombinaison

Les cassures double brin de l'ADN sont les lésions les plus délétères. Elles peuvent être prises en charge par les systèmes de réparation par recombinaison homologue (ou HR « homologous recombination ») ou recombinaison non homologue (ou NHEJ pour « non-homologous end joining »). Le système HR prend place en fin de phase S ou G2 car il utilise la chromatide sœur comme matrice pour la réparation. Les principales protéines impliquées dans ce système sont Rad50, Rad51, Rad52, Rad54, NBS1 (Nijmegen Breakage Syndrome-1) et MRE11 (Meiotic Recombination-11) [130,131]. La protéine Rad52 reconnaît les cassures double brin de l'ADN et permet le recrutement des protéines RPA (Replicative Protein A), Rad51, Rad52, Rad54, BRCA1, BRCA2 (Breast Cancer Susceptibility Proteins) et BARD1 (BRCA1-associated RING domain-1) au niveau de l'extrémité lésée. Ces dernières assurent la détection de la chromatide sœur non endommagée et le complexe nucleoprotéique formé permet une recombinaison homologue entre la séquence d'ADN altérée et la séquence homologue non modifiée de la chromatide sœur [132]. Le duplex ainsi formé permettra l'échange d'un brin d'ADN, qui sera utilisé comme matrice par l'ADN polymérase pour produire le brin complémentaire dépourvu d'erreur.

Le système de recombinaison non homologue (NHEJ) établit une liaison covalente entre deux extrémités clivées sans contrôle de l'homologie entre ces séquences. Les réparations par ce système s'accompagnent souvent de délétions. Son fonctionnement repose sur la reconnaissance du site de coupure par l'hétérodimère Ku70/Ku80 et le recrutement secondaire des effecteurs DNA-PKcs (DNA-dependent Protein Kinase), XRCC4 (X-ray Repair Cross-Complementing protein 4), XRCC7, Rad50, MRE11, NBS1 et de la DNA Ligase-IV (**Figure 7**) [132].

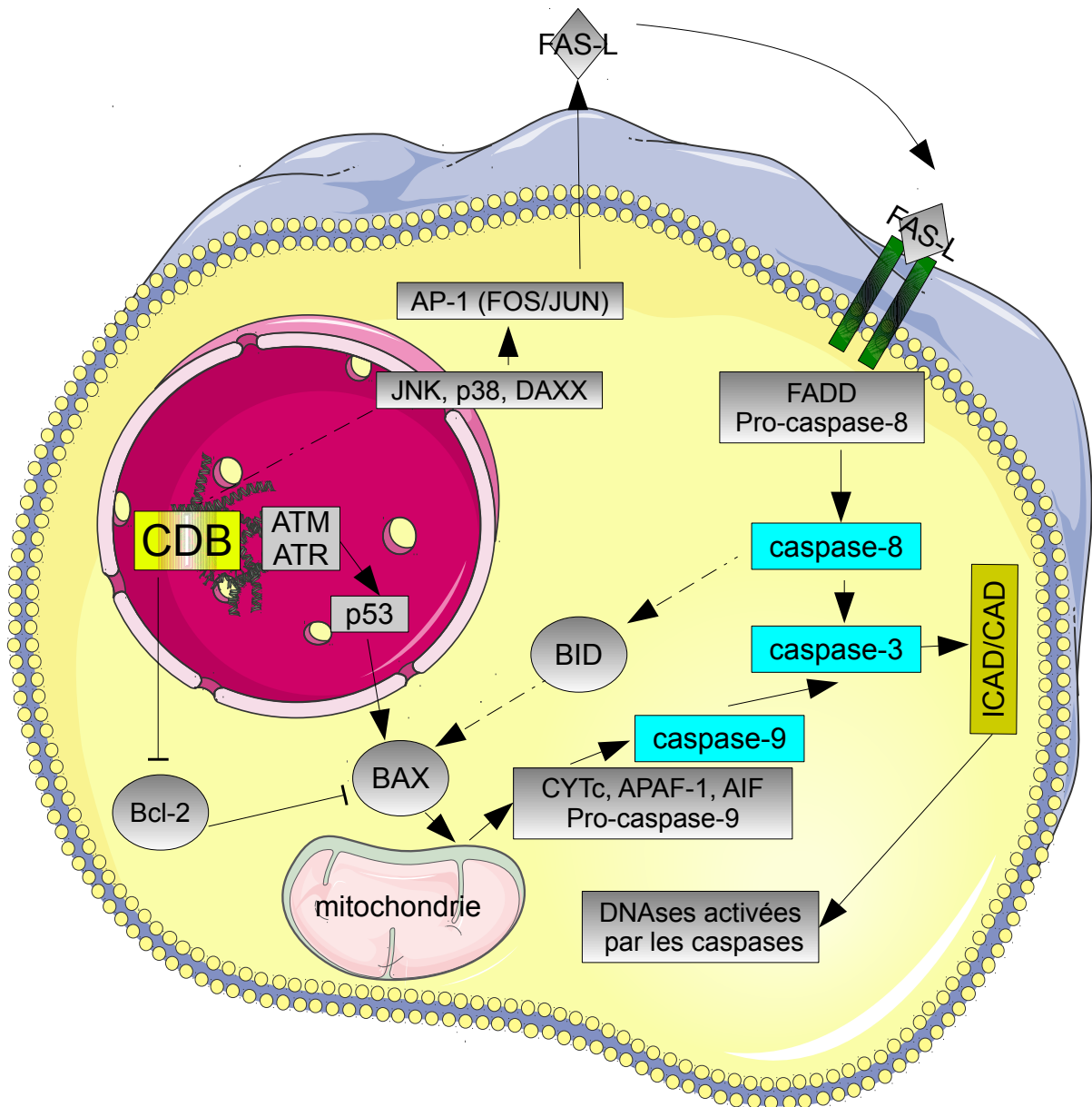


Figure 8 : signalisation cellulaire associée à l'apoptose. Les deux voies principales associées à l'apoptose sont les voies intrinsèque et extrinsèque. **La voie extrinsèque** fait intervenir des récepteurs de mort membranaire (FAS) couplés à la pro-caspase-8 via leur domaine FADD. La fixation de FAS-L à son récepteur entraîne l'activation de la caspase-8 conduisant à l'activation de la caspase-3 et BID. Cette dernière permet l'activation de la voie mitochondriale. **La voie intrinsèque** ou mitochondriale fait intervenir un déséquilibre entre les niveaux de protéines anti-apoptotiques (Bcl-2) et pro-apoptotiques (BAX) au profit de ces dernières. Suite à ce déséquilibre, il y a une perte de perméabilité de la membrane mitochondriale et relargage d'effecteurs (Cytochrome C, APAF-1, AIF, Pro-caspase-9) dans le cytoplasme permettant la formation d'un complexe CYTc/APAF-1/pro-caspase-9 activant la caspase-9, qui à son tour, active la caspase-3. La caspase-3 active les nucléases CAD qui fragmentent l'ADN nucléaire. AP-1 : Activator Protein 1 ; APAF-1 : Apoptotic Protease Activating Factor 1 ; ATM : *Ataxia Telangiectasia Mutated* ; ATR : *Ataxia Telangiectasia Rad3-related* ; BAX : Bcl-2-Associated X protein ; Bcl-2 : B-cell lymphoma 2 ; BID : BH3 Interacting-Domain death agonist ; CAD : Caspase-Activated Dnase ; CDB : Cassure Double Brin ; CYTc : Cytochrome C ; DAXX : Death-associated protein 6 ; FADD : Fas Associated Death Domain ; Fas-L : Fas ligand ; JNK : c-Jun N-terminal kinases ; ICAD : Inhibitor of Caspase-Activated DNase. D'après Lowe et collaborateurs, [217].

2.C. La mort cellulaire par apoptose

En réponse à des stress extrinsèques ou intrinsèques comme des lésions de l'ADN irréparables, les cellules eucaryotes peuvent mourir selon différents processus. La nature du stress, son intensité et le type cellulaire vont définir le processus enclenché. Nous traiterons ici de l'apoptose, qui joue un rôle prépondérant dans la prévention de la carcinogenèse.

2.C.a. Généralités

L'apoptose est un processus de mort cellulaire contrôlé qui est impliqué dans l'homéostasie cellulaire [133], le développement embryonnaire [134], la réponse au stress ou à des signaux de suicide [135–137]. Une cellule normale a constamment besoin que son environnement lui confirme son utilité aux moyens de facteurs de croissance. La perte de ces signaux peut déclencher un processus apoptotique. Les dommages de l'ADN non réparables, la dégradation des télomères, le stress du réticulum endoplasmique (accumulation de protéines anormalement repliées), la perte des contacts entre cellules ou avec la matrice extracellulaire environnante, et des signaux hormonaux (glucocorticoïdes notamment) sont aussi capables de déclencher l'apoptose. Deux voies principales d'activation de l'apoptose ont été décrites, la voie intrinsèque et la voie extrinsèque. Elles aboutissent à l'activation de protéases à cystéine active appelées caspases, et notamment un effecteur commun aux deux voies : la caspase 3, qui conduit au suicide de la cellule.

2.C.b. La voie extrinsèque

L'activation de la voie extrinsèque repose sur des signaux extracellulaires et des récepteurs de mort cellulaire transmembranaires appelés Fas et appartenant à la famille du récepteur du TNF (**Figure 8**). La fixation du ligand Fas-L (Fas-ligand) ou TNF- α à l'extrémité extra-cellulaire du récepteur entraîne l'activation de ce dernier sous forme de trimère. L'extrémité intracellulaire C-terminale comporte une séquence formant un « domaine de mort » (« Death Domain » ou DD). Les trois DD du trimère FAS activé recrutent une protéine cytoplasmique FADD (« Fas Associating protein with Death Domain ») comportant deux domaines : un domaine DD qui forme un dimère avec le domaine DD de FAS et un domaine DED (Death Effector Domain) qui fixe et active par clivage la pro-caspase 8 en caspase 8

active. FAS induit ainsi une cascade d'activations protéiques débouchant sur l'activation de la caspase 3

2.C.c. La voie intrinsèque ou mitochondriale

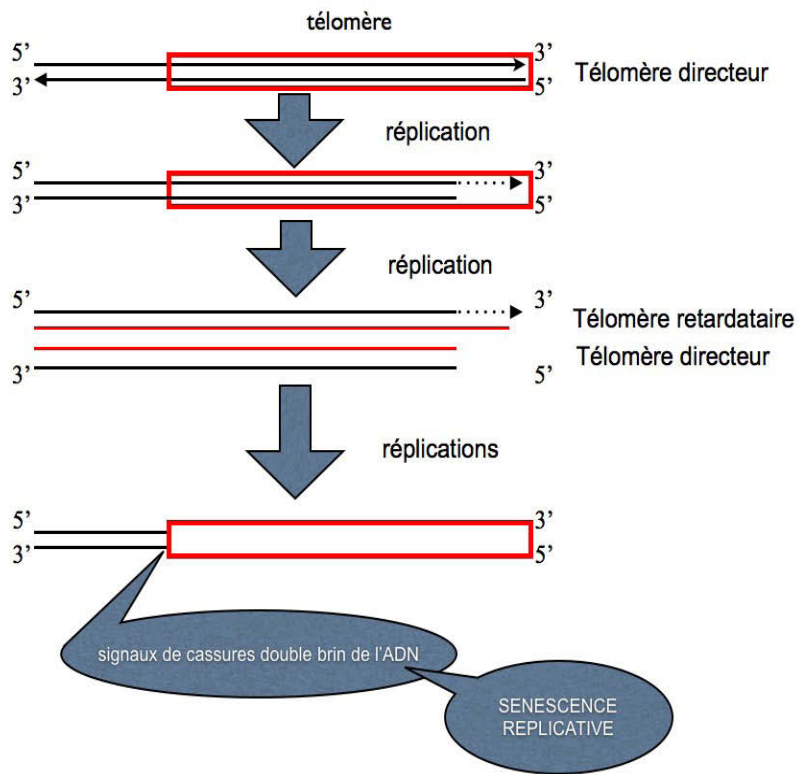
La voie intrinsèque ou mitochondriale est sous le contrôle de protéines intracellulaires appartenant à la famille de Bcl-2 (**Figure 8**). Ces protéines sont pro- (ex. BAX) ou anti- (ex. Bcl2) apoptotiques. Lorsque l'équilibre entre les signaux pro- et anti-apoptotiques est rompu au profit de la voie pro-apoptotique, le cytochrome C mitochondrial est libéré dans le cytoplasme. Ce dernier s'associe à APAF-1 (Apoptotic peptidase activating factor 1) pour former l'apoptosome qui active en cascade la caspase 9 puis la caspase 3.

2.C.d. Les caractéristiques des cellules apoptotiques

L'activation de la caspase 3 initie une cascade d'activations de caspases pour entraîner le clivage de protéines cytosoliques (Kératine, Gelsoline, Actine...), nucléaires (Lamine, Actine...), et une dégradation intranucléosomale de l'ADN par les protéines CAD (Caspase activated DNase). Ce processus aboutit à une fragmentation de l'ADN en séquence de 180 nucléotides. Sur le plan histologique, on constate dans les cellules concernées une compaction et une marginalisation de la chromatine nucléaire, une condensation du cytoplasme ainsi qu'une convulsion des membranes cytoplasmiques et nucléaires. L'intégrité des membranes est donc conservée et les corps apoptotiques sont éliminés par le système immunitaire sans réaction inflammatoire. L'apoptose permet ainsi d'éliminer des cellules sans dommage pour les tissus adjacents.

2.D. La sénescence cellulaire

Les cellules disposent d'une alternative à l'apoptose pour limiter la prolifération de cellules anormales : la sénescence cellulaire [138]. Ce processus cellulaire fait référence à un état physiologique marqué par un arrêt de la croissance et un blocage du cycle cellulaire souvent décrit comme irréversible. Deux types de sénescence cellulaire ont été identifiés à ce jour : la sénescence répllicative (SR) et la sénescence induite par le stress (SIS) [139,140].



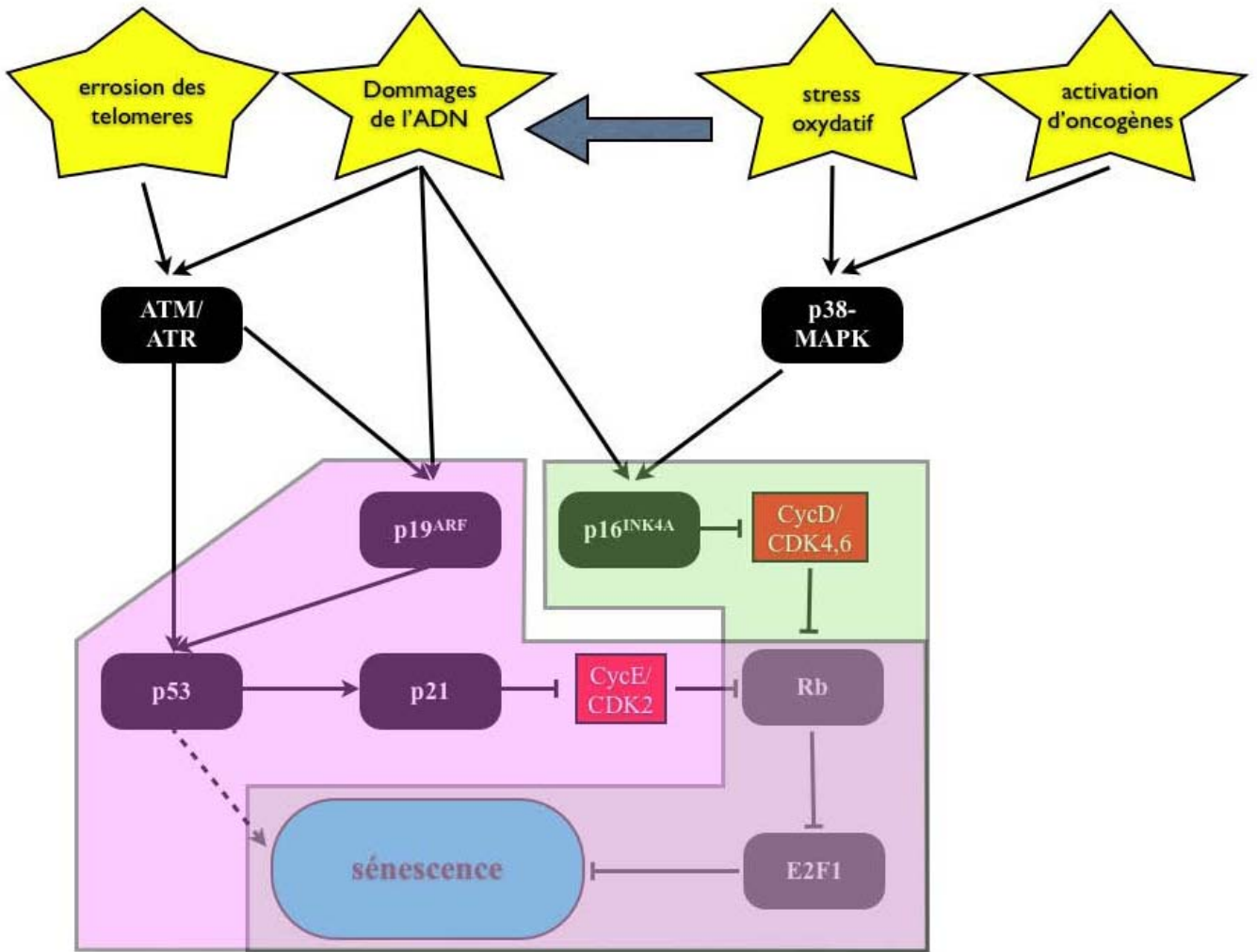
2.D.a. Les sénescences répliquatives et induites par le stress

La sénescence répliquative permet de réguler le nombre de réplifications des cellules eucaryotes. Elle est déclenchée par le raccourcissement, jusqu'à épuisement, des télomères après un nombre défini de réplifications (**Figure 9**). Ce nombre de réplifications est variable selon le type cellulaire. Les régions télomériques sont constituées de répétitions de la séquence TTAGGG associées à des complexes protéiques spécifiques comprenant ATM (*ataxia telangiectasia mutated*), le complexe MRN (MRE11/RAD50/NBS1) et l'hétérodimère Ku80/Ku70. L'ensemble adopte une conformation particulière appelée T-loop protégeant physiquement l'extrémité 3' de l'ADN des activités DNase ou de fusion télomérique. La stabilité de ces boucles est assurée par la longueur du télomère et la longueur des télomères est maintenue par des enzymes spécifiques appelées télomérases. Cependant, ces enzymes ne sont pas produites dans la plupart des cellules. Aussi, l'ADN télomérique s'érode à chaque phase S car l'ADN polymérase, fonctionnant de façon unidirectionnelle, ne peut pas produire de séquence télomérique. Après un certain nombre de divisions cellulaires, les chromosomes rognés à leurs extrémités télomériques présentent alors des dommages persistants, qui s'apparentent à des cassures double brin de l'ADN. La signalisation cellulaire détectant les cassures double brin est activée et aboutit au blocage du cycle cellulaire généralement en phase G1. Ainsi, les cellules eucaryotes ont une capacité de réplification limitée [141].

La sénescence induite par un stress, encore appelée sénescence cellulaire prématurée, a été décrite secondairement. La SIS peut être observée sur des lignées cellulaires immortelles contrairement à la SR et l'induction de la SIS dans des cellules cancéreuses constitue une approche thérapeutique des cancers [142,143]. Cette sénescence cellulaire peut être occasionnée par différents types de stress induisant des dommages directs ou indirects à l'ADN, comme les agents alkylants de l'ADN ou les stress oxydatifs, mais aussi la surexpression d'oncogènes (RAS, RAF, BRAF, c-MYC,...) [139] ou la perte de gènes suppresseurs de tumeur (PTEN, VHL, NF1...).

2.D.b. Les voies de signalisation cellulaire de la sénescence cellulaire

Deux voies de signalisation cellulaire engendrent la sénescence. Elles sont désignées par leurs effecteurs majeurs : les protéines suppresseurs de tumeurs p53 et Rb (Retinoblastoma) (**Figure 10**). Dans un modèle de fibroblastes embryonnaires murins, l'inactivation de p53 ou Rb prévient l'entrée en sénescence des cellules [144,145].



Con.pv³ ré Co
 é e Co
 u x é
 d u c 1
 é Oé c s a à / (5

➤ La voie p53/p21^{Cip}

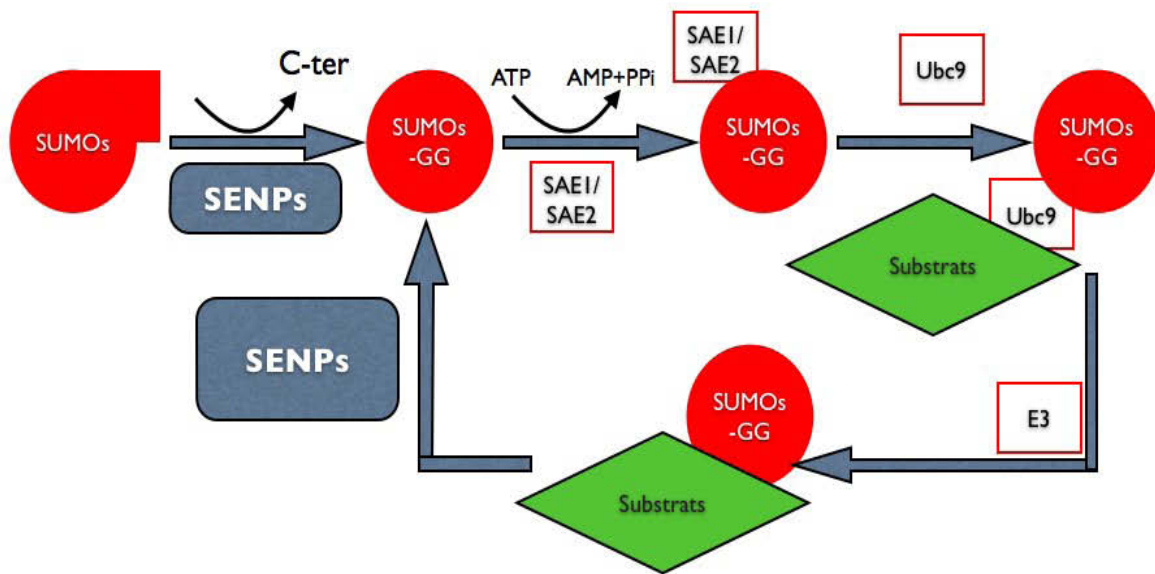
La protéine p53 est qualifiée de « gardien du génome » car elle joue un rôle central dans la prolifération, la différenciation, l'apoptose et la sénescence cellulaire [144]. Son importance dans l'homéostasie cellulaire est confirmée par son statut muté dans plus de 50% des cancers.

Dans la sénescence cellulaire, l'activation de la voie p53 peut faire suite à des dommages à l'ADN. Les voies de signalisation empruntant les kinases ATR/Chk1 ou ATM/Chk2 sont activées par les cassures de l'ADN simples et doubles brins respectivement [140,146].. Ces processus aboutissent à la phosphorylation de p53 sur la serine 15. Cette modification post-traductionnelle induit un blocage de l'ubiquitination par la E3-Ubiquitine ligase MDM2, un blocage de sa translocation au cytoplasme et donc de sa dégradation par le protéasome [147]. Le deuxième axe de régulation de p53 fait intervenir l'expression du gène codant p19^{ARF} suite à un stress métabolique (déplétion en nutriments) ou le stress oncogénique (expression d'oncogènes tels que c-Myc ou RAS) [145]. La protéine p19^{ARF} interagit alors avec MDM2, qui ne peut plus orienter p53 vers une dégradation dépendante du protéasome.

Après son activation la protéine p53 active la transcription de nombreux gènes dont celui codant p21^{Cip} (**Figure 10**) [148,149]. L'accumulation de p21^{Cip} entraîne un blocage du cycle cellulaire en inhibant le complexe CDK2-cycline E et à niveau moindre CDK4-cycline D [127]. L'inhibition de ces complexes, qui phosphorylent la protéine Rb, induit l'hypophosphorylation de Rb et donc la fixation du facteur de transcription E2F1. Or, ce dernier est nécessaire à la production de cycline A, qui est indispensable à la progression du cycle cellulaire [127]. Ce blocage du cycle cellulaire peut aboutir à la sénescence cellulaire.

➤ La voie p16^{ink4A}/Rb

La surexpression de la protéine p16^{ink4A} suite à un stress comme l'activation de l'oncogène RAS, conduit à la sénescence cellulaire (**Figure 10**) [150]. La protéine p16^{INK4A} bloque la phosphorylation de Rb par le complexe CDK4-cycline D. Cette hypophosphorylation de Rb stabilise l'interaction de Rb avec le facteur de transcription E2F1, qui est nécessaire à la progression du cycle cellulaire comme indiqué précédemment [127,144].



1' t o n E m
 1p3 4S L i 22
 1' t o L S 1 p n s l p E i
 L U a 4 r c L r p m
 L 1' t o L v x - p
 U L r p m O 1 n O 1 3 E
 1' t o L M
 1p3 U 1' t o 22 e 2 2 M p e 1' t o
 L p r M 1 p 3 e 1' t o n M 1' t o e 1
 t o é i u f n CPN

➤ Les autres voies de signalisation

Il émerge de la littérature des voies alternatives induisant la sénescence cellulaire [138,145,151]. La signalisation dépendante des MAPK (Mitogen-activated protein (MAP) kinases) conduit à la phosphorylation de p53 via p16^{INK4A} et ainsi à son activation comme précédemment. Alimonti et collaborateurs montrent que la phosphatase PTEN active p19^{ARF} et la protéine mTor [151]. Ces dernières contribuent à l'activation de p53. Dans cette signalisation dépendante de PTEN (« Phosphatase and tensin homolog »), la transcription du gène codant la protéine p27^{Kip1} est augmentée. Cet homologue fonctionnel de p21^{Cip} conduit au blocage du cycle cellulaire par un mécanisme identique.

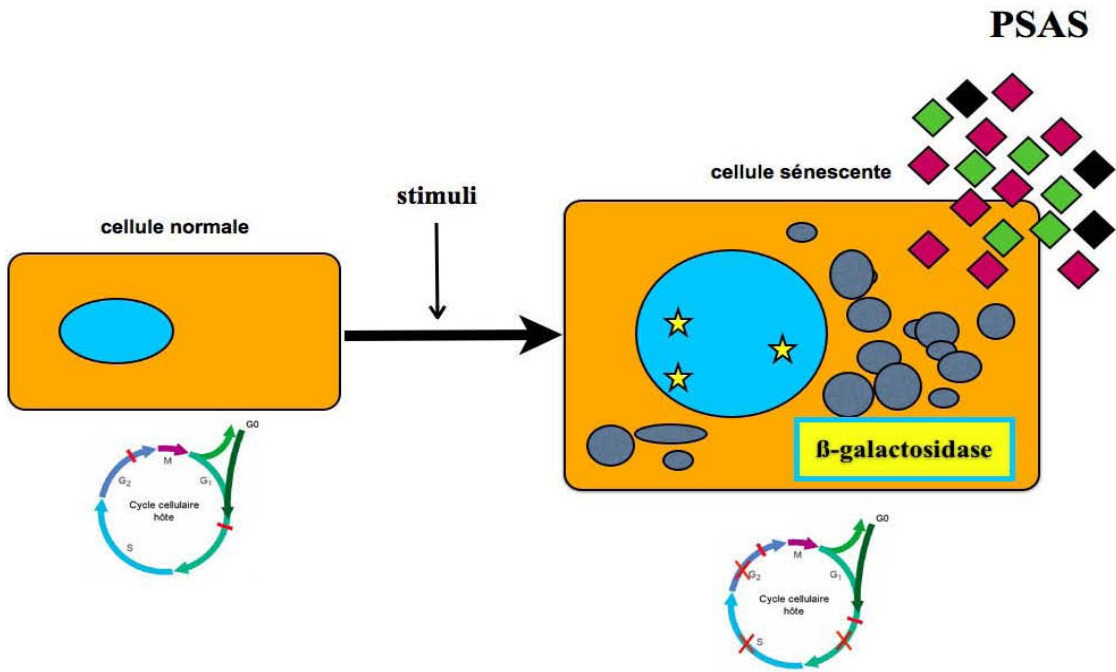
2.D.c. La SUMOylation des protéines et sénescence

La SUMOylation est une modification post-transcriptionnelle des protéines fortement associée à la sénescence cellulaire. Elle repose sur la fixation réversible de peptides SUMO (« Small Ubiquitin-related Modifier ») sur une séquence consensus cible [152,153]. Les peptides SUMO appartiennent à la famille des peptides affiliés à l'ubiquitine [153]. Bien qu'ils ne partagent que 18% d'identité avec l'ubiquitine, des travaux de cristallographie ont toutefois révélé un repliement similaire de leurs extrémités C-terminales, une zone critique pour leur activité biologique [154]. Chez l'Homme, quatre isoformes de SUMO ont été décrites (SUMO-1 à -4). SUMO-1 présente 50% d'identité avec SUMO-2 et 3, qui elles-mêmes partagent 95% d'identité. Ces protéines sont exprimées de façon ubiquitaire contrairement à SUMO-4, la dernière décrite, qui est plus divergente (86% d'homologie avec SUMO-2 [155]) et plus rare. Cette dernière est exprimée dans les cellules dendritiques et les macrophages où elle modulerait l'expression des interférons [156]. Ces données suggèrent des rôles différents pour les différents isoformes. De plus, il existe une certaine spécificité de fixation des isoformes SUMO aux cibles, même si la majorité peut être SUMOylée *in vitro* indifféremment par SUMO-1, -2 ou -3.

L'association de SUMO à sa cible est réalisée par une machinerie enzymatique comparable à celle utilisée lors de l'ubiquitination [153]. La SUMOylation exige une maturation au préalable de SUMO (**Figure 11**) [157]. En effet, SUMO est produit sous la forme d'un précurseur inactif, qui doit être clivé au niveau d'un motif GG par des peptidases à cystéine active spécifiques de SUMO appelées SENP (« sentrin-specific peptidase »). Après maturation, SUMO subit une activation dépendante de l'adénosine triphosphate (ATP) par un complexe

d'activation E1 composé de l'hétérodimère SAE1/SAE2 (« SUMO-activating enzyme subunit 1 » / « SUMO-activating enzyme subunit 2 »). Les protéines SUMO activées sont alors transférées sur l'enzyme de conjugaison E2 ou Ubc9 (« Ubiquitin-like protein SUMO1 conjugating enzyme 9 »). SUMO est ensuite conjuguée de manière covalente à la fonction amine primaire d'une lysine présente dans le motif consensus de SUMOylation ψ KxE (ψ acide aminé hydrophobe) à l'aide de ligases E3. L'enzyme Ubc9 reconnaît le motif consensus et la spécificité pour la protéine cible est apportée par les ligases E3 comme PIAS (« Protein inhibitors of activated STAT ») ou RANBP2 (« RAN binding protein 2 ») [157,158]. Tout comme la conjugaison de l'ubiquitine, la fixation de SUMO à sa cible modifie ses propriétés et le plus souvent, elle augmente sa stabilité ou modifie sa localisation cellulaire [159]. Cependant, la fixation de SUMO à sa cible est rapidement réversible. La « déconjugaison » est un processus catalysé par les peptidases SENP, qui assurent une double fonction de maturation des peptides SUMO et de déconjugaison [153,159]. Chez l'Homme, il existe six enzymes SENP 1, 2, 3, 5, 6 et 7. La comparaison de leurs structures et de leurs activités a révélé une haute affinité d'interaction entre SUMO-1 et SENP1, qui présente une activité préférentielle de désumoylation [160,161]. SENP6 et 7 ont une action de déconjugaison préférentielle sur SUMO-2 et 3. De plus, les enzymes SENP ont une localisation cellulaire différente probablement dictée par leur extrémité N-terminale peu conservée. SENP1 et 6 ont une localisation nucléaire et cytoplasmique alors que SENP2 est associé aux pores nucléaires. SENP3 et 5 sont localisés dans le nucléole.

La SUMOylation joue un rôle dans les processus de réponse au stress comme la sénescence cellulaire [140,152,157,162]. La SUMOylation de plusieurs protéines associées à la sénescence a été mise en évidence et renforce cette voie en stabilisant les protéines cibles ou en augmentant leurs activités [140,162]. Les acteurs majeurs de la signalisation cellulaire associée à la sénescence, p53 et Rb, peuvent être SUMOylés *in vivo* et *in vitro* par SUMO-1, -2 ou -3. L'expression de la ligase E3 PIASy est augmentée pendant la sénescence répllicative et corrélée avec l'hypersumoylation observée dans les cellules en culture [163]. De plus, la surexpression de PIASy induit la sénescence, *via* une SUMOylation accrue de p53, tandis que la mutation inactivatrice ou la délétion de PIASy abolit cet effet. La fixation de SUMO-1 sur p53 augmente son activité, et en conséquence, une transcription accrue du gène codant p21^{Cip} ainsi que la mise en place de la sénescence cellulaire [164]. Par ailleurs, la désSUMOylation de Rb est associée à une répression de la sénescence. La surexpression des peptides SUMO, en conduisant à une accumulation de protéines SUMOylées, peut également induire la



a n u u
 c d c x
 e mi ★
 t e ' é m
 r
 e a a

sénescence cellulaire. Cette augmentation du nombre de protéines SUMOylées peut également résulter d'une inhibition des enzymes de déconjugaison SENP ou d'une extinction du gène codant SENP1 [140,165,166]. La diminution par ARN interférence de SENP1, SENP2 et SENP7 induit la sénescence de fibroblastes humains en culture [140]. La répression même indirecte de SENP2 par la protéine de pore nucléaire Tpr induit également la sénescence prématurée de cellules en culture [167]. Il existe donc très clairement une relation forte entre les processus de SUMOylation et la sénescence cellulaire, deux processus impliqués dans la réponse aux stress.

2.D.d. Les caractéristiques des cellules sénescentes

La sénescence cellulaire est caractérisée par d'importants changements physiologiques, morphologiques, et métaboliques (**Figure 12**).

Une caractéristique de la sénescence cellulaire est le blocage du cycle cellulaire. Ce blocage est souvent décrit comme irréversible et cet arrêt peut survenir à différents stades du cycle cellulaire. Les cellules en sénescence répliquative sont généralement bloquées en phase G1 (2n) et n'entrent pas en phase S même après stimulation par des facteurs de croissance [168]. Les cellules en sénescence prématurée (SIS) peuvent être bloquées en phase G1, S ou G2 [138,169]. Ces cellules présentent, en outre, une résistance à la mort par apoptose.

La morphologie des cellules sénescentes est marquée par une mégalocytose associée à un aplatissement de la cellule et un contour irrégulier, une vacuolisation importante, une masse lysosomiale augmentée, la présence de vésicules de lipofuschine (produits de la peroxydation lipidique), des mitochondries peu denses en microscopie électronique, et des nucléoles proéminents [141,170].

Les deux caractéristiques biochimiques classiquement décrites [171,172] sont une augmentation de l'activité β -Galactosidase lysosomale à un pH 6 et un phénotype sécrétoire (**Figure 12**). Dimri et collaborateurs ont démontré que la détection d'une activité β -galactosidase à pH 6 est spécifique des cellules sénescentes et n'est pas retrouvée dans les cellules quiescentes [168]. Si les cellules sénescentes ne se divisent plus, elles ne sont pas pour autant inactives sur le plan métabolique. En effet, elles sécrètent un nombre important de facteurs, qui forment le phénotype sécrétoire associé à la sénescence (ou PSAS). Le phénotype sécrétoire varie selon le type cellulaire et le facteur déclenchant. Parmi les facteurs

sécrétés, ont été identifiés principalement des cytokines et des chimiokines pro-inflammatoires ainsi que des facteurs de croissance et des métalloprotéases. Ainsi, il émerge depuis peu, l'idée que la sénescence aurait à la fois une action anti-tumorale en bloquant le cycle cellulaire et pro-tumoral en stimulant la croissance des cellules environnantes.

3 Le cancer colorectal

3.A. Généralités

Le cancer se définit comme une pathologie ayant pour origine des mutations et des modifications épigénétiques qui conduisent à la transformation de cellules normales en cellules cancéreuses dont l'une des caractéristiques est de proliférer de façon incontrôlée. Les principaux gènes altérés appartiennent à trois familles : les oncogènes, les gènes suppresseurs de tumeurs et les gènes mutateurs.

Les oncogènes sont des gènes codant des protéines impliquées dans la prolifération cellulaire. A l'état normal, ils influent positivement la croissance et la division cellulaire ; ils sont alors appelés proto-oncogènes. Lors d'un cancer, leur activité est augmentée à cause de mutations et/ou d'une expression exagérée suite à des modifications épigénétiques. Ils prennent alors la dénomination d'oncogène. Parmi les oncogènes, on trouve le gène codant pour la protéine (Myc) et les gènes codant des petites protéines G (KRAS, RAF, RAB) impliqués dans la transduction de signaux associés à la prolifération cellulaire, la différenciation et la survie [173].

Par opposition, les gènes suppresseurs de tumeurs ou anti-oncogènes sont requis pour le fonctionnement normal des cellules et répriment la prolifération des cellules en agissant sur le cycle cellulaire et la différenciation cellulaire. Parmi ces derniers se trouvent les gènes codant pour les protéines p53 ou Rb. Les mutations carcinogènes entraînent une perte de leurs fonctions qui aboutit à une prolifération cellulaire dérégulée.

Le troisième type de gène pouvant affecter la carcinogenèse sont les gènes de type mutateur. Ce sont les gènes impliqués dans les mécanismes de réparation (systèmes BER, NER, MMR) de l'ADN ou correcteur de modifications épigénétiques. Leur mutation va favoriser l'apparition de mutations et/ou de modifications épigénétiques pour aboutir à un gain ou une perte de fonction des proto-oncogènes et des gènes suppresseurs de tumeurs.

La carcinogenèse est un processus séquentiel. Les altérations génétiques et épigénétiques vont avoir pour conséquence de modifier la physiologie des cellules par étapes.

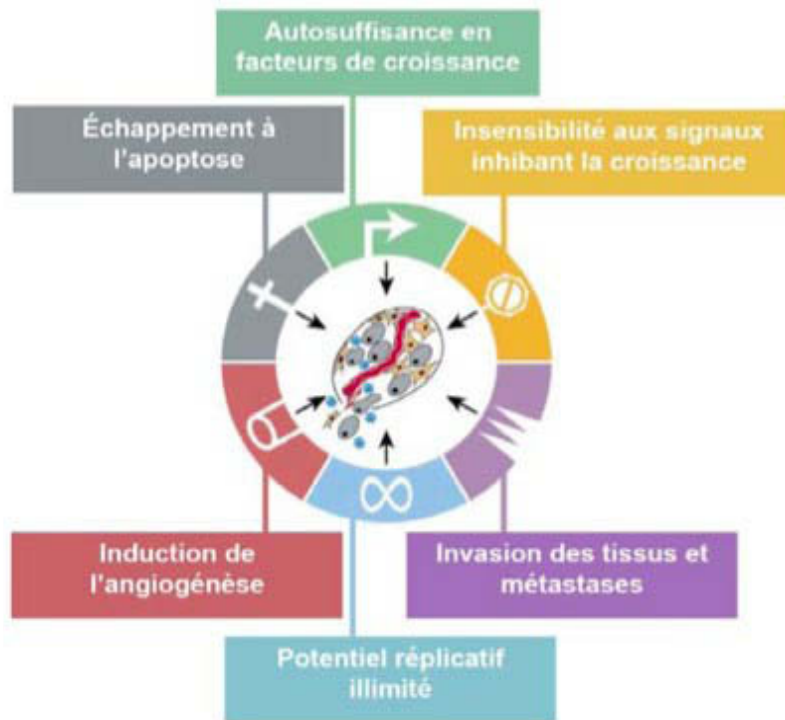


Figure 13 : les modifications acquises au cours de la tumorigenèse. Représentation des 6 caractères que doivent acquérir les cellules pour devenir cancéreuses. L'ordre d'apparition de ces caractères est variable selon le type de cancer. D'après Hanahan et Weinberg, [173]

En adéquation avec cette idée, Hanahan et Weinberg ont proposé que la formation d'un cancer nécessite l'acquisition séquentielle de 6 propriétés cellulaires (**Figure 13**) [174] :

1- Autosuffisance en signaux de croissance. La tumeur produit elle-même des signaux de croissance afin d'être indépendante de son microenvironnement.

2- Insensibilité aux signaux inhibant la croissance. Les cellules cancéreuses échappent aux signaux de contrôle cycle cellulaire.

3- Potentiel réplicatif illimité. Les cellules cancéreuses limitent l'érosion des télomères et échappent ainsi aux mécanismes de sénescence.

4- Échappement à l'apoptose. Les cellules cancéreuses deviennent insensibles aux signaux pro-apoptotique.

5- Induction de l'angiogenèse. Les cellules ont besoin d'oxygène et de nutriments pour survivre. Cet apport se fait par l'intermédiaire de vaisseaux sanguins. Les cellules cancéreuses peuvent initier la formation de nouveaux vaisseaux sanguins par un mécanisme appelé angiogenèse.

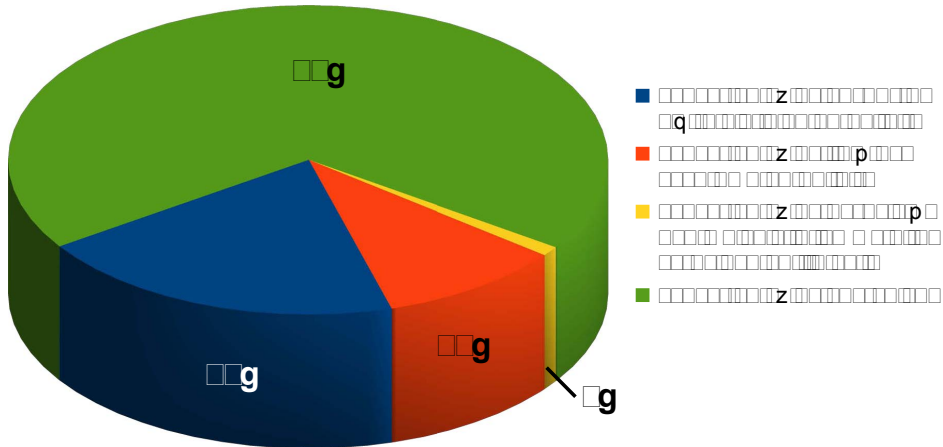
6- Invasion des tissus et métastases. Durant le développement de la plupart des cancers, des cellules vont se détacher de la tumeur et être à l'origine de la formation de tumeurs dans d'autres organes.

3.B. Épidémiologie du cancer colorectal

Selon l'Organisation Mondiale de la Santé, chaque année environ 6.000.000 personnes décèdent des suites d'un cancer et la lutte contre ces maladies est devenue un problème de santé publique dans tous les pays industrialisés. Après les cancers de la prostate et du sein, le cancer colorectal est le troisième cancer le plus fréquent chez l'homme et la femme [175]. Sa prévalence est en progression constante depuis 1975 et représentait environ 1,2 million de cas en 2008 dans le monde. Son incidence serait plus élevée dans les pays occidentaux que dans les pays en voie de développement [6]. Avec environ 609.000 décès imputés par an, c'est également la quatrième cause de décès par cancer. En France, on dénombre 37.400 nouveaux cas par an et il est responsable de 16.800 décès par an [176]. Son incidence est identique chez l'homme et la femme jusqu'à 50 ans et devient plus importante chez l'homme après 50 ans [175]. Le sexe-ratio est de 1,5. La survie à 5 ans est de 43% chez l'homme et de 46% chez la femme tous stades confondus.



I	
H	xi à u t f i
u	xf à n p u
r	n mp gi
ô	gi à n u
ô	a f p si à g p y à u (sà a p é o
e n	



F i u n t ô ' c é è g s v f s x m



3.B.a. Facteurs de risque du cancer colorectal

Les facteurs de risque du cancer colorectal (CCR) sont de deux types : les facteurs endogènes, liés à la prédisposition génétique, et les facteurs exogènes, liés au contact du sujet avec son environnement [2].

➤ Les facteurs de risque endogènes

Les facteurs de risque endogènes sont génétiques, congénitaux ou acquis, liés à une affection ou un terrain prédisposant (**Figure 14**).

▪ Les pathologies génétiques

La Polypose Adénomateuse Familiale (PAF) représente environ 1% des cancers colorectaux (**Figure 14**). C'est une maladie héréditaire rare à transmission autosomique dominante, due à une mutation du gène suppresseur de tumeur « *adenomatous polyposis coli* » (*APC*) codant une protéine APC inactive et donc incapable de limiter l'accumulation de β -caténine. Elle se caractérise par l'existence de centaines, voire de milliers, de polypes (ou adénomes) au niveau du côlon et du rectum. Le risque de développer un cancer à partir de ces adénomes est de 100%, sur une période de 10 à 20 ans, s'ils ne sont pas traités. La PAF peut apparaître chez un patient atteint de cette maladie dès l'âge de 10 ans, 80 % des individus présentent cette polypose à l'âge de 25 ans.

Le cancer colorectal héréditaire sans polypose (HNPCC pour Hereditary Non Polyposis Colorectal Cancer) ou syndrome de Lynch est la forme la plus fréquente des cancers colorectaux familiaux : 2 à 4% des CCR (**Figure 14**). Il s'agit d'une maladie génétique, autosomique dominante liée à la mutation dans un gène de réparation de l'ADN [177,178], notamment les gènes hMSH2, hMLH1 et hMSH6 [179–181], qui sont impliqués dans la réparation des mésappariements générés lors de la réplication. Ce défaut de réparation de l'ADN s'exprime sur des séquences courtes et répétées de 1 à 5 nucléotides appelées séquences microsatellites. Ces séquences étant répétées un grand nombre de fois, leur réplication comporte un risque d'erreur important : c'est l'instabilité microsatellitaire (MSI). Dans le CCR, 5 séquences microsatellitaires spécifiques sont susceptibles d'être mutées. L'expression clinique du syndrome HNPCC apparaît lorsque 2 séquences ou plus sont mutées. Contrairement à la PAF qui n'affecte que le tube digestif, le syndrome HNPCC prédispose à d'autres cancers (endomètre, grêle, voies urinaires, estomac, pancréas, voies biliaires,

glioblastome, adénome sébacé, kératoacanthome).

La polypose juvénile familiale est une maladie à transmission autosomique dominante, qui se caractérise par la présence de 50 à 200 polypes siégeant majoritairement dans le côlon et le rectum (98%), plus rarement dans l'estomac, le duodénum et l'intestin grêle [182]. Le risque de dégénérescence de ces polypes est d'environ 50% pour le côlon et le rectum [183]. S'y associent d'autres anomalies extra-digestives. La polypose juvénile peut résulter de l'altération de différents gènes. Les mutations germinales du gène suppresseur de tumeur codant SMAD4, une protéine clé de la voie de signalisation du TGF- β , sont retrouvées dans environ 20 à 50 % des familles atteintes de polypose juvénile [183,184]. D'autres gènes (*PTEN*, *BMPR1A/ALK3*) ont été impliqués dans la polypose juvénile familiale, mais leur rôle reste controversée [185].

▪ Les pathologies coliques prédisposantes

Il a été montré que 1 à 2% des cancers colorectaux seraient consécutifs à des maladie inflammatoire chronique de l'intestin [6]. Les deux principales maladies incriminées sont la maladie de Crohn et la rectocolite hémorragique [186,187] (**Figure 14**). Ces maladies créent un contexte inflammatoire chronique, décrit pour être responsable de dysplasie et favoriser la carcinogenèse [188,189].

La maladie de Crohn (MC) est une affection inflammatoire chronique d'origine multifactorielle qui peut atteindre de façon segmentaire le plus souvent l'iléon, le côlon et l'anus. Le risque de cancer colorectal chez les patients atteints de maladie de Crohn est multiplié par 3. Elle touche 5% des cas à 20 ans [190].

La rectocolite hémorragique est une affection inflammatoire évoluant par poussées qui touche exclusivement les muqueuses du côlon et du rectum. Elle atteint constamment le rectum et s'étend de manière continue plus ou moins haut vers le cæcum, respectant le grêle. Le risque de cancer colorectal est évalué à 2% après 10 ans d'évolution, 8% après 20 ans et 18% après 30 ans. Ce risque est multiplié par 3 en cas d'atteinte du côlon distal, et par 15 en cas d'atteinte généralisée du côlon.

➤ Les facteurs exogènes

L'hypothèse d'une composante environnementale dans le cancer du côlon est suggérée par les différences d'incidence entre pays occidentaux et d'Afrique. Les migrants de pays de faible incidence présentent le même niveau de risque du pays d'accueil dès la première génération.

En 2000, Fabre et collaborateurs évoquent un rôle majeur joué par l'alimentation [191]. La consommation excessive de viande rouge et de charcuterie constitue un facteur de risque [192]. Les fruits auraient au contraire un rôle protecteur. La consommation de légumes, de poissons et de volailles n'a en revanche aucun effet sur l'incidence du cancer colorectal [193].

Le tabac, la consommation excessive d'alcool et l'obésité sont des facteurs de risque. Une activité physique régulière est au contraire un facteur protecteur [193].

Enfin, le diabète de type 2 est un facteur de risque, surtout pour les cancers du côlon gauche [193].

3.B.b. Développement du cancer colorectal

La carcinogenèse colorectale débute à partir d'une cellule épithéliale intestinale qui, sous l'effet progressif de mutations, acquiert des atypies cyto-nucléaires (adénome dysplasique), puis échappe progressivement aux mécanismes de contrôle de la prolifération cellulaire et devient une cellule cancéreuse. Cette longue séquence, depuis la première cellule dysplasique jusqu'au cancer invasif, a été baptisée séquence adénome-cancer par Fearon et Vogelstein, et s'étend sur 10 à 15 ans d'évolution [5]. La prolifération tumorale cancéreuse reste d'abord confinée à l'épithélium (cancer *in situ*), puis s'organise progressivement et acquiert une néo-vascularisation qui lui permet de s'étendre en profondeur dans la paroi colique. Lorsqu'elle atteint la sous-muqueuse, elle peut s'étendre vers les ganglions lymphatiques régionaux le long des axes de drainage vasculaire. Au fur et à mesure de son évolution, des cellules tumorales peuvent se détacher, migrer dans la circulation générale, et être le point de départ de métastases viscérales.

Les foyers de cryptes aberrantes (FCA) sont les premières anomalies décelables au sein de la muqueuse colique [194]. Ils sont marqués par un aspect élargi des cryptes avec une

augmentation de l'espace intercryptique. Cette lésion ne comporte pas d'atypies cellulaires et reste donc bénigne. Cependant, lorsqu'un contingent d'atypies cellulaires est présent, le FCA devient alors dysplasique et définit ainsi le micro-adénome, lésion initiale de la séquence adénome-cancer.

Les adénomes coliques sont des proliférations intra-épithéliales de cellules néoplasiques. Les adénomes se présentent dans plus de 90% des cas sous forme d'excroissances muqueuses appelées polypes dont la forme et la taille peuvent varier (un millimètre à plusieurs centimètres) [195]. La prévalence des adénomes augmente avec l'âge et atteint 30% chez des sujets de plus de 65 ans [1]. 2,5% des adénomes évoluent dans un délai de 10 à 20 ans en adénocarcinomes, correspondant à la définition histologique du cancer colorectal. L'adénocarcinome est d'abord limité à l'épithélium colique, puis s'étend progressivement en profondeur. A ce stade de l'évolution, l'adénocarcinome colique infiltre la paroi colique riche en vaisseaux sanguins et lymphatiques. Il peut ainsi s'étendre aux ganglions et aux organes de voisinage, constituant ainsi des métastases.

L'extension ganglionnaire est possible dès que le cancer dépasse la *lamina propria* de la muqueuse colique. Le cancer devient alors plus avancé, et les chances de guérison diminuent. L'extension métastatique aux autres organes se fait par voie sanguine. Les métastases sont hépatiques dans 75% des cas, pulmonaires dans 15%, osseuses dans 5% et cérébrales dans 5% des cas.

➤ **La carcinogenèse colorectale**

La carcinogenèse regroupe l'ensemble des mécanismes qui conduisent une cellule normale à devenir cancéreuse. Cette transformation requiert, pour une cellule donnée, l'acquisition de 6 à 10 événements mutationnels clonaux. Ces mutations, germinales ou somatiques, intéressent les oncogènes, les gènes suppresseurs de tumeur et les gènes mutateurs. Elles conduisent à des modifications des voies de signalisation clés de l'homéostasie cellulaire (cycle cellulaire, réplication et réparation de l'ADN, sénescence cellulaire et apoptose), et provoquent l'instabilité génomique favorisant la survenue d'autres mutations.

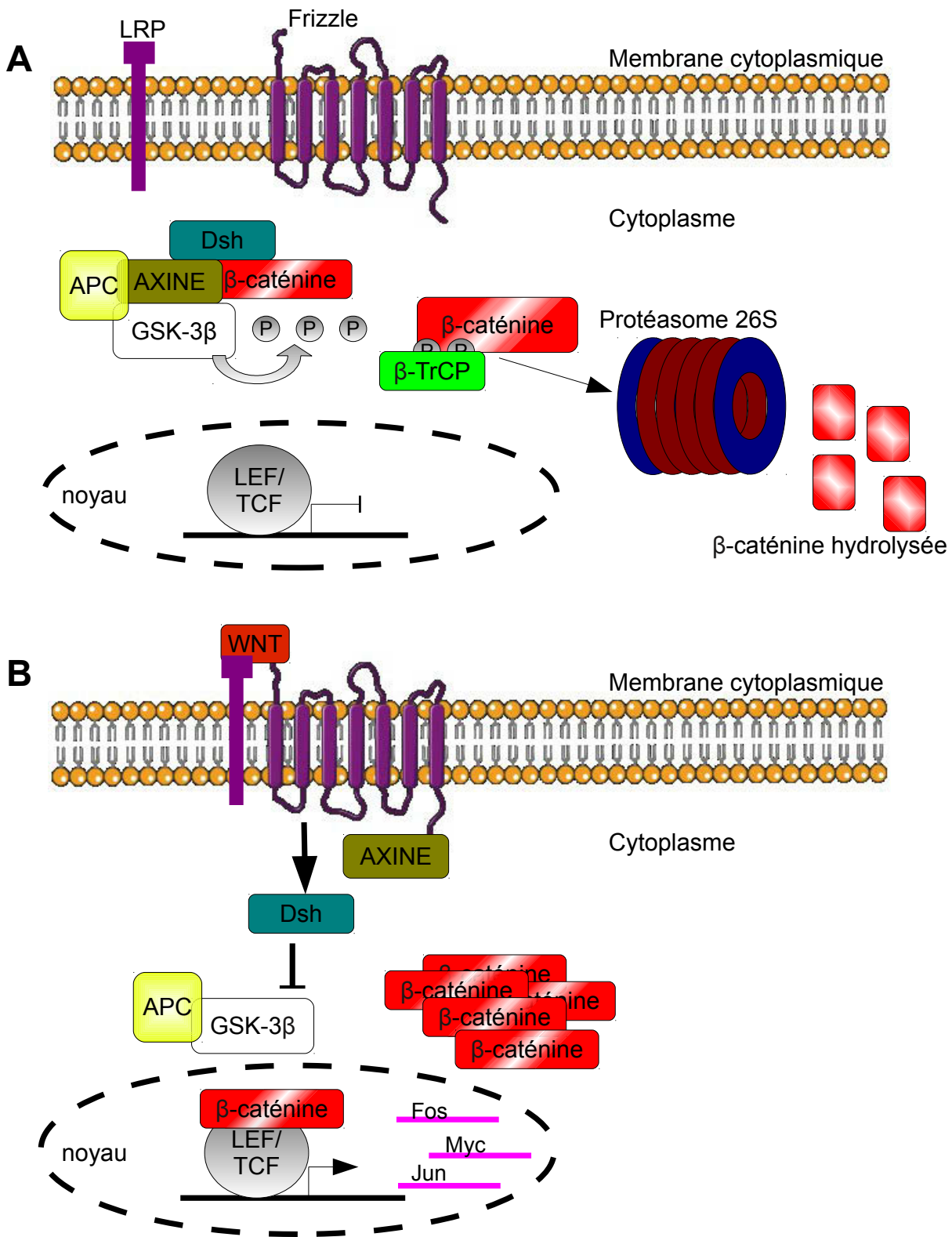


Figure 15 : la voie de signalisation WNT/APC/β-caténine. En l'absence de ligand Wnt (A), la β-caténine est associée au complexe de destruction composé de l'AXINE, Dishevelled, APC et GSK3β. GSK3β phosphoryle l'extrémité N-terminale de la β-caténine, permettant sa reconnaissance par la E3-ubiquitine ligase β-TrCP. Cette dernière conduit à l'ubiquitination de la β-caténine et sa dégradation par le protéasome. En présence de ligand Wnt (B), la β-caténine sort de ce complexe, est transloquée dans le noyau où elle s'associe aux facteurs TCF-LEF pour activer la transcription de gènes cibles (Fos, Jun, Myc, ...). β-TrCP : Beta-transducing repeat-containing protein ; APC : adenomatous polyposis coli ; Dsh : Dishevelled ; GSK-3β : Glycogen synthase kinase 3 beta ; LEF : Lymphoid enhancer-binding factor ; LRP : LDL Receptor Related Protein ; TCF : Transcription factor. D'après Bienz et Clevers, [208].

➤ Les voies de signalisation impliquées dans le cancer colorectal

▪ Le gène suppresseur de tumeur p53

Le gène *TP53* est localisé sur le bras court du chromosome 17 et code pour une protéine nucléaire de 53 kDa, p53. En réponse aux différents stress cellulaires de type intrinsèques (ex. activation de voies oncogéniques) ou extrinsèques (ex. agents génotoxiques), l'augmentation de l'expression de p53 stimule l'expression de nombreux gènes impliqués dans le maintien de l'intégrité du génome, la progression du cycle cellulaire, l'apoptose et la sénescence [196–198]. Sa mutation est très fréquente dans le cancer colorectal.

▪ La voie Wnt/APC/ β -caténine

La voie Wnt/APC/ β -caténine est une des voies de signalisation majeure dans le processus de carcinogenèse (**Figure 15**) [45]. Dans le cas du processus normal d'activation, les récepteurs *frizzle* sont activés par la fixation du ligand Wnt, qui conduit à la phosphorylation de la protéine Dishevelled. Cette protéine cytoplasmique active alors la voie Jun kinase (JNK) et dissocie le complexe APC-Axine-GSK3 en se fixant à l'Axine. La dissociation de ce complexe abolit son activité kinase sur la β -caténine. Alors que la forme phosphorylée de la β -caténine est ubiquitinée par β -TrCP et dégradée par le protéasome 26S [199], la forme non phosphorylée est transportée au noyau où elle s'associe aux facteurs de transcription TCF/LEF (TCF-4 exprimé dans le côlon) pour activer la transcription de gènes cibles, qui favorisent la prolifération cellulaire [200]. Parmi ces gènes, on dénombre : *c-myc* connu pour réguler la prolifération cellulaire, l'apoptose et la sénescence [201–204], la cycline D1 qui joue un rôle central dans la régulation du cycle cellulaire [127], le récepteur PPAR δ (« Peroxisome Proliferator-Activated Receptor delta ») impliqué dans l'inflammation et l'apoptose [205]. La β -caténine accumulée dans le cytoplasme s'associe à l'actine ainsi qu'à la E-cadherine, une glycoprotéine du cytosquelette, et ainsi altère l'adhésion cellulaire. Par ailleurs, la β -caténine libérée du complexe de destruction APC-Axine-GSK3 β peut s'associer aux microtubules et ainsi agir sur la migration cellulaire et la stabilité chromosomique en contrôlant l'organisation des kinétochores (qui lient les chromosomes aux microtubules) et des chromosomes pendant la métaphase [206].

Dans le cancer colorectal, la voie Wnt/APC/ β -caténine peut être activée en absence du ligand Wnt par la mutation de gènes codant pour les protéines de cette voie. Le plus souvent, il

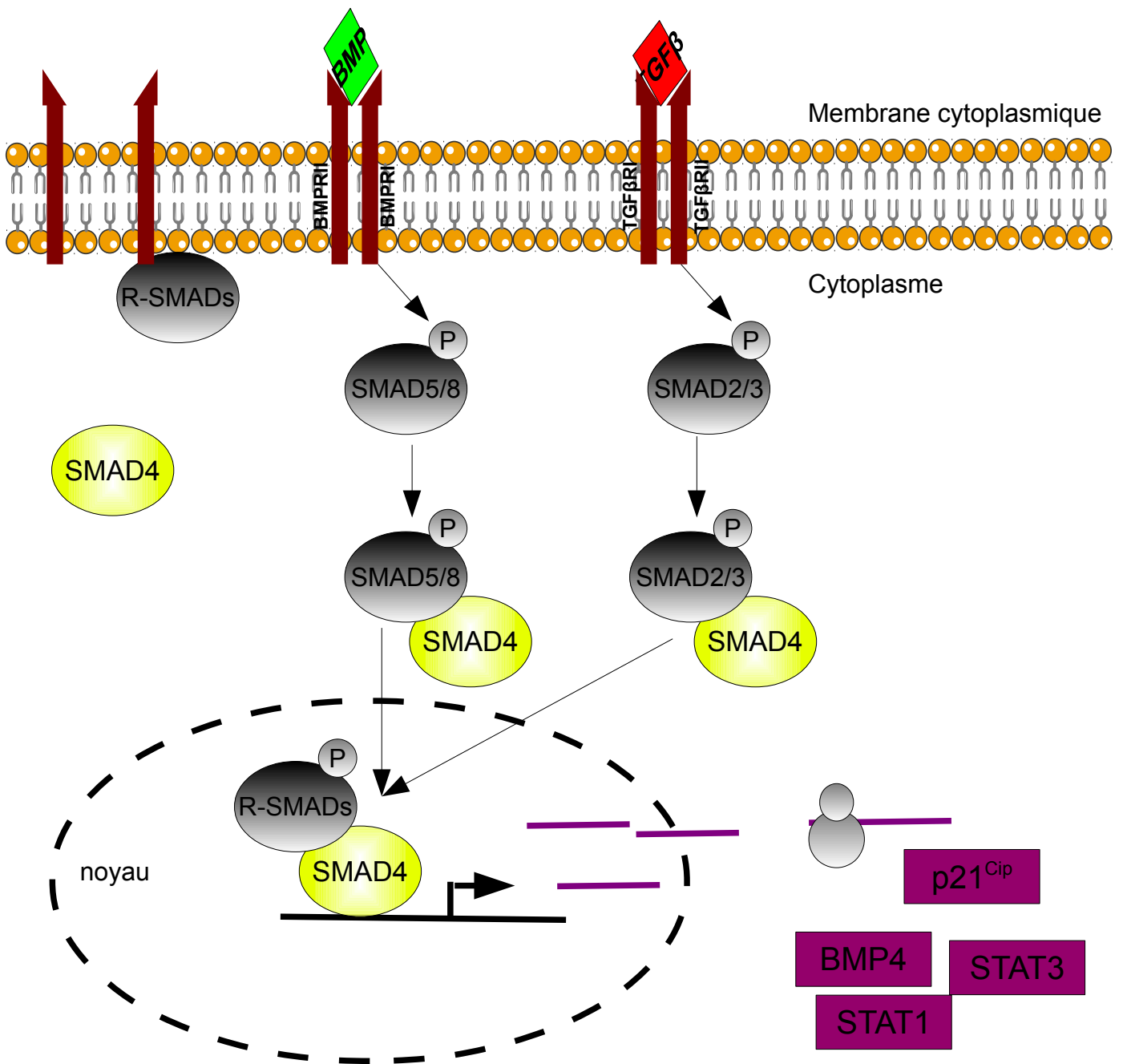


Figure 16 : la voie de signalisation TGFβ/SMAD. En se fixant à leurs récepteurs, les membres de la famille du TGFβ (BMP, TGFβ, ...) entraînent la phosphorylation des R-SMAD grâce à l'activité sérine/thréonine kinase du récepteur. R-SMAD correspond aux SMAD 2 et 3 dans la voie TGFβ/Activin et SMAD 1, 5 et 8 en réponse à la fixation de BMP. Les protéines SMAD ainsi phosphorylées peuvent établir des complexes avec SMAD4. Ces complexes migrent dans le noyau et stimulent la transcription de nombreux gènes codant des protéines impliquées dans la différenciation cellulaire (BMP4,...), l'apoptose (STAT1,...), l'angiogenèse (STAT3,...) et le blocage du cycle cellulaire (p21^{Cip},...). BMP : Bone morphogenetic proteins ; BMPRI : Bone morphogenetic protein receptor I ; BMPRII : Bone morphogenetic protein receptor II ; SMAD : Mothers against decapentaplegic homolog ; STAT : Signal transducer and activator of transcription ; TGFβ : Transforming growth factor beta ; TGFβRI : Transforming growth factor beta receptor I ; TGFβRII : Transforming growth factor beta receptor II. D'après Heldin et collaborateurs, [53].

s'agit de l'inactivation par mutation du gène APC induisant la synthèse d'une protéine APC tronquée incapable de former un complexe APC-Axine-GSK3 actif [207,208]. Bien que l'activation de la voie Wnt/APC/ β -caténine soit retrouvée dans de nombreux cancers, les mutations du gène *apc* semblent spécifiques du cancer colorectal [209,210]. La protéine APC est cependant intacte dans 15% des cancers colorectaux sporadiques. D'autres mutations affectant la voie Wnt/APC/ β -caténine expliquent alors son activation notamment au niveau de séquences répétées du gène de l'Axine et dans l'exon 3 du gène codant pour la β -caténine (*CTNNB1*) au niveau des sites de phosphorylation [211], ou dans le gène codant TCF-4 [212]. L'altération de la voie WNT est indispensable à la carcinogenèse colorectale, cependant elle n'est pas suffisante pour la transformation en cellules malignes.

▪ La voie TGF- β /SMAD

La superfamille du TGF- β (« Tumor Growth Factor beta ») comprend différents membres dont les BMPs (Bone Morphogenic Proteins), les activines et le TGF- β . Leurs effets dépendent de l'intensité et de la durée des signaux induits [213] et une de leurs fonctions essentielles chez l'adulte, en particulier au niveau de l'épithélium digestif, est une inhibition du cycle cellulaire en phase G1 [49].

La liaison de ces morphogènes (ex. TGF- β) à leur récepteur (ex. TGFBRII) permet l'activation du récepteur par phosphorylation. En fonction du couple ligand-récepteur, le récepteur activé phosphoryle des protéines R-SMAD (SMAD 2 et 3 / SMAD 1, 5 et 8) et les phospho-R-SMAD s'associent à SMAD4. Les complexes formés sont transloqués au noyau où ils induisent la transcription de gènes cibles dont des inhibiteurs majeurs du cycle cellulaire comme les CDKi p15^{Ink4B}, p16^{Ink4A} et p21^{Cip} (**Figure 16**). Dans les cancers colorectaux, on retrouve des mutations inactivatrices du récepteur TGFBRII ou du récepteur ACVR2 des activines [214].

▪ La voie RAS/MAPK

La voie Ras/MAPK (Ras/Raf/MEK/Erk) transmet des signaux mitogènes *via* des récepteurs à tyrosine kinase comme EGFR ou VEGFR (**Figure 17**). Cette voie est importante pour la croissance et la prolifération des cellules tumorales mais également pour l'angiogenèse [215].

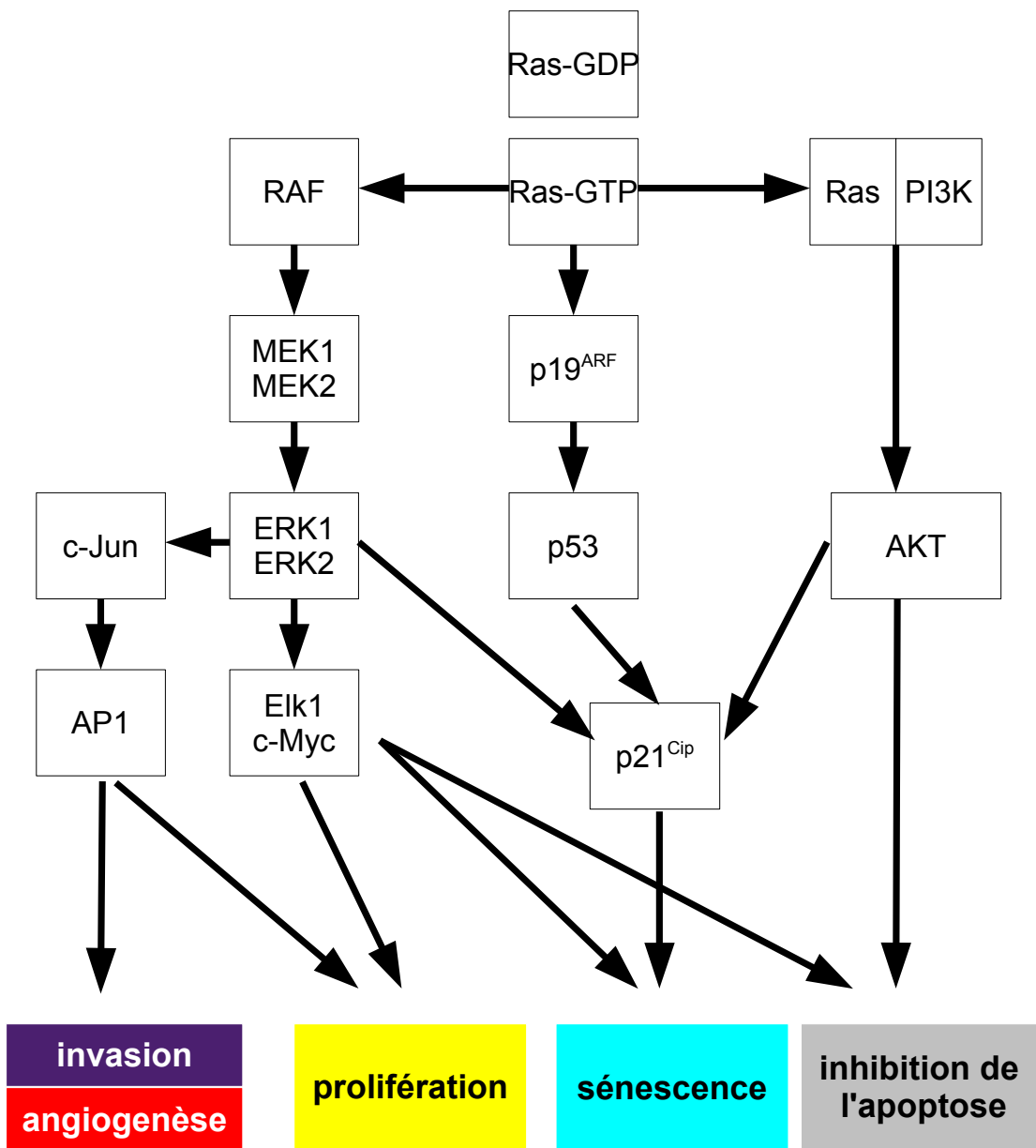


Figure 17 : la signalisation cellulaire associée à RAS. L'activation de RAS peut conduire à l'activation de trois principales voies de signalisation aboutissant à la sénescence, la prolifération, l'invasion, l'angiogenèse et l'inhibition de l'apoptose. AKT : Protein Kinase B ; AP-1 : activator protein 1 ; ELK1 : E twenty-six-like transcription factor 1 ; ERK : extracellular signal-regulated kinases ; MEK : mitogen-activated protein kinase kinases ; PI3K : Phosphatidylinositol 3-kinases ; RAF : Rapidly Accelerated Fibrosarcoma. D'après Malumbres et Barbacid, [172].

Chez l'Homme, il existe 3 protéines RAS : H-RAS, K-RAS et N-RAS appartenant toutes à la superfamille des petites protéines G. Dans le cancer colorectal, on retrouve des mutations de *K-ras* qui entraînent l'activation constitutive de la voie. L'activation de la voie RAS/MAPK peut également être associée à une surexpression de récepteurs de facteurs de croissance, comme l'EGFR qui est surexprimée dans 30 à 70% des cancers colorectaux [216].

Parmi les gènes activés par la voie RAS/MAPK, on dénombre le gène codant p19^{ARF} qui va inhiber la E3-ubiquitine ligase Mdm2 impliquée dans la dégradation de p53. Ainsi, la voie RAS/MAPK favorise un arrêt prolongé du cycle cellulaire et la sénescence cellulaire [139]. Comme dans les cancers colorectaux, la protéine p53 est fréquemment inactive, l'activation de la voie induit au contraire à une prolifération accélérée [217]. La cascade Raf/MEK/ERK joue également un rôle important dans la survie cellulaire *via* l'induction de signaux anti-apoptotiques, en particulier par phosphorylation de la protéine pro-apoptotique Bad, qui est alors séquestrée par les protéines de la famille 14-3-3.

▪ La voie PI3K/Akt

La voie PI(3)K/Akt transmet, comme la voie Ras/MAPK, des signaux provenant de nombreux récepteurs à tyrosine kinase (**Figure 18**). Les mutations de différents composants de cette voie sont très fréquentes dans la majorité des cancers et leur caractère oncogénique est démontré. Dans les cancers colorectaux, la plupart des mutations affectent le gène PIK3CA codant, la sous-unité catalytique alpha de la PI(3)Kinase (p110). Cependant, 40% des cancers colorectaux sporadiques présentent une mutation sur un autre acteur de la voie, dont PTEN, Akt2 ou p85 [218]. Ces mutations sont généralement exclusives, suggérant qu'une atteinte unique de la voie est suffisante pour procurer un avantage de sélection.

La kinase Akt possède de très nombreux substrats inducteurs de la survie cellulaire ou ayant un rôle mitogénique (**Figure 18**). Par exemple, phospho-Mdm2, transloquée au noyau, assure l'ubiquitination de p53 afin de promouvoir sa dégradation par le protéasome [147] et phospho-Bad, séquestrée par 14-3-3, inhibe l'apoptose [219].

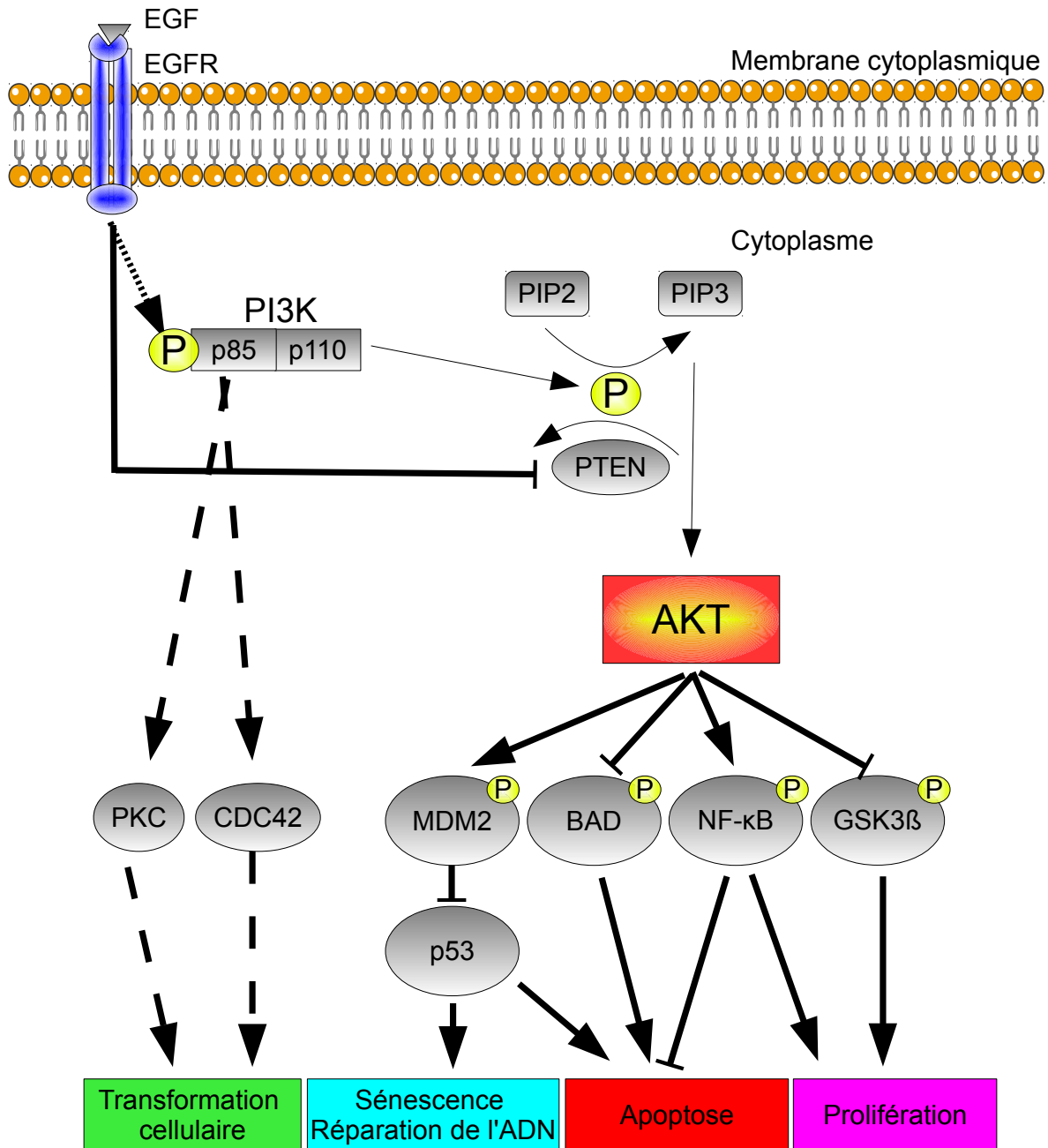


Figure 18 : la voie de signalisation PI3K/AKT. En se fixant à son récepteur, EGF entraîne la phosphorylation de la sous-unité p85 de la PI3K et réprime la phosphatase PTEN. La PI3K peut alors convertir le PIP2 en PIP3 conduisant à l'activation de la kinase AKT, qui va moduler l'activité de nombreux effecteurs : MDM2, NF-κB, BAD, GSK3β, ... Ces modifications vont favoriser la prolifération cellulaire, inhiber les voies de signalisation associées à l'apoptose et à la sénescence. L'activation de la sous-unité p85 de la PI3K peut également favoriser la transformation cellulaire *via* l'activation des protéines PKC et CDC42. AKT : Protein Kinase B ; BAD : Bcl-2-associated death promoter ; CDC42 : Cell division control protein 42 homolog ; EGF : epithelial growth factor ; EGFR : epithelial growth factor receptor ; NF-κB : nuclear factor kappa B ; PI3K : phospho-inositide-3-kinase ; PIP2 : Phosphatidyl inositol diphosphate ; PIP3 : Phosphatidyl inositol triphosphate ; PKC : Protein Kinase C ; PTEN : Phosphatase and tensin homolog ; GSK3β : Glycogen synthase kinase 3 beta ; Ub : Ubiquitine. D'après Cristofalo et collaborateurs, [142].

3.B.c. L'instabilité génomique

Les systèmes de réparation protègent les cellules des mutations induites par les agents mutagènes. La mort cellulaire, la sénescence cellulaire ou encore le système immunitaire tendent à limiter l'expansion des cellules anormales, si bien qu'il paraît impossible pour une personne donnée d'acquérir au cours de la vie autant de mutations sur un même clone cellulaire [220]. C'est ainsi qu'a été introduite la notion d'environnement cellulaire permissif ou instabilité génomique [221], qui fait appel à deux grandes voies de la carcinogenèse : l'instabilité chromosomique (CIN) et l'instabilité microsatellitaire (MIN). L'instabilité génomique augmente la probabilité d'obtenir des mutations successives sur un même clone, et rend possible la transformation d'une cellule normale en cellule cancéreuse, malgré les systèmes de protection.

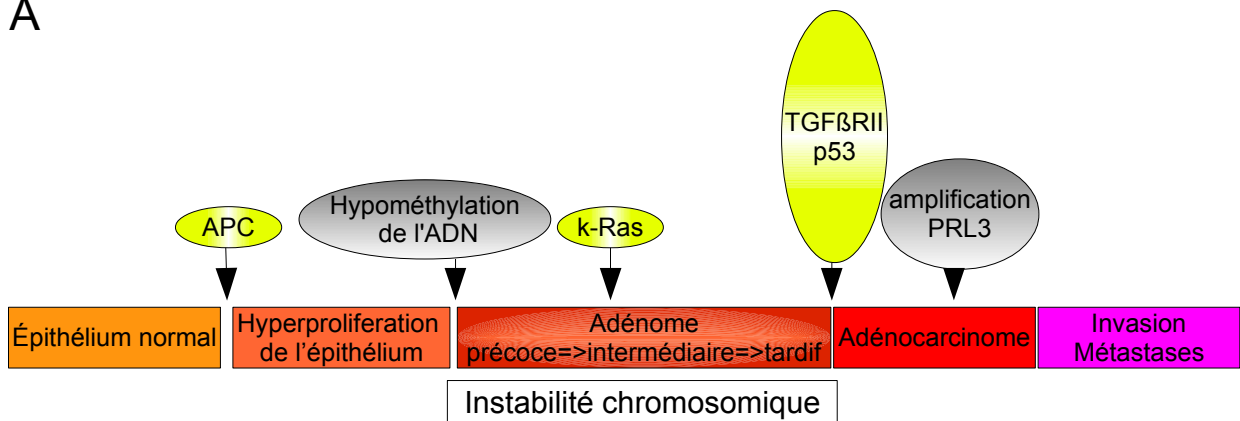
➤ L'instabilité chromosomique

L'instabilité chromosomique concerne 70 à 85% des cancers colorectaux. Elle est donc appelée « voie traditionnelle » ou encore « voie suppressive » car elle se caractérise par la perte de matériel chromosomique entraînant l'inactivation de gènes suppresseurs de tumeurs [222]. L'instabilité chromosomique conduit au phénotype LOH+ (« *loss of heterozygosity* ») et se caractérise par une aneuploïdie. La séquence adénome-cancer est le support clinique de cette voie de carcinogenèse. Selon ce modèle proposé par Fearon et Vogelstein, à chaque étape intervient une mutation qui s'ajoute à la précédente et induit une modification phénotypique : mutation du gène *APC*, du gène *K-ras*, des gènes *SMAD4* ou *DCC* et du gène *TP53* [184,213] (**Figure 19**). Le défaut d'une mutation stoppe l'expression phénotypique de la séquence qui ne se poursuit pas, même si les mutations suivantes s'opèrent. Cependant, seule une petite proportion de cancer colorectal possède l'ensemble des mutations, ce qui tend à montrer que d'autres événements génétiques interviennent et court-circuitent la voie CIN pour délivrer le message biologique nécessaire à la poursuite de la séquence [223].

➤ L'instabilité des microsatellites

L'instabilité microsatellitaire est associée au phénotype MSI+ (*MicroSatellite Instability*) typique du syndrome HNPCC. Elle est retrouvée dans 15 à 20% des cancers colorectaux. C'est la « voie mutationnelle ». Il s'agit de mutations sur des séquences courtes répétées appelées séquences microsatellites. Ces mutations sont normalement réparées par les systèmes de

A



B

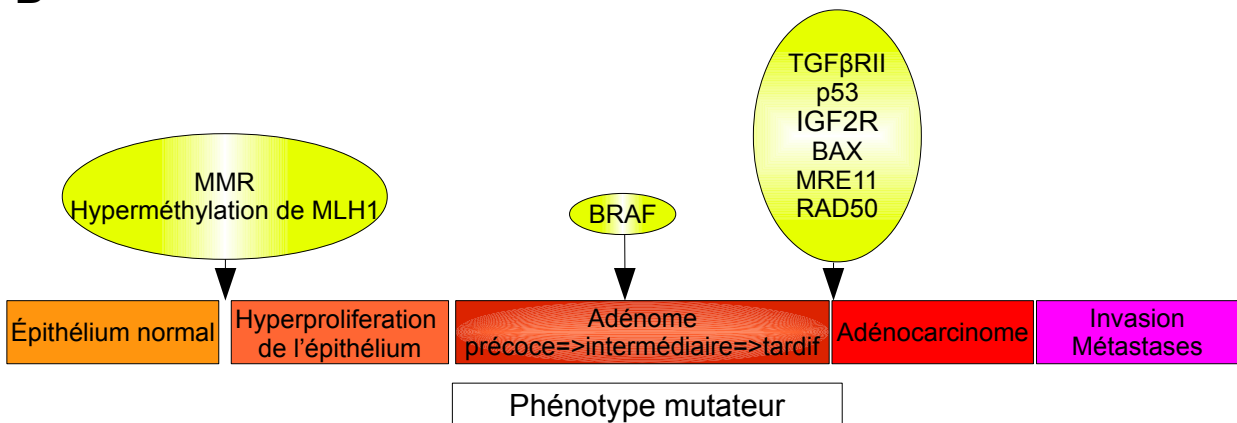


Figure 19 : les modèles de progression du cancer colorectal. Séquence d'apparition des altérations génétiques et épigénétiques lors de la progression du cancer colorectal dans le modèle d'instabilité chromosomique (A) ou d'instabilité microsatellitaire (B). ● : mutation ou modification épigénétique ; APC : Adenomatous Polyposis Coli ; BAX : Bcl-2-associated X protein ; BRAF : B-RAF ; IGF2R : Insulin-like Growth Factor 2 Receptor ; MLH1 : MutL homolog 1 ; MMR : mismatch repair ; MRE11 : meiotic recombination 11 homolog ; MSS : microsatellite stable ; MSI : microsatellite instable ; PRL3 : Phosphatase of regenerating liver 3 ; RAF : Rapidly Accelerated Fibrosarcoma ; TGFβRII : Transforming Growth Factor beta receptor II. D'après Vilar et Gruber [225]

réparation de l'ADN *via* les protéines du système MMR (« *MisMatch Repair* »). Or, dans le syndrome HNPCC, le malade est porteur de mutations germinales dans au moins un des 7 gènes codant les effecteurs du système MMR [178–180]. Dans 15% des cancers colorectaux sporadiques, il y a une hyperméthylation du promoteur du gène hMLH1, mais pas de mutation [224]. Cette altération épigénétique conduit à une inactivation de la transcription de ce gène et donc au phénotype MSI+ (**Figure 19**).

Dans le cancer colorectal sporadique, 5 séquences microsatellites (BAT25, BAT26, D5S346, D2S123 et D17S250) sont affectées. Ces séquences sont retrouvées dans des gènes impliqués dans le cancer colorectal comme les gènes codant le récepteur TGF- β RII, la protéine pro-apoptotique Bax, la caspase 5, la β -caténine, la protéine APC, le facteur de croissance IGF-II, ou encore MSH3 et MSH6 (**Figure 19**). Selon le nombre de séquences microsatellites touchées, 3 phénotypes sont décrits [225] :

- phénotype MSI-H (*High*) lorsque 2 séquences microsatellites ou plus sont mutées : ce phénotype est rencontré dans pratiquement 100% des cancers colorectaux du syndrome HNPCC. hMLH1 et hMSH2 sont les 2 gènes les plus fréquemment mutés (90% des HNPCC) [225]. De même 50% des adénomes ont le phénotype, ce qui montre une implication plutôt précoce dans la séquence adénome-cancer. Dans le cancer colorectal sporadique, le phénotype MSI-H est rencontré dans 15% des cas. Les mutations somatiques de hMLH1 et hMSH2 sont peu fréquentes. Cependant, la perte d'expression de ces gènes est retrouvée dans 95% des cancers colorectaux sporadiques. En revanche, le phénotype MSI n'est retrouvé que dans 1,8 à 2,5% des adénomes sporadiques, ce qui sous-entend que cette voie de carcinogenèse intervient très tardivement dans la séquence adénome-cancer du cancer colorectal sporadique [225].
- phénotype MSI-L (*Low*) lorsqu'une seule séquence microsatellite est mutée : ce phénotype peut être rencontré dans le syndrome HNPCC, par mutation du gène hMSH6 [225]. Deux mécanismes sont avancés pour expliquer ce phénotype dans le cancer colorectal sporadique : une altération du gène MGMT (O-6-Méthylguanine DNA Methyltransferase) du système MMR et une perte d'expression du gène hMLH1 suite à une méthylation aberrante.
- phénotype MSS (*Stable*) lorsqu'aucune séquence microsatellite n'est mutée. Ce phénotype est le plus fréquent dans la voie mutationnelle. Il représente 54 à 78% des

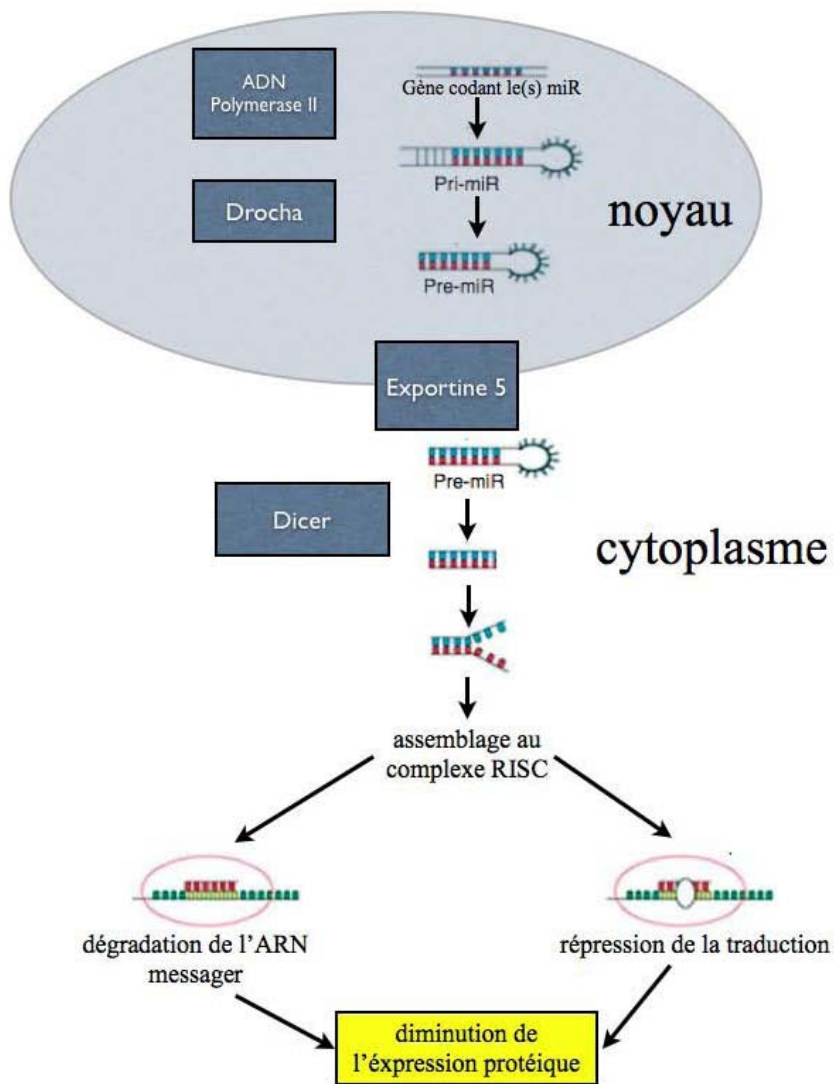
cas [225].

3.B.d. Les altérations épigénétiques

Les modifications épigénétiques sont des altérations qui modifient l'expression de gènes sans altérer la séquence de l'ADN. Elles peuvent néanmoins se transmettre au cours de la division cellulaire, mais sont réversibles. Trois types de modifications épigénétiques affectant la transcription sont rencontrés : l'acétylation, la méthylation et l'ubiquitination des histones, et la méthylation de l'ADN. Une modification de la méthylation de l'ADN est fréquemment observé dans le cancer colorectal.

La méthylation de l'ADN chez les mammifères s'effectue sur les di-nucléotides CpG par l'adjonction d'un groupement méthyl (-CH₃) au niveau du carbone 5 d'une cytosine. Cette modification est catalysée par les DNA méthyltransferases DNMT1, DNMT3a et DNMT3b [226]. Elles maintiennent l'empreinte de méthylation du génome après chaque réplication lors de la phase S du cycle cellulaire. Les dinucléotides CpG sont globalement sous-représentés dans le génome humain, hormis dans de courtes régions appelées îlots CpG situées au niveau du promoteur et du 1^{er} exon des gènes [227]. Leur hyperméthylation entraîne une extinction transcriptionnelle par inhibition de la liaison des facteurs de transcription, et par l'acétylation d'histones [228]. Ce processus peut entraîner l'arrêt de l'expression de gènes suppresseurs de tumeurs [229].

Dans le cancer colorectal, il existe une hypométhylation globale du génome, qui peut rendre compte d'une activation de gènes réprimés dans des conditions normales et une hyperméthylation des îlots CpG. Certains individus semblent prédisposés à l'hyperméthylation aberrante de séquences promotrices, notamment de gènes suppresseurs de tumeurs (hMLH1, p19^{ARF}, p16^{INK4A}) [227,230]. Dans les cancers, l'hyperméthylation des gènes suppresseurs correspond au phénotype « *CpG Island Methylator Phenotype* » (CIMP) [231]. Les mécanismes à l'origine de ce phénotype sont encore mal connus. Deux hypothèses expliquent cette méthylation aberrante : une surexpression des DNMT [232] ou une altération des facteurs protégeant l'ADN de la méthylation, comme les histones acétyltransférases (HAT) [233]. A l'heure actuelle, il est clairement établi que l'inactivation des gènes suppresseurs de tumeurs par méthylation est une cause majeure d'instabilité génomique et que l'association des altérations génétiques et épigénétiques est nécessaire à la carcinogenèse. Suite à ces



o
oi
oi
r oc s o o L
A t r e. o
L o e n oi
L N oi
oi

observations, des inhibiteurs des DNMTs (decitabine et 5-azaC) ont été développés [234,235].

3.B.e. Les oncomirs

Les microARN ou miR sont de petits ARN non codant de 19 à 23 acides nucléiques découverts initialement chez le nématode *Caenorhabditis elegans* [236]. Ces ARN non codant se fixent sur la portion 3' non traduite d'ARN messagers cibles. Cette fixation entraîne la formation de complexes ribonucleoprotéiques dont la spécificité est due à des interactions ARN-ARN imparfaites conduisant le plus souvent à la dégradation ou à l'altération de la traduction des ARN messagers cibles.

Les gènes de miARN sont transcrits par l'ARN polymérase III sous la forme de longs précurseurs appelés « pri-miARN » (**Figure 20**). Ces précurseurs sont clivés dans le noyau par un complexe nommé microprocesseur, formé des enzymes Drosha et DGCR8 (« Di George Critical Region 8a »). On obtient alors un produit intermédiaire appelé « pré-miARN ». Le pré-miARN est un ARN d'environ 70 nucléotides, replié en tige-boucle imparfaite par complémentarité de bases entre la première et la deuxième moitié de sa séquence. Ce pré-miARN est transporté du noyau au cytosol par un transport actif GTP-dépendant grâce à une interaction avec l'exportine 5. Le pré-miARN est ensuite clivé par une enzyme de la famille DICER qui permet l'hydrolyse de la structure boucle, pour libérer un petit ARN double-brin appelé miARNdb. Ce dernier est alors incorporé dans un complexe d'extinction appelé RISC (« *RNA-induced silencing complex* »), qui assure le passage du miARN double brin à l'état simple brin [237]. Ce complexe nucléoprotéique, peut alors reconnaître les ARN messagers cibles et deux voies de blocage de la traduction sont alors possibles, soit la dégradation de l'ARNm cible si le complexe contient la protéine Argonaute 2, soit la répression de la traduction de ce dernier si le complexe contient la protéine Argonaute 1.

Actuellement, un peu plus de 1000 miR ont été identifiés dans le génome humain. Chacun cible potentiellement de nombreux ARN messagers (dus à leur appariement imparfait avec la cible) et on considère qu'ils interviennent dans la régulation d'environ 30 % du génome humain. Si leurs fonctions précises sont encore mal connues, les microARN sont impliqués dans un grand nombre de processus physiologiques essentiels tels le métabolisme, l'apoptose, la sénescence, la croissance et la différenciation cellulaire [238]. De plus, l'expression des microARN est dérégulée dans de nombreuses maladies, notamment les maladies

inflammatoires chroniques de l'intestin et les cancers.

De nombreuses études se sont intéressées à l'expression des micro-ARN dans les cancers [239]. Les microARN dérégulés dans les cancers sont appelés oncomirs. Ils ont pour cible notamment des transcrits de gènes suppresseurs de tumeurs [240]. L'oncomir let-7 cible les produits des gènes Ras, c-Myc et CDK6. L'oncomir miR-24 régule c-Myc et E2F1. L'expression de ces deux oncomirs est diminuée dans plusieurs cancers. D'autres oncomirs sont au contraire surexprimés, par exemple mir-21, un répresseur de BCL2, de PTEN et de EGFR [239]. Dans le cancer du côlon, les oncomirs 19a, 20, 21, 34a, 92, 96, 106a, 135b, 181b, 183, 203, 210 et 222 sont sous-exprimés tandis que les oncomirs 1, 9, 30-3p, 34b, 124a, 129, 133a, 143, 145, 192, 192-2, 215, 301, 328, 342, 486 et 592 sont surexprimés [241,242]. Une étude menée par Savers et collaborateurs en 2009 a mis en évidence que le profil d'expression des microARN dépendait du stade du cancer du côlon et de la nature des mutations observées dans le système de réparation des mésappariements [242]. Bien que nous commençons à avoir une bonne représentation des miARN dérégulés dans les cancers, peu de données ont été recueillies sur les cibles de ces miR et les conséquences fonctionnelles de leurs dérégulations. Toutefois, plusieurs oncomirs sont décrits pour jouer un rôle dans la progression du cancer colorectal. Le mir-196a promeut l'apparition de métastases *via* son action sur les facteurs de transcription « homeo-box » HoxA7, HoxB8, HoxC8 et HoxD8. Leurs répressions stimulent la migration cellulaire et l'invasion [243]. Le micro-ARN 34a inhibe des facteurs de transcription de la famille E2F comme E2F1 qui est impliqué dans la progression du cycle cellulaire [239]. L'expression de ce microARN est réduite dans le cancer colorectal suite à la méthylation de son promoteur [243].

3.B.f. Carcinogenèse et inflammation

L'inflammation et l'immunité jouent un rôle important dans l'initiation et la progression du cancer. Bien que l'immunité innée et adaptative collaborent à la réduction de l'incidence des tumeurs (cet aspect ne sera pas développé dans ce travail), il est actuellement admis que l'inflammation et l'activation chronique de l'immunité innée sont globalement pro-carcinogènes [244,245]. Ces processus affectent de nombreux aspects de la carcinogenèse comme la prolifération et la survie des cellules malignes, l'angiogenèse et la formation des métastases [246,247].

➤ Les étiologies de l'inflammation associée au cancer

Les connections entre inflammation et cancer peuvent être schématisées selon deux voies : une voie intrinsèque liée aux altérations génétiques et épigénétiques des cellules néoplasiques, qui génère de l'inflammation, et une voie extrinsèque reposant sur le recrutement de cellules immunitaires [248,249].

Des études récentes ont établi un lien entre l'activation des oncogènes et l'inflammation. En plus de promouvoir la prolifération cellulaire, des oncogènes activent des voies de l'inflammation. La voie de signalisation RAS/MAPK induit la voie de l'inflammation dépendante de NF- κ B et ainsi la production de cytokines et de chimiokines pro-inflammatoires [250,251]. L'oncogène *c-myc* code un facteur de transcription qui stimule l'angiogenèse en favorisant l'expression et la sécrétion de la cytokine IL-1 β [252]. L'inactivation de gènes suppresseurs de tumeurs peut également induire la production de médiateurs de l'inflammation. L'inactivation du gène codant le récepteur de type du TGF- β stimule la production des chimiokines pro-inflammatoires CXCL5 et CXCL12 [253].

Des études épidémiologiques ont montré que de nombreuses infections chroniques prédisposent aux cancers. L'infection par *Helicobacter pylori* est associée aux cancers gastriques et aux lymphomes de la muqueuse gastrique ; des infections virales (virus des hépatites B et C, Papillomavirus) sont associées aux cancers du foie et du col de l'utérus. De même, des pathogènes intestinaux pourraient être impliqués dans le cancer colorectal (cf. ci-dessous). Les pathologies inflammatoires chroniques de l'intestin sont également des facteurs de risques bien identifiés du cancer colorectal (cf. ci-dessus), de même que des agents chimiques irritants (ex. fumée de cigarette, silice,...). L'inflammation est alors liée au recrutement de cellules immunitaires, comme les macrophages, les cellules dendritiques, les polynucléaires neutrophiles et les lymphocytes T, dans le microenvironnement tumoral. Ces leucocytes produisent des espèces oxygénées ou azotées réactives, des cytokines, des chimiokines, des facteurs de croissance, des facteurs angiogéniques, et des protéases dégradant la matrice extra-cellulaire (ex. les metalloproteinases MMP1, MMP3 et MMP9). Ces acteurs favorisent la prolifération et la migration des cellules tumorales [248,249].

En conséquence, l'inflammation, patente ou non sur le plan clinique, est une composante du cancer. Un microenvironnement inflammatoire localisé au niveau tumorale est toujours présent et affecte la carcinogenèse. Plusieurs effecteurs pro-inflammatoire jouant un rôle dans

la carcinogenèse ont été décrits.

➤ **NF-κB et carcinogenèse**

Plusieurs cytokines pro-inflammatoires comme le TNF- α , l'IL-1, l'IL-6 et l'IL-8 (CXCL8), connues pour être associées à la progression tumorale, sont codées par des gènes cibles de la voie NF-κB et parallèlement de nombreux oncogènes activent la voie NF-κB [247,248]. De plus, un lien direct a été établi entre une activation de la voie NF-κB et la promotion de tumeurs chez la souris et chez l'Homme. [247,254,255].

L'activation de NF-κB augmente la transcription de 2 types de gène : des gènes de l'inflammation et de l'immunité innée pro-inflammatoire (COX2, iNOS, TNF, cytokines, chimiokines et protéases) et des gènes impliqués dans la survie des cellules (BCL-XL, cIAPs, GADD45 β , BFL1 et SOD2) [254–256]. Dans les cellules de l'inflammation, l'activation de NF-κB dépendante de IKK- β favorise le développement de tumeurs en induisant l'expression des gènes codant des cytokines pro-inflammatoire, des facteurs de croissance et de survie, qui vont agir sur les cellules malignes ou de l'environnement tumoral (fibroblastes). Dans les cellules en cours de transformation ou déjà malignes, l'activation de NF-κB conduit surtout à l'expression de gènes codant des protéines anti-apoptiques comme GADD45 β (DNA-damage-inducible 45 β), BCL-XL (B-cell lymphoma XL) et BFL1 (B-cell-lymphoma-2-related protein) [256]. L'activation de la voie NF-κB contrôle donc à la fois la survie des cellules transformées et l'inflammation conduite par les cellules de l'immunité. Ces dernières produisent des facteurs sécrétés soutenant la croissance tumorale.

➤ **Il-6 et carcinogenèse**

L'interleukine IL-6 occupe une place particulière dans la carcinogenèse colorectale associée à l'inflammation. L'IL-6 est un acteur majeur dans la transition entre l'inflammation aiguë et l'inflammation chronique. L'IL-6 se fixe à ses récepteurs (IL-6R α) qui interagissent avec l'adaptateur membranaire gp130. En conséquence, il active les tyrosines kinases Janus (JAK), qui phosphorylent l'effecteur STAT3. STAT3 phosphorylé forme un dimère, qui migre dans le noyau pour activer la transcription des gènes cibles [257]. Dans le cancer colorectal, STAT3 induit ainsi l'expression de gènes importants pour la prolifération cellulaire (cycline D et PCNA) et la suppression de l'apoptose (Bcl-XL, Bcl-2 and Mcl-1).

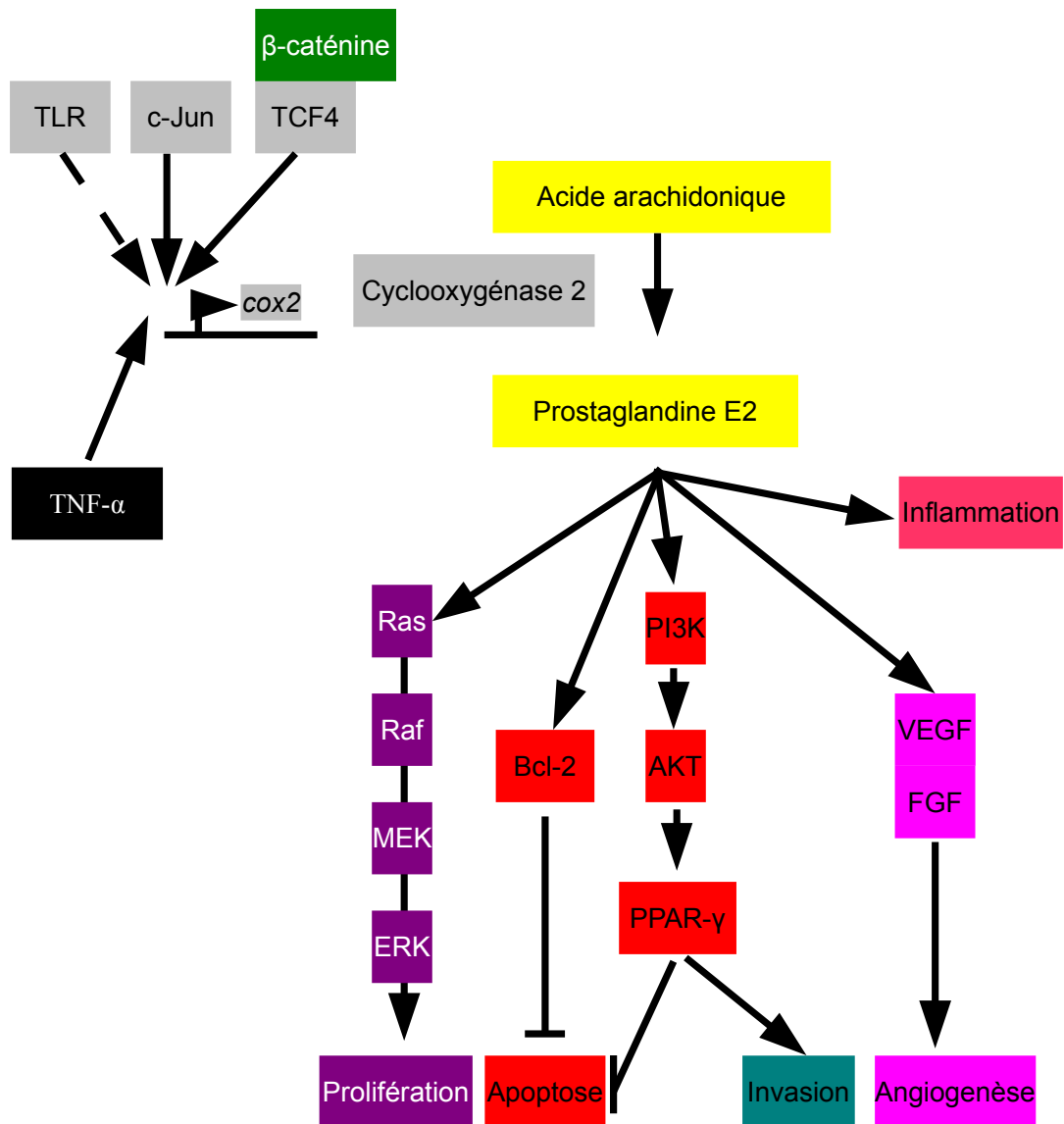


Figure 21 : la voie de signalisation COX-2. La transcription du gène codant le cyclooxygénase (cox-2) est régulé par l'activation des TLRs, les facteurs de transcription TCF/ β -caténine ou par la présence de TNF- α dans le milieu extra-cellulaire. COX-2 est une enzyme responsable de la conversion de l'acide arachidonique en prostaglandine E2. Les prostaglandines ainsi produites vont favoriser l'inflammation, la prolifération cellulaire, l'apoptose, l'invasion et l'angiogénèse. AKT : Protein Kinase B ; Bcl-2 : B-cell lymphoma 2 ; ERKs : extracellular signal-regulated kinases ; FGF : Fibroblast Growth Factor ; MEKs : mitogen-activated protein kinase kinases ; PI3K : phospho-inositide-3-kinase ; PPAR- γ : peroxisome proliferator-activated receptor gamma ; RAF : Rapidly Accelerated Fibrosarcoma ; TLR : Toll-Like receptors ; TNF- α : Tumor Necrosis Factor alpha ; VEGF : Vasculer Endotelial Growth Factor. D'après [267-270].

Aussi, le rôle de IL-6 dans le cancer colorectal a fait l'objet de nombreux travaux. Plusieurs études montrent que cette cytokine favorise la croissance des cellules issues de cancers colorectaux en stimulant leur capacité d'invasion et semble également promouvoir la formation de tumeurs secondaires [258–260]. L'IL-6 est produite par les cellules de l'immunité sous la dépendance de NF- κ B [261,262] et serait donc un déterminant essentiel de la voie extrinsèque de la carcinogenèse associée à l'inflammation.

➤ **Carcinogenèse et chimiokines**

De nombreuses chimiokines et leurs récepteurs sont exprimés dans les tissus cancéreux. La chimiokine CCL2 jouerait un rôle essentiel dans le cancer colorectal. En effet le blocage de la voie TNF- α /TNFR réduit l'expression de CCL2 et en conséquence diminue l'infiltrat macrophagique et la progression du cancer colorectal [263]. De plus, des antagonistes de CCL2 inhibent l'expression de COX-2 (cf. ci-dessous), atténuent l'angiogenèse et diminuent le nombre et la taille des tumeurs [263].

Les chimiokines et leurs récepteurs joueraient également un rôle important dans la formation des métastases [250]. La migration des cellules issues du cancer du côlon est augmentée *in vitro* sous l'action de chimiokines [264]. Des études cliniques montrent une corrélation entre l'expression de CXCR3, CXCR4 et CCR7, la récurrence des tumeurs et l'apparition de métastases hépatiques [264,265].

➤ **COX-2 et carcinogenèse**

COX-2 est une enzyme impliquée dans la synthèse des prostaglandines. Son expression est associée à la carcinogenèse colorectale chez l'homme (**Figure 21**) [266]. La surexpression de COX-2 stimulerait la croissance tumorale grâce à un effet pro-angiogénique (stimule la production de VEGF) et son absence inhibe le développement de polypes colorectaux adénomateux [267,268]. Par ailleurs, il existerait un polymorphisme du gène codant COX2 qui prédisposerait au cancer colorectal [269].

➤ **Récepteurs Toll-like et carcinogenèse**

Une activation aberrante des récepteurs Toll-like (TLR) contribuerait à l'activité pro-

tumorale de NF- κ B. Ces activations peuvent être liées à des polymorphismes dans les gènes codant les TLR et certaines mutations sont associées à un mauvais pronostic [270]. Une déficience en MyD88, une protéine adaptatrice des TLR, réduit significativement le nombre et la taille des tumeurs intestinales dans un modèle murin APCmin [271] et la diminution d'expression de TLR4 dans des souris inoculées avec des cellules cancéreuses coliques induit une augmentation de la survie et une diminution de la taille des tumeurs [272].

➤ **TNF- α et carcinogénèse**

La cytokine TNF- α joue un rôle clé dans la régulation de l'inflammation et la défense de l'hôte. L'activation des récepteurs TNF (TNFR) induit la voie NF- κ B et donc la survie cellulaire ou peut induire la caspase 8 et donc induire l'apoptose [273]. Le TNF- α induit également l'expression de chimiokines pouvant avoir des effets pro- ou anti-tumoraux. Cependant, une inhibition de la voie de signalisation du TNF- α dans un modèle murin de tumeurs coliques induites entraîne une diminution de l'inflammation et du nombre de tumeurs [274], ce qui suggère que l'action pro-tumorale du TNF- α prend le pas sur l'effet anti-tumoral.

➤ **Carcinogénèse et stress oxydant**

Le stress oxydant est défini par une production d'espèces chimiques oxygénées ou azotées réactives dépassant les mécanismes de résistance anti-oxydant [275]. De nombreuses études réalisées *in vivo* supportent l'implication du stress oxydatif dans la carcinogénèse. L'injection de fer induit la production d'espèces oxygénées réactives et l'apparition de cancers chez le rat. Les agents infectieux carcinogènes comme les virus des hépatites B et C ou *Helicobacter pylori* induisent également un stress oxydatif massif qui pourrait contribuer à l'apparition de cancers [276]. Des facteurs de risques décrits dans le cancer colorectal telles que la consommation d'alcool, de tabac, de viande rouge ainsi que l'inflammation stimulent le relargage de radicaux libres oxygénés ou nitrosés. Ces composés induisent directement des dommages à l'ADN [220]. Les dommages à l'ADN occasionnés peuvent être extrêmement variables (formation de 8-oxoguanine conduisant à la transversion GC-TA, rupture de la molécule d'ADN,...) [277]. L'oxydation des bases de l'ADN par ces composés réactifs favorise l'apparition de mutations. De même l'oxydation des lipides peut générer des époxydes ou des aldéhydes, qui sont des adduits capables d'induire des mutations [275].

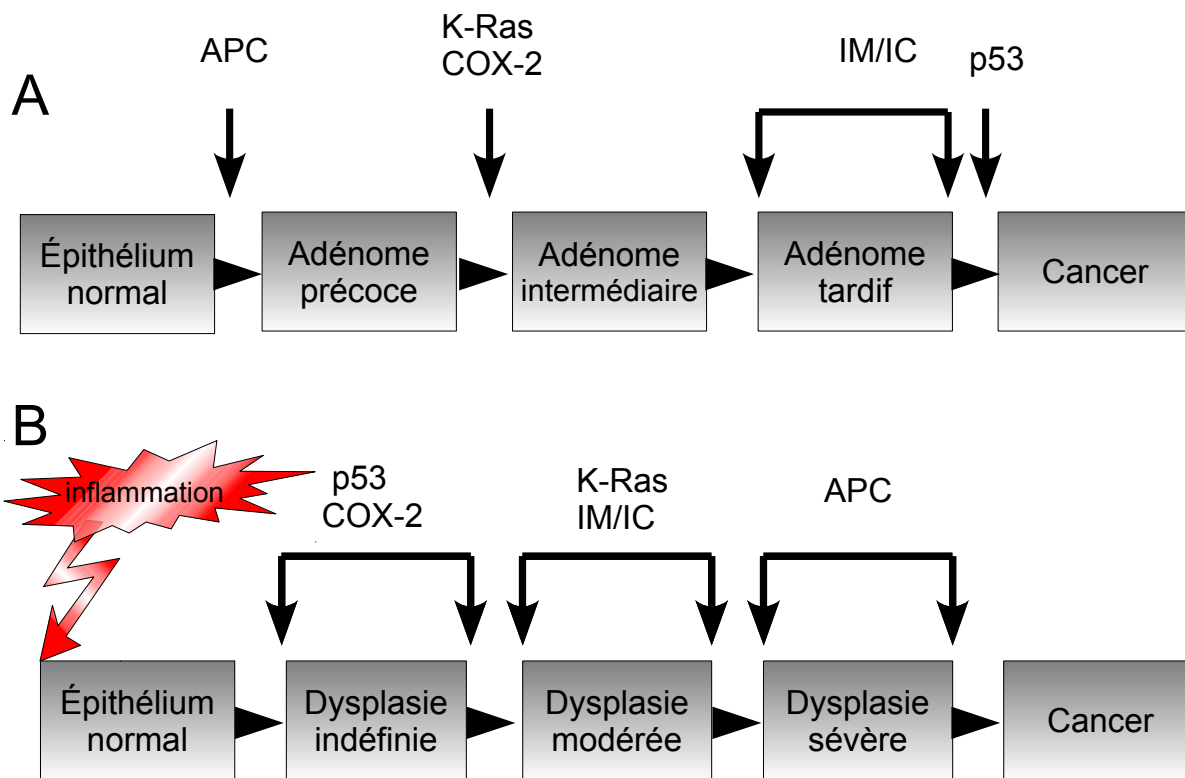


Figure 22 : les différentes séquences adénome / carcinome. La séquence adénome / carcinome correspondant au modèle de tumorigenèse colorectale proposé par Fearon et Vogelstein en 1990 (A). Séquence adénome / carcinome associée à une pathologie inflammatoire chronique (B). APC : *adenomatous polyposis coli* ; COX-2 : cyclooxygénase 2 ; IL-17 : interleukine 17 ; IM/IC : instabilité microsatellitaire/instabilité chromosomique.

3.B.g. Modèles de carcinogenèse

Dans le côlon, la transformation d'une cellule normale en cellule cancéreuse a été modélisée. Cette modélisation retrace les événements mutationnels qui font progresser la séquence adénome-cancer.

➤ Le modèle de Vogelstein : la séquence adénome cancer

Fearon et Vogelstein ont modélisé la carcinogenèse colorectale selon une séquence d'événements génétiques (**Figure 22**) [5]. Le processus s'initie au niveau des cellules progénitrices situées à la base des cryptes. Pour résumer, la séquence adénome-cancer par cette voie comporte quatre étapes :

1. la lésion initiale est le foyer de cryptes aberrantes : la mutation d'APC est présente lorsqu'il s'agit d'un PAF, mais absente dans le cancer colorectal sporadique. En revanche dans ce dernier cas, *K-ras* se trouve muté dans 68% des cas.
2. La mutation de *APC* conduit au développement d'un adénome précoce.
 - Si la mutation intervient chez un sujet *K-ras* sauvage (le plus fréquent), il n'y a pas d'évolution vers les stades ultérieurs de la séquence, ce qui explique que la majorité des adénomes n'évoluent pas vers un adénome plus grave ;
 - Si la mutation intervient chez un sujet déjà muté *K-ras*, cette dernière accélère le passage de l'adénome précoce vers un adénome plus avancé (adénome intermédiaire).
3. Les mutations de *DCC*, *SMAD2* et surtout *SMAD4* font avancer la séquence adénome-cancer vers le développement d'un adénome tardif.
4. La mutation de *p53* est la dernière étape de la séquence, signant la transformation d'une lésion non cancéreuse (adénome tardif) en lésion cancéreuse, l'adénocarcinome.

➤ Autres modèles

Une autre séquence adénome cancer a été proposée dans le contexte de pathologies inflammatoires chroniques [186,278]. Elle fait intervenir les mêmes mutations mais leur ordre d'apparition est bouleversé (**Figure 22**). Dans cette séquence, les mutations du gène codant pour *p53* seraient initiatrices et les altérations de *APC* surviendraient lors des dernières étapes de la carcinogenèse. L'apparition de ces mutations est liée à un contexte d'inflammation

chronique comme observé dans la maladie de Crohn. Elle serait induite par le stress oxydatif [276,279].

4. Microbiote et cancer colorectal

10^{13} bactéries commensales colonisent le côlon alors que le microbiote du duodénum et de l'iléon ne représente que 10^2 et 10^8 bactéries respectivement. Cette augmentation de la densité bactérienne est associée à une augmentation du risque de cancer d'un facteur 12 [10].

4.A. Le microbiote du cancer colorectal

Le séquençage des amplicons d'ARNr 16s a permis d'établir une carte à haute résolution de la composante bactérienne du microbiote intestinal associée au cancer colorectal [280–282]. Il émerge de trois études indépendantes une surreprésentation de *Fusobacterium spp.*, de membres de la famille des *Streptococcaceae*, et de microorganismes proposés comme probiotiques (membres de la famille des *Coriobacteriaceae*). A l'inverse, les sites sains adjacents aux tumeurs semblent former une niche pour des membres potentiellement pathogènes de la famille des *Enterobacteriaceae*, comme *Salmonella*, *Citrobacter*, *Cronobacter* et *Shigella*, ainsi que des membres de l'ordre des Actinomycetales. Les Firmicutes affichent une distribution disparate où certaines espèces sont enrichies dans le tissu tumoral tandis que d'autres vivent dans la muqueuse saine adjacente. Il semble donc émerger de ces travaux que le tissu cancéreux engendre l'apparition d'un microbiote spécifique qui peut s'expliquer par la présence de protéines membranaires surexprimées dans le cancer colorectal. Ces dernières peuvent servir de récepteurs spécifiques comme les CEACAM et les intégrines [283–287]. CEACAM1 et CEACAM6 sont reconnus par des adhésines bactériennes [288,289]. L'adhésine fimbriale FimH de *Escherichia coli* fixe les résidus mannoses de CEACAM6 [290]. Les intégrines sont également des glycoprotéines exprimées à la surface de nombreux types cellulaires [286], qui peuvent servir de récepteurs pour des bactéries [286,291].

Malgré les précieux renseignements fournis par ces études du microbiote, il convient de noter qu'elles ont des limites. Bien que l'abondance des espèces bactériennes soit mesurable, la profondeur de séquençage de ces études empêche l'identification des espèces minoritaires qui peuvent être importantes comme *E. coli*. Bien que l'une des études utilise une approche meta-transcriptomique [282], ces travaux déterminent uniquement la présence ou l'absence d'espèces bactériennes, alors que l'expression de gènes bactériens tels que les facteurs de virulence peut avoir un impact sur la carcinogenèse. Par ailleurs, l'analyse du microbiote associé au cancer colorectal a seulement été réalisés à partir de formes évoluées de la

maladie. Il est fort probable que le microbiote évolue avec le processus de carcinogénèse et que les bactéries initiatrices du cancer disparaissent au cours de cette évolution pour laisser la place à des bactéries plus adaptées au tissu tumoral. Aussi, Tjalsma et collaborateur ont récemment proposé une vision dynamique de la composition du microbiote (le modèle « driven-passenger ») et de son rôle au cours de la carcinogénèse colique en classant les bactéries en deux groupes : les bactéries « conductrices » et les bactéries « passagères » [10]. Les premières joueraient un rôle dans l'initiation du cancer en induisant des dommages à l'ADN (*Enterococcus faecalis* [292], *Streptococcus gallolyticus* [293,294], *Bacteroides fragilis*, entérobactéries [7,13]), en activant des oncogènes (*Bacteroides fragilis* [203,295]) ou en induisant une réaction inflammatoire exacerbée (*Shigella*, *Citrobacter*, *Salmonella*, *Bacteroides fragilis* [10]). Les bactéries « passagères » (ex. *Fusobacterium* spp., *Streptococcaceae*, et *Coriobacteriaceae*), installées sur des lésions cancéreuses établies, pourraient simplement accompagner, ralentir ou favoriser la progression de la pathologie. Cependant, l'identification des bactéries favorisant la carcinogénèse colique reste difficile à cause (i) du délai survenant entre l'initiation du processus et l'apparition des signes cliniques, (ii) de la diversité du microbiote intestinal et (iii) de la variabilité du génome entre les individus d'une même espèce bactérienne. Aussi, les données épidémiologiques et cliniques concernant le rôle de bactéries dans la carcinogénèse colorectale sont peu nombreuses voire absentes.

4.A.a. Microbiote intestinal, alimentation, et carcinogénèse

Le côlon est exposé en permanence à un mélange complexe de composés, qui sont d'origine alimentaire, microbienne ou le résultat de processus de digestion et d'excrétion. Ces composés peuvent endommager la muqueuse intestinale et être à l'origine de cancers colorectaux.

Le métabolisme des substrats alimentaires par certains groupes bactériens du microbiote intestinal peut potentiellement produire des métabolites réactifs carcinogènes [8,279]. En conséquence, certains groupes bactériens du microbiote associés à une alimentation pourraient être un facteur de risque du cancer colorectal [8,279,296,297].

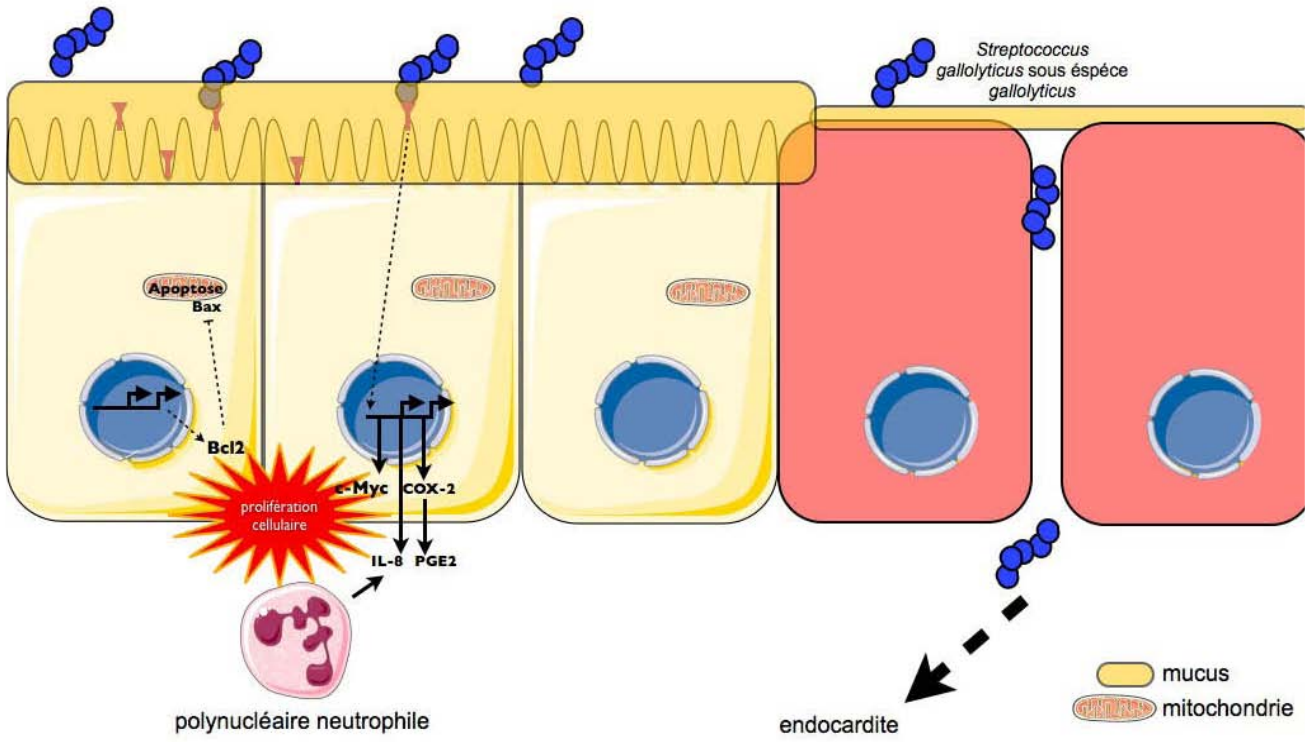
L'alcool déshydrogénase microbienne (ADH) exprimée par certains groupes bactériens du côlon peut être impliquée dans l'induction de dommages à l'ADN dans les cellules épithéliales [8,279]. En présence d'éthanol, l'ADH microbienne peut conduire à la production de

l'acétaldéhyde, un carcinogène chimique connu [298]. Les lésions de l'ADN dans les cellules épithéliales sont également provoquées par certaines bactéries commensales comme *Eubacterium faecalis*, qui produisent des espèces chimiques oxygénées réactives tels que le superoxyde, le peroxyde d'hydrogène, les radicaux d'hydroxyle et le peroxyde nitrite. Les bactéries sulfogéniques sont des membres du microbiote du côlon normal. Ces bactéries utilisent le sulfate comme oxydant pour la dégradation de la matière organique. Du sulfure d'hydrogène est formé lors de ces réactions ; or une association entre le sulfure d'hydrogène et le cancer colorectal a été démontrée [299]. En outre, il est intéressant de noter que pendant la circulation entéro-hépatique, les sels biliaires rencontrent les bactéries anaérobies facultatives du tractus digestif et deviennent le substrat d'une biotransformation microbienne [300]. La déhydroxylation des acides biliaires effectuée par certaines bactéries du côlon conduit à la formation d'acides biliaires secondaires dont le rôle promoteur de tumeur est établi [301]. En particulier, l'acide deoxycholic et l'acide lithocholic sont liés à la carcinogénèse du côlon dans plusieurs études épidémiologiques humaines [302]. Ainsi, la consommation d'aliments à haute teneur en graisses, qui induit une sécrétion forte d'acides biliaires, est corrélée avec le risque de cancer colorectal [301].

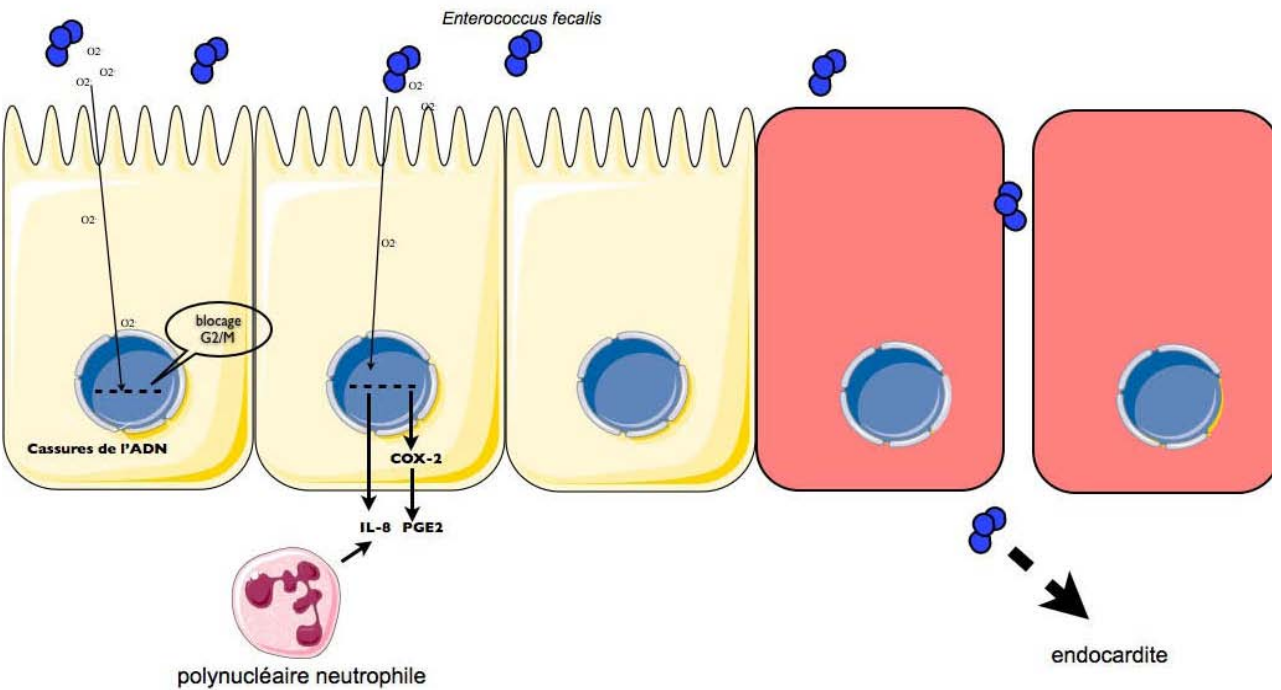
4.A.b. Microbiote intestinal, inflammation et carcinogénèse

Certains microorganismes du côlon peuvent intervenir à différents stades du processus néoplasique par l'induction d'une inflammation. Une réponse inflammatoire aberrante d'origine infectieuse ou non dans le tractus digestif peut conduire à une dysbiose intestinale en faveur de pathobiontes ou de pathogènes allochtones [107]. Les pathobiontes ou pathogènes allochtones peuvent faire basculer un état inflammatoire physiologique vers un état inflammatoire pathologique [103,104]. L'envahissement par ces microorganismes pathogènes contribue donc à la consolidation de l'état inflammatoire par l'induction d'une boucle pro-inflammatoire auto-entretenu [13,109,115,117,303], qui peut alors favoriser le développement du cancer colorectal comme évoqué précédemment.

La meilleure preuve du lien existant entre l'inflammation induite par le microbiote et la carcinogénèse est l'augmentation du risque de cancer colorectal dans les maladies inflammatoires chroniques de l'intestin comme la rectocolite hémorragique ou la maladie de Crohn [13,97,117]. Dans la maladie de Crohn, plusieurs études montrent une dysbiose pro-inflammatoire avec notamment une colonisation anormalement élevée de l'épithélium intestinal



ub d s p
c b n d C d n u b d C i t d n s p C
p n e d C e d



i c h o
é
i d s p c
c u b d C i t d n s p C p n
e d C e d

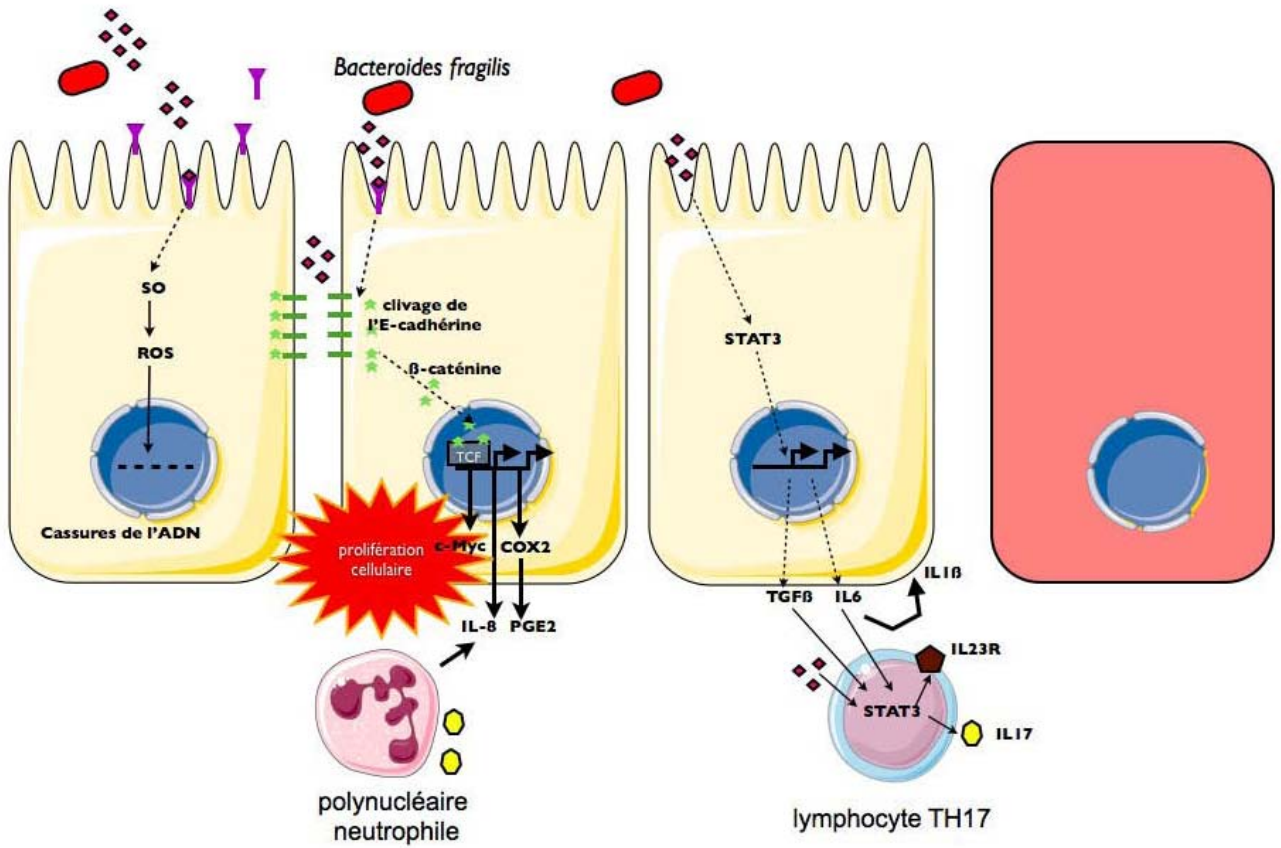
par des *E. coli* adhérents invasifs stimulant la sécrétion de cytokines pro-inflammatoires (TNF- α et IL-8) [290]. Lors de la mise en place de la dysbiose, certaines bactéries ayant des propriétés anti-inflammatoires voient leur nombre diminuer, principalement l'espèce *Faecalibacterium prausnitzii* qui sécrète des métabolites capables de bloquer l'activation de la voie NF- κ B et la production d'IL-8 [304].

Par ailleurs, le rôle du microbiote dans la carcinogenèse colique a été exploré dans des modèles murins d'inflammation chronique de l'intestin (souris TGF- $\beta^{-/-}$ et IL-10 $^{-/-}$) [303,305]. Il ressort de ces travaux que les animaux élevés dans des conditions axéniques développent une inflammation réduite sans transformation maligne [13,306]. La présence du microbiote ou l'introduction de bactéries, tel que *E. coli* ou *E. faecalis* conduit par contre à une colite suivie par l'apparition de tumeurs [13,303,307].

Les mécanismes associés à la carcinogenèse n'ont pas été spécifiquement explorés en réponse aux infections bactériennes, mais rejoignent très probablement les processus évoqués précédemment avec un rôle prépondérant du stress oxydatif, de la voie NF- κ B, des cytokines et des chimiokines pro-inflammatoires et de l'infiltrat leucocytaire [308].

4.B. *Streptococcus gallolyticus gallolyticus* et carcinogenèse

Streptococcus gallolyticus sous espèce *gallolyticus*, anciennement nommé *Streptococcus bovis* biotype I, est une bactérie commensale entérique commune. Elle est responsable d'endocardites et plus rarement de septicémies [309,310]. Plusieurs études ont rapporté une fréquence anormalement élevée du portage digestif et des bactériémies à *S. gallolyticus* spp. *gallolyticus* chez les patients atteints de cancer colorectal [310–312]. Plus récemment, Bakar et collaborateurs ont montré que *S. gallolyticus* spp. *gallolyticus* présente une activité pro-inflammatoire en induisant la production d'IL-1, de COX-2 et d'IL-8 (**Figure 23**) [313]. Cette activité était associée à la surexpression de l'oncogène c-Myc et de la protéine anti-apoptotique Bcl2. Cependant, le rôle de cette bactérie dans la carcinogenèse colique reste discuté [310,314], car des études récentes ont remis en cause les données épidémiologiques obtenues précédemment [309,311,312,315,316].



n C n
 ug N) vgu) qf
 n C t C
 p a C c d
 n C a b g à S C t
 p ch d x O C uc c
 ô C
 coyD
 e ô s
 g à Se n Ss e ô s h t Se t Ss.vgu s.v n
 n suc c(eu c C c (sch dpe
 p

4.C. *Enterococcus faecalis* et carcinogénèse

Enterococcus faecalis est une autre bactérie commensale de la flore digestive humaine. C'est également une bactérie pathogène opportuniste responsable d'infections extra-intestinales telles que des infections du tractus urinaire ou des endocardites [317,318]. Comme avec *S. gallolyticus* spp. *gallolyticus*, un taux anormalement élevé de bactériémies à *E. faecalis* a été observé chez des patients atteints de cancer colorectal [318–321]. Par ailleurs, *E. faecalis* produit *in vitro* et *in vivo* des ions superoxydes extracellulaires [322]. Ces ions sont convertis en peroxyde d'hydrogène, qui est génotoxique (**Figure 24**) [292]. Cette activité génotoxique induit une instabilité chromosomique [320,323]. De plus, l'activité protumorale de *E. faecalis* a été mise en évidence dans un modèle murin d'inflammation chronique [307].

4.D. *Bacteroides fragilis* et carcinogénèse

Bacteroides fragilis est une bactérie pathogène opportuniste du microbiote intestinal. Elle est responsable d'infections intestinales et extra-intestinales [324]. Toprak et collaborateurs ont montré une colonisation anormalement élevée des tumeurs colorectales par des souches de *Bacteroides fragilis* entérotoxiques, des bactéries habituellement responsables de diarrhées inflammatoires aiguës chez l'enfant [325]. Cette bactérie produit une toxine, appelée BFT (« *Bacteroides fragilis* toxin ») ou fragilysine [11].

La toxine clive l'E-cadherine *in vitro* (**Figure 25**) [295]. La protéine ainsi clivée libère la β -cathénine jusqu'alors séquestrée dans le cytoplasme et active la voie Wnt/APC/ β -cathénine, qui joue un rôle essentiel dans la prolifération cellulaire et la carcinogénèse colique.

En réponse à l'infection par des *Bacteroides fragilis* entérotoxigènes, les cellules épithéliales intestinales HT29 et T84 surexpriment également la spermine oxydase. Cette enzyme produit des radicaux libres, qui engendrent des cassures de l'ADN [326]. Ces lésions induisent la production de IL-8, l'expression de COX-2 et en conséquence la production de prostaglandine pro-inflammatoire par les cellules épithéliales intestinales. La production de cytokines pro-inflammatoires et de radicaux libres au niveau des foyers d'infection correspond à une réponse immunitaire normale en présence de bactéries [119,327]. Elle devient par contre délétère pour l'hôte si elle devient chronique ou exagérée comme cela est observé en modèle murin [12].

La fragilysine induit *in vitro* et *in vivo* d'importantes altérations de la muqueuse intestinale [11,203,326,328]. L'analyse histologique des intestins d'animaux infectés avec la bactérie toxigénique ou traités avec la toxine purifiée révèle une destructuration de l'épithélium ainsi que des infiltrats de la *lamina propria* constitués de cellules immunitaires [329,330]. De plus, des souris infectées par les bactéries toxigènes développent une inflammation chronique de l'intestin et une prolifération accrue des cellules conduisant à l'apparition de tumeurs [331].

L'analyse de cette réponse inflammatoire dans un modèle murin montre que les bactéries productrices de BFT entraînent une colite caractérisée par la différenciation de lymphocytes de type TH17 producteurs d'IL-17 et IL-23, suite à l'activation sélective du facteur de transcription STAT3 et la production d'IL-6, d'IL-1 et de TGF- β . De façon intéressante, le blocage de IL-17, et du récepteur de IL-23, une cytokine amplifiant la réponse TH17, bloque la colite, l'hyperplasie et les tumeurs induites par la bactérie. Il faut noter que la réponse immunitaire de type TH17 a été associée à une action carcinogène dans d'autres modèles mais des effets inverses ont été aussi rapportés (**Figure 25**) [12,332,333].

L'ensemble de ces données suggèrent que les *Bacteroides fragilis* entérotoxigènes pourraient jouer un rôle dans la carcinogenèse colique.

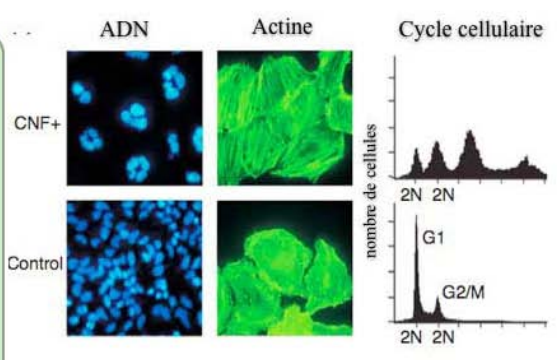
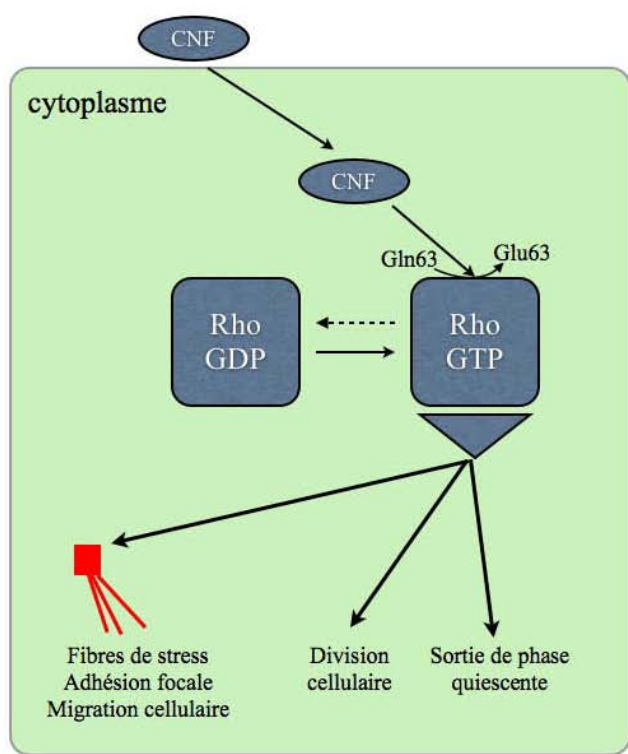
4.E. *Escherichia coli* et carcinogenèse

L'espèce *E. coli*, comme *B. fragilis*, est le prototype du pathobionte. Elle se comporte le plus souvent comme une bactérie commensale. Cependant, elle est surreprésentée dans les tumeurs de patients atteints de cancer colorectal [110,334]. *E. coli* peut acquérir de nombreux facteurs de virulence lui permettant d'accéder à de nouvelles niches écologiques et d'être responsable d'infections intestinales ou extra-intestinales. Un intérêt croissant est porté à l'activité potentiellement carcinogène de toxines affectant le cycle cellulaire [25].

4.E.a. Cytotoxic Necrotizing factor

La cyclomoduline CNF (« Cytotoxic Necrotizing factor ») est la seule cyclomoduline de *E. coli* qui stimule l'avancée du cycle cellulaire. Trois autres cyclomodulines favorisant la prolifération cellulaire ont été caractérisées : la toxine dermonecrotique (DNT) de *Bordetella*

□



x p u ni cF u cN u
 r d s L q L u ni é a
 s o Ce L
 s o Ce. I a C a v d j

□

□

spp., PMT (« Pasteurela Multocida Toxin ») de *Pasteurela multocida* et CagA chez *Helicobacter pylori* [25,335–337].

➤ Support génétique et structures

CNF a été décrite initialement par Caprioli *et al.* en 1983 [338]. Il existe trois variants chez *E. coli*, CNF1 à CNF3. Un quatrième a été décrit chez *Yersinia enterocolytica* (CNFy). Ces toxines partagent 65 à 90 % d'identité. CNF1, CNF3 et CNFy sont codés par des gènes chromosomiques, CNF2 par un gène plasmique [339]. Le gène *cnf1* est situé sur l'îlot de pathogénicité PAI-II en aval de l'opéron *hlyCABD*, qui code l'hémolysine alpha. CNF3 est retrouvé chez les souches responsables d'infections chez les ovins et les caprins.

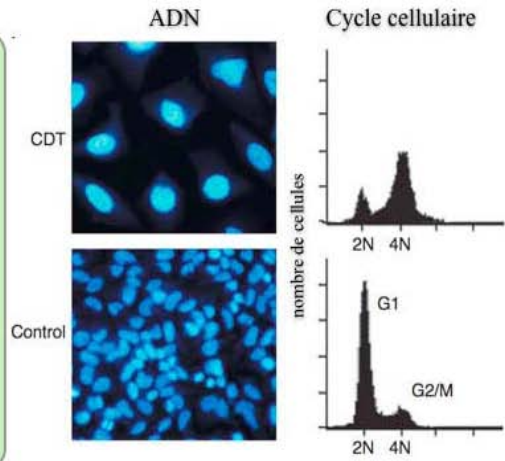
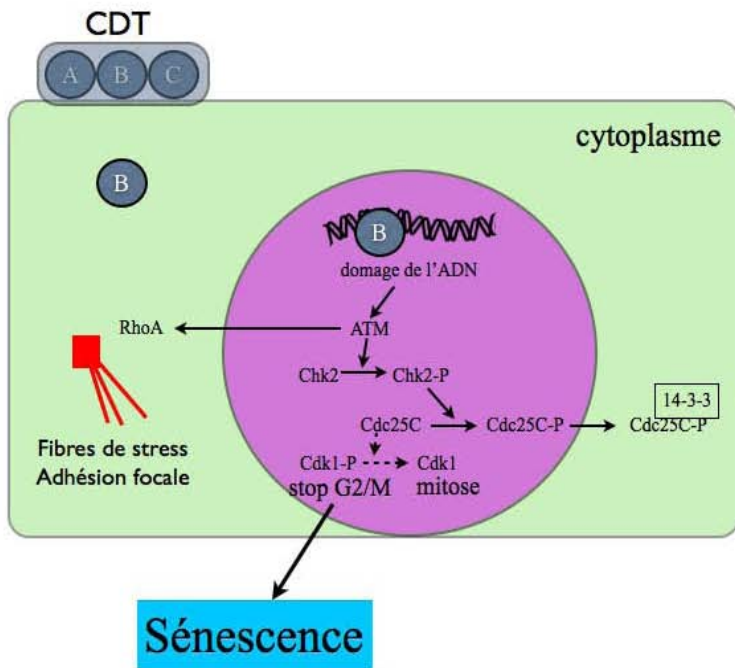
Les toxines de la famille CNF sont des protéines d'environ 115kDa organisées en trois domaines fonctionnels [25,340]. L'extrémité N-terminal contient un domaine de liaison au récepteur cellulaire. Le domaine central assure la translocation du domaine C-terminal dans le cytoplasme. Enfin l'extrémité C-terminal contient le domaine enzymatique responsable de l'activité toxique.

➤ Mécanisme d'action dans la cellule cible.

CNF (comme DNT) possède une activité transglutaminase et catalyse la désamination ou la polyamination des protéines de la famille des Rho GTPase : Rho, Rac et Cdc42 sur la glutamine en position 63 [25,341]. Il s'ensuit une activation de ces GTPases qui contribuent à la dérégulation d'un large panel de processus cellulaires en interagissant avec des effecteurs protéiques spécifiques (**Figure 26**) [342].

Ainsi, CNF favorise l'entrée des cellules dans le cycle cellulaire et la transition G1/S en affectant la production de cycline D1, et des CDKi p21^{CIP} et p27^{KIP} [25,340]. CNF entraîne le contournement du point de contrôle, induit des altérations du fuseau mitotique, de la cytokinèse, de la ségrégation des chromosomes et peut ainsi occasionner des aneuploïdies ou des endoreduplications de l'ADN (**Figure 26**) [25,343]. CNF bloque également la différenciation des cellules et augmente l'expression des protéines anti-apoptotiques Bcl-2 et Bcl-XL [336,344].

Plusieurs arguments soutiennent le rôle potentiel de CNF dans la transformation cellulaire. La modification biochimique des protéines Rho induite par les CNF correspond à la modification des protéines Ras observée dans de nombreuses tumeurs [173,345]. Les toxines



dh u

D

C

Cdr

cdr

a

a

pm

c

C

e

l

-: f A

CNF dérèglent l'activité des GTPases Rho comme les effecteurs SopE et SptP de *Salmonella* spp., bactéries dont le portage est associé à la survenue de cancer de la vésicule biliaire [346,347]. L'activité de CNF s'apparente par de nombreux aspects à celle des protéines CagA de *H. pylori*, un des meilleurs exemples de toxines carcinogènes. Dans des cellules épithéliales, CNF1 induit une activation prolongée du facteur de transcription NF- κ B [348], un acteur reconnu de la carcinogénèse [256]. Cette activation induit l'expression de *cox2* [345], un autre acteur de la carcinogénèse colorectale [268]. CNF1 en activant de façon prolongée la voie NF- κ B favorise donc un état inflammatoire chronique connu pour favoriser la carcinogénèse [246,248]. L'activité de CNF1 favorise la prolifération cellulaire, protège les cellules épithéliales de l'apoptose et induit des aneuploïdies ; des altérations pouvant conduire à la transformation cellulaire [348]. CNF1 promeut la rupture des jonctions cellulaires et la motilité cellulaire, par un mécanisme dépendant de RhoA [280,288] ; un mécanisme proche de la dédifférenciation. Enfin, les *E. coli* producteurs de CNF1 sont surreprésentés dans les tumeurs colorectales [349,350].

4.E.b. Cytolethal distending toxin

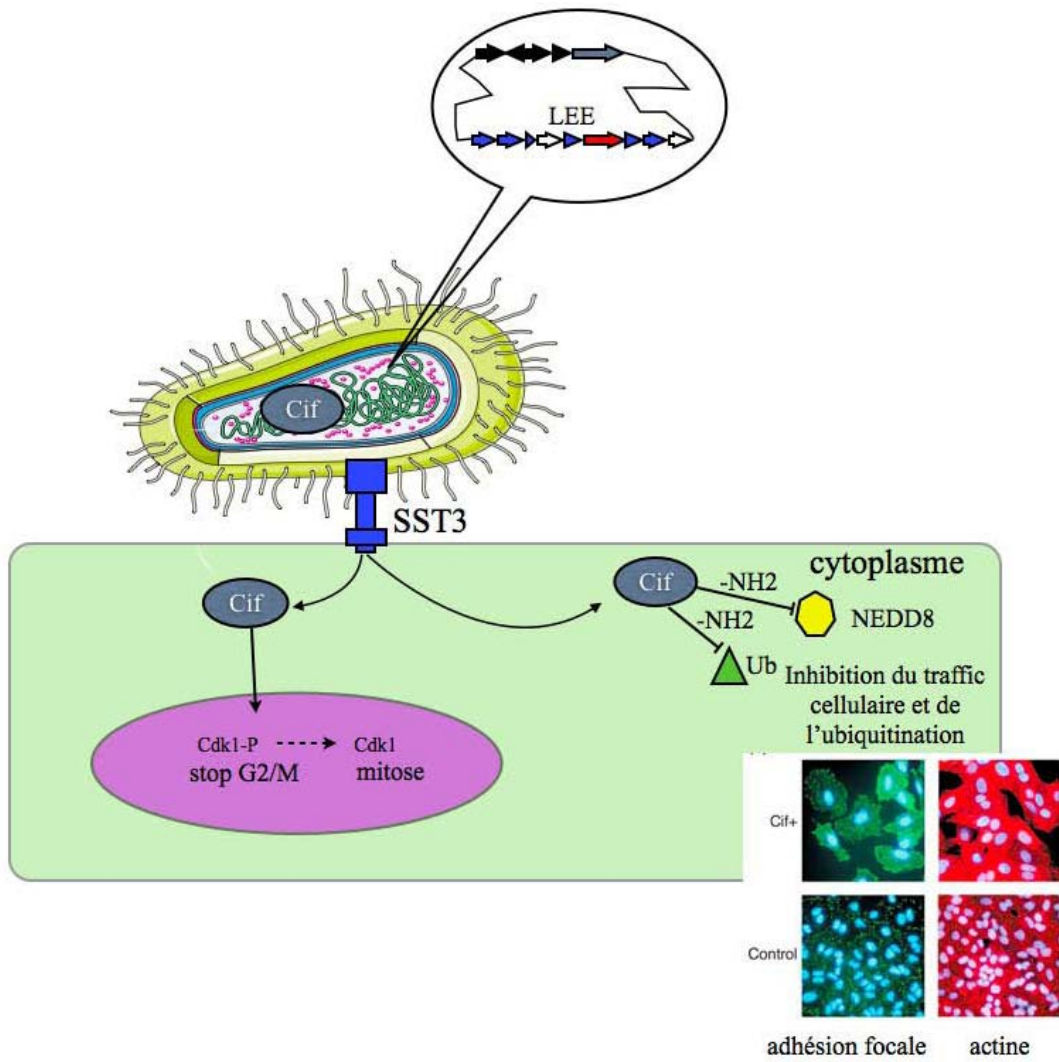
La toxine CDT (« Cytolethal distending toxin ») induit un blocage du cycle cellulaire en phase G2/M [25,351]. Elle est observée chez de nombreuses bactéries Gram négatif dont les entérobactéries : *Escherichia coli*, *Shigella dysenteriae*, *Salmonella enterica* serotype *typhi* [25,340,352].

➤ Support génétique et structure

CDT est une toxine constituée de 3 sous-unités CdtA, CdtB et CdtC codées par un opéron. Les sous-unités A et C sont impliquées dans la liaison de la toxine à la membrane plasmique des cellules cibles. La sous-unité CdtA bénéficie d'un greffage lipidique au niveau d'un domaine consensus appelé « lipobox » et la majorité des CDT reste étroitement associées aux vésicules de la membrane externe. CdtB, la sous-unité la plus conservée, présente une homologie structurale et fonctionnelle avec des métallo-enzymes dont des nucléases et possède un domaine d'adressage au noyau [351].

➤ Mécanisme d'action dans la cellule cible.

Les radeaux lipidiques riches en cholestérol de la membrane cytoplasmique sont



s C C Cp C CC
 f r r oh f
 è r r oh d Pi t
 (u Cx x u Cx x L n a s x f
 u f PmH

impliqués dans l'adressage de CDT dans la cellule cible [351]. La sous-unité CdtB est internalisée par endocytose, transportée par voie rétrograde vers l'appareil de Golgi et le réticulum endoplasmique, puis transloquée dans le noyau [25,351]. La sous-unité B y exerce une activité enzymatique comparable à celle de la DNase I [351]. Elle provoque des cassures double brin de l'ADN, qui active le système de réparation Mre11-Rad50 ainsi que la voie de signalisation ATM (**Figure 27**) [131]. ATM active par phosphorylation la kinase Chk2, qui intervient dans le point de contrôle G2/M [351]. Cette dernière phosphoryle la phosphatase Cdc25c qui est alors séquestrée par la protéine 14-3-3 dans le cytoplasme. Cette séquestration hors du noyau de Cdc25c maintient le complexe Cdk1-cycline B de l'état phosphorylé inactif et bloque la transition G2/M. Par ailleurs, les cassures de l'ADN induites par CDT entraînent l'activation du proto-oncogène c-MYC [353] et l'entrée de cellules en culture en sénescence [352].

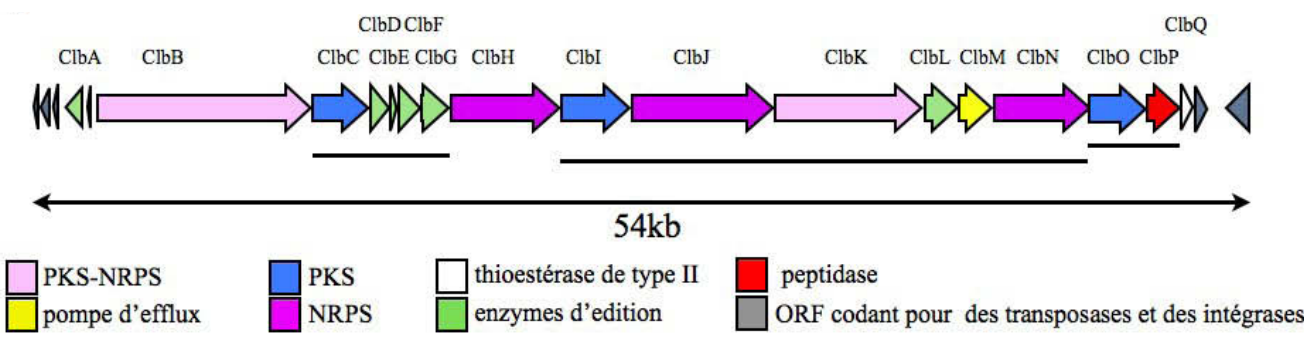
Par ailleurs, des expérimentations chez l'animal soutiennent l'hypothèse d'un rôle de CDT dans la carcinogenèse. La toxine CDT produite par la bactérie hépato-carcinogène *H. hepaticus* favorise le développement d'une dysplasie hépatique chez la souris [354]. L'activité génotoxique de CDT favoriserait l'acquisition d'événements mutationnels pouvant activer des gènes oncogènes et/ou inactiver des gènes suppresseurs de tumeurs.

4.E.c. Cell cycle inhibiting factor

La toxine Cif (« Cell cycle inhibiting factor ») est produite par les *E. coli* entéropathogènes ou entérohémorragiques. Ces bactéries sont caractérisées par leur capacité à détourner le cytosquelette à leur profit suite à l'injection d'effecteurs dans la cellule hôte *via* un système de sécrétion de type III (SST3). Ce système est codé par l'îlot de pathogénicité LEE (« Locus of enterocyte effacement »). Bien que non codée par LEE, la cyclomoduline Cif est injectée directement dans les cellules hôtes *via* le SST3 [355,356].

➤ Support génétique et structure

Le gène codant Cif est porté par un prophage lambdaïde [356]. La toxine Cif est organisée en deux régions. La séquence N-terminale porte un peptide d'adressage au SST3, et le domaine C-terminal porte l'activité enzymatique de la toxine. Sa structure est apparentée à celle des peptidases à cystéine active, des acétyltransférases et des transglutaminases.



□□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□
 □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□
 t □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□
 L □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□
 L □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□ □□□□□□□□□□□□□□□□□□□□

➤ Mécanismes d'action dans la cellule cible

Dans la cellule hôte, Cif inactive l'enzyme E2 du système d'ubiquitination en désaminant la glutamine située en position 40 entraînant une inhibition de la voie ubiquitine/protéasome, une accumulation des CKI p21^{CIP} et p27^{KIP} et un blocage du cycle cellulaire en G₂/M ou G1/S selon le stade du cycle lors de l'intoxication [357] (**Figure 28**).

4.E.d. Colibactine

La colibactine est une génotoxine découverte chez *Escherichia coli* en 2006 par l'équipe du Professeur E. Oswald [26]. Sa synthèse est assurée par une chaîne d'assemblage enzymatique codée par un cluster de gènes (54kb) appelé « *pks* » pour « polyketide synthetase ». Cet îlot est inséré dans le gène codant pour un ARN de transfert de l'asparagine, un site courant d'intégration des îlots de pathogénicité.

➤ Support génétique

L'îlot génomique *pks* se compose de 19 gènes *clbA* à *clbR* (**Figure 29**) [26]. L'activité transcriptionnelle des gènes de l'îlot a révélé un découpage en sept ARN messagers polycistroniques de tailles variables [358]. Un polycistron de 23,3kb englobe les gènes de *clbI* à *clbN*. Le deuxième recouvre les gènes *clbC* à *clbG* et mesure 6,2kb. Deux autres transcrits codent 2 protéines, un premier de 3,9Kb code *clbO* et *clbP*, le suivant de 0,9 kb code *clbR* et *clbA*. Les autres transcrits forment des monocistrons.

Les gènes *clbL*, *clbR* et *clbM* codent respectivement une intégrase de type P4, une protéine possédant un domaine de liaison à l'ADN de type hélice-boucle-hélice et une protéine apparentée aux pompes d'efflux MATE (« multidrug and toxic compound extrusion »). Les autres gènes codent des enzymes potentiellement responsables de la synthèse de la colibactine. L'analyse des séquences suggère la présence d'une phosphotransférase de type II (*clbA*), de trois synthétases de peptides non ribosomiaux (*clbH*, *clbJ* et *clbN*), trois synthétases de polykétides (*clbC*, *clbI* et *clbO*), deux synthétases d'hybrides polycétides-peptides non ribosomiaux (*clbB* et *clbK*), une thioesterase de type II (*clbQ*), et six enzymes d'édition (*clbD*, *clbE*, *clbF*, *clbG*, *clbL* et *clbP*). Toutes ces protéines sont nécessaires à l'activité

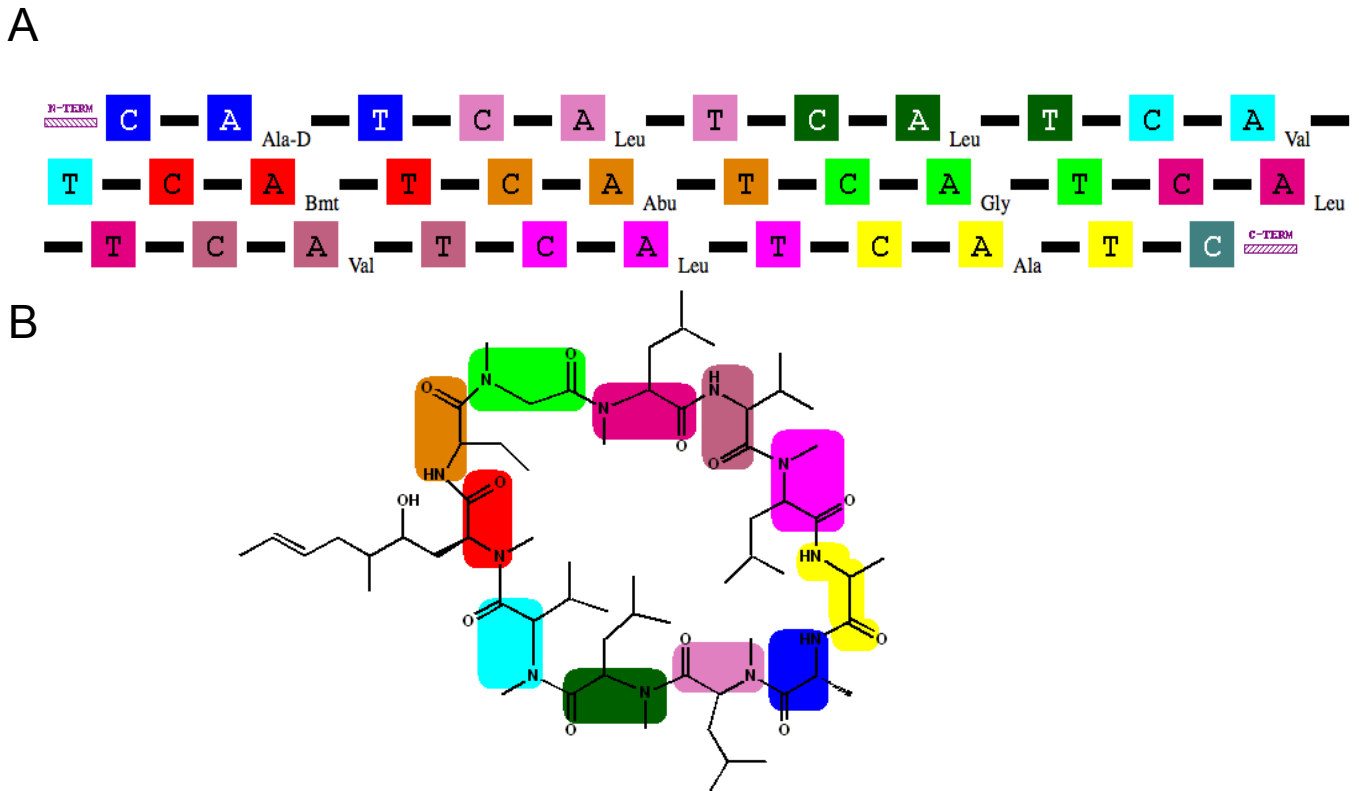


Figure 30 : la cyclosporine A, un exemple de composé issu d'une mégasyntèse de type NRPS. L'ensemble des réactions enzymatiques nécessaire à l'assemblage de la cyclosporine A est réalisé par une enzyme de 1,4MDa, contenant 34 domaines répartis en 11 modules. Chacun de ces modules est présenté de couleur différente et les acides aminés incorporés sont indiqués (A). Le squelette des acides aminés incorporés par chaque module présente la même couleur sur la molécule de cyclosporine A (B). A : Adénylation ; C : Condensation ; T : Transfert. D'après Mohanty (<http://www.nii.res.in/nrps-pks.html>).

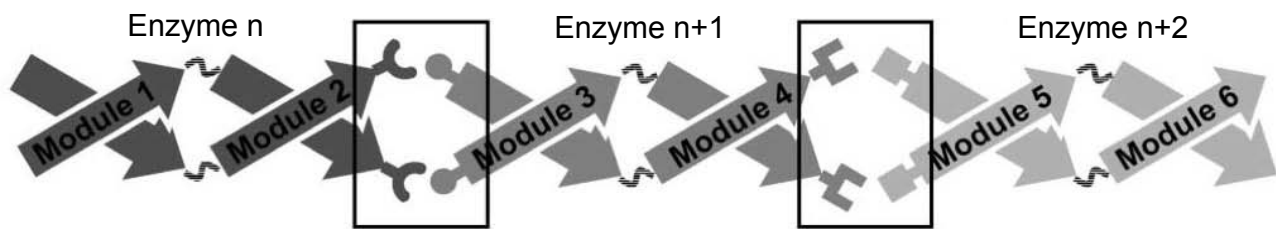


Figure 31 : organisation des mégasyntèses d'une chaîne d'assemblage. A l'extrémité de chaque enzyme se trouve des domaines d'interaction protéine-protéine permettant l'association spécifique de l'enzyme n à l'enzyme n+1. D'après Broadhurst et collaborateur [376].

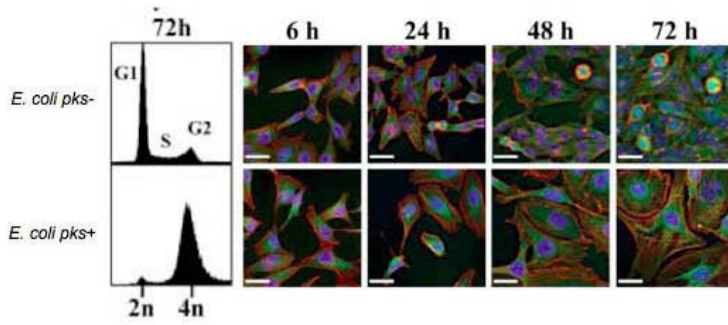
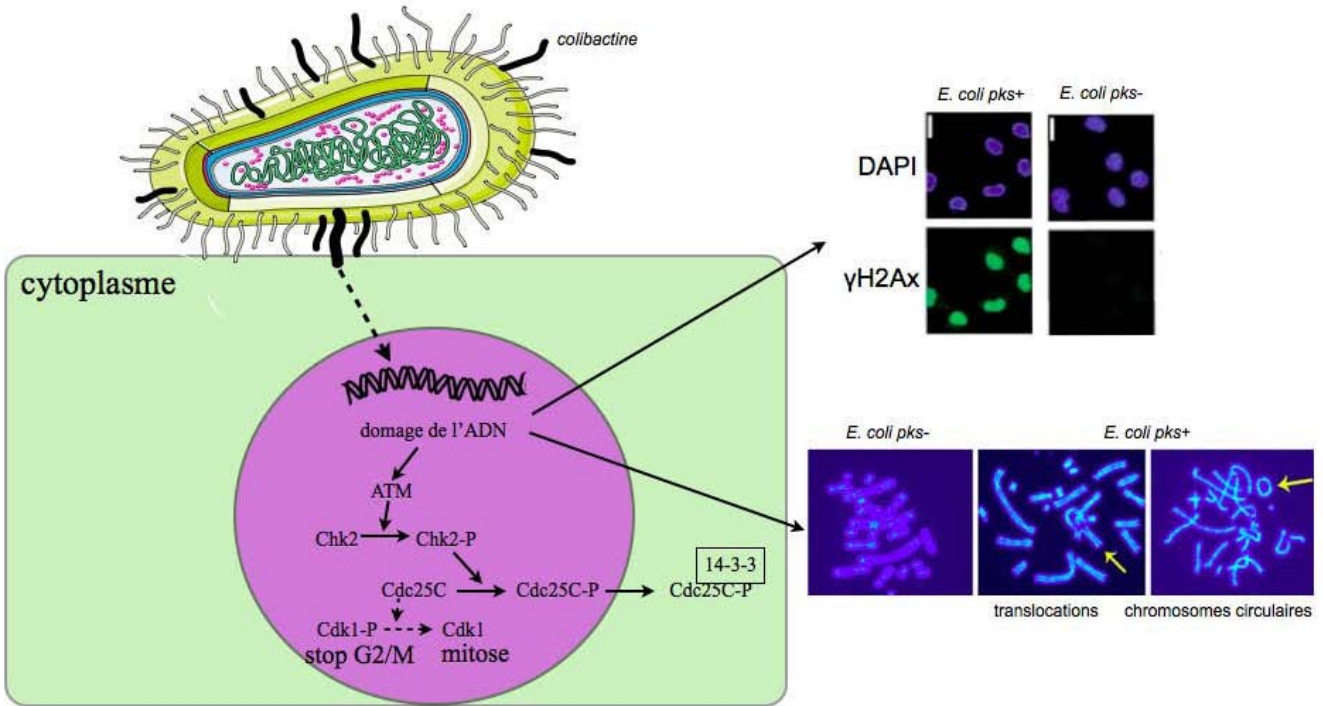
toxique des bactéries productrices de colibactine [14,26,359].

➤ Structure

Malgré de multiples essais, la colibactine n'a pas été isolée à ce jour et sa structure précise reste inconnue. Cependant, l'analyse *in silico* de la fonction des gènes montre qu'elle se compose d'un ou plusieurs composés hybrides constitués de polycétides et de peptides non ribosomiaux. Ce type de composés, plus fréquemment observés chez les bactéries appartenant aux Firmicutes (ex. *Bacillus spp.*) et aux Actinomycetales (ex. *Streptomyces spp.*), regroupe de nombreuses molécules bioactives souvent toxiques, parfois utilisées en thérapeutique. Elles constituent l'essentiel de la pharmacopée d'origine naturelle et sont une source d'inspiration majeure pour les chimistes. Parmi les premiers utilisés en thérapeutique, on dénombre notamment les pénicillines et les céphalosporines. Depuis, un grand nombre de composés a été découvert et utilisé en thérapeutique : d'autres familles d'antibiotiques majeurs (glycopeptides, macrolides, aminosides, tetracyclines, daptomycine etc...), des antifongiques, des anticancéreux, des antiviraux, des immunosuppresseurs, des hypocholestérolémiants mais aussi des insecticides, des tensioactifs, ... [27,360–366].

Ces composés sont produits par des chaînes d'assemblage enzymatique [360] grâce aux enzymes NRPS (« Synthase ») (**Figure 30**), PKS (« PolyKetide Synthase ») et PK-NRPS (« PolyKetide - Non-Ribosomal Peptide Synthase »). Chaque enzyme porte un ou plusieurs modules enzymatiques qui polymérisent des monomères. Ces modules portent en général 3 domaines : (i) le domaine Thiolation (T) de fixation des monomères, (ii) le domaine KetoSynthase (KS) des PKS ou le domaine de Condensation (C) des NRPS, et (iii) le domaine AcylTransférase (AT) des PKS ou le domaine d'Adénylation (A) des NRPS. Ces deux derniers domaines assurent l'assemblage des monomères au composé en cours de synthèse.

Les enzymes s'organisent en méga-complexes grâce à des domaines d'interactions spécifiques permettant d'ordonner les enzymes en une chaîne d'assemblage (**Figure 31**) [360,367]. L'organisation s'apparente à la machinerie de traduction des ARN messagers où la séquence des nucléotides est remplacée par l'ordre des enzymes et le ribosome par les activités enzymatiques de chaque module. Si dans la traduction, le ribosome se déplace sur l'ARN, dans la chaîne d'assemblage enzymatique, le composé synthétisé est transféré d'un module enzymatique au suivant comme dans une chaîne d'assemblage d'automobiles. A



b' æP c pé e/ b xdD pé y / A s y / A xEid d / A r rx y g/ æ Eé b p n y m uo L o t. p' L. p t. e' L. e r t p n L E R t g/ Lg / p é K DN

chaque étape, les modules enzymatiques NRPS assurent l'incorporation d'un monomère acide aminé en catalysant la formation d'une liaison OC-N (pont peptidique). Les enzymes PKS incorporent des monomères acyl-Coenzyme A en catalysant la formation de liaison C-C. La nature des monomères incorporés varie selon les enzymes ; on a dénombré plus de 300 acides aminés différents (protéinogènes ou non) et presque autant de monomères acyl (ex. acethyl, malonyl, methylmalonyl...) dans les composés NRP, PK et PK-NRP identifiés. En fin de chaîne d'assemblage, le composé est libéré à partir de l'enzyme terminal par une thioestérase, qui le plus souvent génère un composé linéaire ou cyclisé [360,365,368]. Par ailleurs, les composés sont modifiés en cours ou en fin de synthèse par des enzymes d'édition. Ces dernières catalysent une grande diversité de réactions comme des méthylations, des réductions, des oxydations, des halogénations et des isomérisations. Ainsi, la famille des composés NRP, PK ou PK-NRP contient une grande diversité de molécules portant des activités très différentes.

➤ Mécanisme d'action de la colibactine

L'action des bactéries productrices de colibactine a été explorée *in vitro* sur plusieurs lignées cellulaires et *in vivo* dans un modèle murin d'anses intestinales ligaturées et chez des souris IL10^{-/-} axénique monocolonisées par une souche porteuse de l'îlot *pks* [13,26,369]. Ces bactéries induisent, *in vitro* et *in vivo*, des cassures double brin de l'ADN [13,14,26]. Cette activité génotoxine, apparentée à celle de CDT, nécessite des bactéries vivantes et un contact direct entre les bactéries et les cellules. Ces cassures de l'ADN induisent une activation du système de réparation de l'ADN par la voie de signalisation ATM. Cette activation se traduit par la phosphorylation de ATM et l'apparition d'histones γ H2Ax phosphorylés en position 139 (**Figure 32**). Cette activation de ATM induit une cascade de phosphorylation passant par les protéines Chk2 (« Check point Kinase 2 ») et Cdc25c (« Cyclin Dépendent kinases phosphatase C ») comme observé avec la cyclomoduline CDT [26,351]. La phosphatase Cdc25c phosphorylée est séquestrée dans le cytoplasme par les protéines 14-3-3. La protéine Cdk1 nucléaire reste donc inactive, ce qui induit un arrêt du cycle cellulaire en phase G2/M. Cet arrêt du cycle cellulaire se traduit par l'accumulation de cellules en phase G2 contenant une quantité d'ADN équivalente à 4n et une mégaloctose. Une analyse cytogénétique de ces cellules a révélé des signes de réparation incomplète marqués par d'importants remaniements chromosomiques (chromosomes en anneaux, des cassures de chromatides, des chromosomes dicentriques, etc...) [14], qui signent une instabilité chromosomique, et une augmentation du taux de mutation [14]. Ces données suggèrent que les bactéries productrices de colibactine

peuvent participer à l'initiation du cancer colorectal.

➤ **Épidémiologie des bactéries productrices de colibactine**

Des études épidémiologiques ont révélé la présence de l'îlot *pks* dans de nombreuses souches de *E. coli* mais également d'autres enterobactéries : *Klebsiella pneumoniae*, *Enterobacter aerogenes*, *Citrobacter koseri* [369,370]. La fonction biologique de la colibactine dans les bactéries n'est pas établie. Elle pourrait favoriser la persistance des bactéries en freinant le renouvellement des épithéliums intestinaux [26,371].

L'îlot génomique *pks* est observé chez les souches de *E. coli* appartenant au phylogroupe B2, comprenant notamment des souches responsables d'infections extra-intestinales comme les souches de références uropathogènes CFT073, 536 et UTI89 [13,26,350,369,370,372–374] ou responsables de méningites [26]. Dans ces souches, il est fréquemment associé la cyclomoduline CNF [349]. Les données épidémiologiques obtenues au laboratoire et par Arthur et collaborateurs [13] ont mis en évidence que les *E. coli* producteurs de colibactine sont plus fréquemment retrouvés dans les biopsies de tumeurs de patients atteints de cancer colorectal (66,7% et 50%) que dans les biopsies de patients atteints de la maladie de Crohn (40% et 32,2%) ou ne souffrant pas de maladie inflammatoire chronique de l'intestin ou de cancer (20,8% et 19%).

II- Présentation des résultats expérimentaux

L'ensemble des travaux présentés dans ce mémoire avait pour objectif d'explorer l'impact des *E. coli* porteurs de l'îlot *pks* sur la progression du cancer colorectal et d'initier une stratégie visant à contrôler les effets délétères de l'îlot *pks* sur les cellules de l'hôte.

- Une première étude présente l'activité pro-tumorale des *E. coli* porteurs de l'îlot *pks*. Elle montre que des lignées cellulaires épithéliales intestinales d'origine humaine secrètent des facteurs de croissance en réponse à l'infection et peuvent ainsi stimuler la croissance tumorale dans un modèle murin de xénogreffe.
- Une deuxième étude présente la caractérisation structurale et fonctionnelle de ClbP, une protéine de l'îlot *pks*. Il émerge de ces travaux que l'activité peptidase de ClbP est nécessaire à l'activité toxique de l'îlot *pks* et qu'elle appartient à un nouveau sous-groupe de peptidases contribuant à la bioactivation extracellulaire de composés de type NRP comme la colibactine.
- Une dernière étude rapporte l'identification de trois inhibiteurs nanomolaires de ClbP. Ces composés se fixent dans le site actif de ClbP et préviennent l'activité génotoxique et pro-tumorale des *E. coli* porteurs de l'îlot *pks*.

1. Les *Escherichia coli* producteurs de colibactine induisent la sénescence cellulaire et stimulent la prolifération cellulaire.

Publication n°1:

Bacterial genotoxin colibactin promotes tumorigenesis and senescence by impairing SUMOylation.

Antony Cougnoux, Guillaume Dalmasso, Ruben Martinez, Emmanuel Buc, Julien Delmas, Lucie Gibold, Denis Pezet, Harald Wodrich, Arlette Darfeuille-Michaud, Richard Bonnet.

Manuscrit soumis.

Les données du laboratoire confortées par des résultats récemment publiés par l'équipe du Professeur C. Jobin montrent que la prévalence des souches *E. coli* porteuses de l'îlot *pkS* est anormalement élevée dans les biopsies humaines issues de cancers colorectaux. Nous avons donc recherché l'impact des *E. coli pkS+* sur la progression du cancer colorectal. Pour ce faire, nous avons analysé la croissance de cellules épithéliales intestinales malignes en réponse à l'infection par un *E. coli* porteurs de l'îlot *pkS*, *in vitro* et dans un modèle murin de xénogreffe. Les infections ont été réalisées avec une souche *E. coli* de laboratoire (DH10B) possédant le chromosome bactérien artificiel pBeloBAC11 contenant ou non l'îlot *pkS* (*E. coli pkS+* et *E. coli pkS-* respectivement) afin de s'affranchir de l'effet des facteurs de virulence autres que la colibactine.

Nous avons observé une augmentation de la croissance tumorale dans les tumeurs infectées par *E. coli pkS+* comparativement à des tumeurs non infectées ou infectées par *E. coli pkS-* (**Figure 1, Publication 1**). *In vitro*, *E. coli pkS+* induisait un arrêt de la prolifération cellulaire. Ces résultats paradoxaux suggéraient un mécanisme indirecte stimulant la prolifération cellulaire. Le surnageant de cultures cellulaires infectées par *E. coli pkS+* et collecté à 5 jours post-infection (appelé milieu conditionné) était capable de stimuler la croissance de cellules non infectées contrairement au milieu conditionné par des cellules infectées par *E. coli pkS-*. Par ailleurs, des xénogreffes constituées de 20% de cellules infectées avec *E. coli pkS+* et de 80% de cellules non infectées, avaient une croissance plus rapide que ces mêmes xénogreffes réalisés avec des cellules infectées par des *E. coli pkS-* (**Figure 1, Publication 1**). L'ensemble de ces résultats suggérait que les cellules infectées par *E. coli pkS+* sécrètent des facteurs de croissance stimulant la prolifération des cellules tumorales.

Le mécanisme associé à la progression tumorale induite par *E. coli pks+* a été exploré *in vitro* et *in vivo*. Contrairement aux *E. coli pks-*, les *E. coli pks+* provoquaient l'induction d'une sénescence cellulaire prématurée (**Figure 2, Publication 1**) caractérisée par une mégaloctose, un blocage du cycle cellulaire, la présence d'une activité β -galactosidase lysosomiale à pH6 et l'activation des voies de signalisation p53/p21^{Cip} et Rb/E2F1. La sénescence cellulaire est généralement associée à un phénotype sécrétoire, qui dépend principalement de la nature des cellules et du stress induisant la sénescence [172,375]. Dans les cinq modèles cellulaires utilisés (HCT116, HCT8, HT-29, Int-407 et Caco2), nous avons observé un phénotype sécrétoire associé à la sénescence marqué par une expression augmentée du gène codant le facteur de croissance HGF et dans une moindre mesure de ceux codant le FGF et le GM-CSF (**Figure 2, Publication 1**). L'utilisation d'inhibiteurs de la voie HGF et d'anticorps neutralisants spécifiques prévenait la prolifération cellulaire et la progression tumorale associée à *E. coli pks+* (**Figure 2, Publication 1**). Ces données montraient ainsi que HGF est le médiateur principal de l'activité pro-proliférative associée à l'infection par *E. coli pks+* dans les modèles utilisés. Ainsi, les cellules rendues sénescents par *E. coli pks+* miment les fibroblastes associés aux tumeurs connus pour stimuler la croissance tumorale en sécrétant des facteurs de croissance [376,377].

La SUMOylation des protéines a récemment émergé comme un régulateur clé de la sénescence cellulaire en augmentant la stabilité et donc l'activité transcriptionnelle de p53 [140,157,162]. De façon intéressante, nous avons observé une modification du profil de SUMOylation des protéines en réponse à l'infection par *E. coli pks+* alors que cette altération n'était pas observée avec *E. coli pks-* (**Figure 3, Publication 1**). Cette modification s'accompagnait d'une augmentation de la fraction SUMOylée de p53 (**Figure 3, Publication 1**), qui est rapportée pour induire la sénescence [140]. Par ailleurs, nous avons observé une diminution concomitante de l'expression de l'enzyme SENP1, qui présente une activité préférentielle de déSUMOylation pouvant favoriser l'état SUMOylé de p53 [140,157]. Afin d'explorer plus avant le rôle de SENP1, nous avons étudié la sénescence cellulaire induite par *E. coli pks+* dans un contexte de SENP1 surexprimé. La surexpression de SENP1 permettait de prévenir les modifications de la SUMOylation des protéines, l'induction de la sénescence et l'effet pro-prolifératif du milieu conditionné avec *E. coli pks+*, contrairement à l'expression d'un mutant ponctuel inactif de SENP1 (**Figure 3, Publication 1**). Nous avons également montré que cette diminution d'expression de SENP1 était liée à l'augmentation d'expression du

microARN-20a-5p (**Figure 4, Publication 1**). Ce dernier était capable de se fixer à l'extrémité 3' non traduite de l'ARNm codant SENP1. La transfection dans les cellules du microARN 20a-5p reproduisait les effets de l'infection par *E. coli pks+* (sénescence, production de HGF, modification de la SUMOylation des protéines et stimulation de la prolifération cellulaire). Inversement une stratégie utilisant un ARN interférent ciblant le microARN-20a-5p prévenait les modifications de la SUMOylation des protéines, l'induction de la sénescence et l'effet pro-prolifératif associés à l'infection par *E. coli pks+* (**Figure 4, Publication 1**). L'expression du microARN-20a-5p est connue pour être sous le contrôle du facteur de transcription c-Myc qui est sur-exprimé dans les cellules présentant des dommages de l'ADN [201]. Nous avons montré que l'infection par *E. coli pks+* induisait l'accumulation de c-Myc, et sa fixation au promoteur du gène codant le microARN-20a-5p. De plus, une stratégie utilisant des ARN interférents ciblant c-Myc permettait de prévenir l'induction du gène codant le microARN-20a-5p, la sénescence et l'effet pro-prolifératif associée à l'infection par *E. coli pks+* (**Figure 4, Publication 1**). Au total, ces résultats révèlent la signalisation cellulaire responsable de l'induction de la sénescence en réponse à l'infection par *E. coli pks+*. Les dommages à l'ADN induits par ces derniers stimulent l'expression de c-Myc, qui se fixe au promoteur du gène codant le microARN-20a-5p et induit sa transcription. Le microARN-20a-5p réprime l'expression de SENP1 et en conséquence modifie la SUMOylation des protéines dont p53 qui est un inducteur majeur de la sénescence cellulaire.

Afin de confronter les résultats obtenus *in vitro* et dans le modèle murin de xélogreffe, nous avons collecté des biopsies de cancers colorectaux colonisées ou non par *E. coli pks+* (n=6/groupe) puis comparé les niveaux d'expression des gènes codant le microARN-20a-5p, SENP1 et HGF (**Figure 5, Publication 1**). De façon intéressante, l'expression des ARNm codant HGF étaient significativement plus élevée dans les tumeurs colonisées par *E. coli pks+*. Cette augmentation était associée à une augmentation de l'expression du microARN-20a-5p et une diminution de SENP1. Les profils d'expression de HGF, SENP1 et du microARN-20a-5p dans les tumeurs des patients colonisés par *E. coli pks+* étaient donc comparables aux profils obtenus lors des infections réalisées *in vitro* ou en modèle murin.

En conclusion, nos résultats révèlent un nouveau paradigme de la carcinogenèse colorectale dans lequel la sénescence cellulaire, induite par les bactéries productrices de colibactine, jouent un rôle essentiel. Nous proposons que ces bactéries modulent le microenvironnement tumoral en favorisant l'émergence de cellules sénescents, qui stimulent la

croissance tumorale par la sécrétion de facteurs de croissance, notamment HGF (**Figure 6, Publication 1**). Ce modèle est en accord avec les données récemment publiées par Arthur et collaborateurs qui montrent un effet carcinogène de *E. coli pks+* indépendant de l'inflammation dans un modèle murin IL10^{-/-} [13]. Une stratégie visant à bloquer la production de colibactine pourrait donc être envisagée pour inhiber son activité pro-tumorale et génotoxique.

Bacterial genotoxin colibactin promotes tumourigenesis by impairing SUMOylation

Antony Cougnoux^{1,2,*}, Guillaume Dalmasso^{1,2,*}, Ruben Martinez³, Emmanuel Buc^{1,2,4}, Julien Delmas^{1,2,4}, Lucie Gibold^{1,2,4}, Denis Pezet^{1,2,4}, Harald Wodrich³, Arlette Darfeuille-Michaud^{1,2,4}, Richard Bonnet^{1,2,4}.

*, These authors contributed equally to this work

1. Clermont Université, UMR 1071 Inserm/Université d'Auvergne, 63000 Clermont-Ferrand, France
2. INRA, USC 2018, 63000 Clermont-Ferrand, France
3. Microbiologie Fondamentale et Pathogénicité, CNRS UMR 5234, Université Bordeaux SEGALEN, Bordeaux, France
4. Centre Hospitalier Universitaire, 63000 Clermont-Ferrand, France

Corresponding authors: Richard Bonnet

Address: CHU, Laboratoire de Bacteriologie

58 rue Montalembert

63000 Clermont-Ferrand

France

e-mail: rbonnet@chu-clermontferrand.fr

Tel: +33 (0)4 73 754 920

Short title: colibactin-producing *E. coli* induce tumour growth.

Abstract:

Microbiota is suspected to be involved in colorectal cancer. Mucosa-associated *Escherichia coli* isolated from colorectal cancer frequently harbour the *pks* genomic island (*pks+* *E. coli*) responsible for the synthesis of the genotoxin colibactin. Here we report that a transient contact of few malignant cells with colibactin-producing *E. coli* increases tumour growth in a mouse xenograft model. Growth is sustained by cellular senescence, which is accompanied by the production of HGF. We demonstrate that cellular senescence is a consequence of *pks+* *E. coli*-induced alteration of protein SUMOylation in eukaryotic cells. The underlying mechanisms involve the expression of c-Myc and microRNA-20a-5p, which targets SENP1, a key protein in the regulation of the SUMOylation. These results are consistent with the expression levels of SENP1, microRNA-20a-5p and HGF found in human colorectal cancers colonized by *pks+* *E. coli*. Overall, these data reveal a new paradigm for carcinogenesis in which colibactin-induced senescence plays a critical role.

Keywords: colorectal cancer, SUMO, microRNA, SENP1, colibactin, *Escherichia coli*, HGF, growth factor

Cancer is defined as uncontrolled cell proliferation caused by accumulated genetic and epigenetic mutations¹⁻³. The origin of these mutations can be multifactorial and overall 20% of cancers worldwide are estimated to be attributable to infections^{4,5}. Many cancers occur in tissues with a high exposure to microbiota, especially colorectal cancers (CRCs), the fourth most common cancer with more than one million cases yearly⁶. The high density of microorganisms observed in the colon, (10^{13} commensal bacteria colonizing the colon vs from 10^2 for duodenum to 10^8 for the distal ileum), is associated with a greater risk of cancer than in the small intestine⁵. The role of microbiota in colon cancer is documented by the fact that mice genetically susceptible to CRC, or mice with chemical-induced chronic inflammation, develop significantly fewer tumours under germ-free conditions than when they harbour conventional microbiota^{7,8}. The role of microbiota in the development of CRC could be related to the pro-carcinogenic features of intestinal bacteria which may initiate cancer development and/or have a competitive advantage in the tumour microenvironment that allows them to play a role in disease progression^{9,10}. Few putative pro-carcinogenesis bacteria have been identified. *Enterococcus faecalis* species produce superoxide able to induce DNA damage and chromosomal instability in colonic epithelial cells¹¹. Enterotoxigenic *Bacteroides fragilis* produce the fragilysin toxin, which is genotoxic¹⁰, causing cell proliferation by cleavage of the tumour suppression factor E-cadherin¹², and increasing tumorigenesis by triggering production of interleukin-17¹³. Finally, colonic malignancies allow *Streptococcus gallolyticus* subsp. *gallolyticus* to colonize established colorectal tumours, leading to tumour development by induction of the pro-inflammatory COX2 pathway¹⁴.

Escherichia coli are the predominant aero-anaerobic Gram-negative bacteria in the normal intestinal microbiota. As a commensal bacterium, *E. coli* coexists with its mammalian host in good harmony, promotes normal intestinal homeostasis and rarely causes disease. However, linked to the presence of pathogenicity island(s), some virulent *E. coli* can colonize the human gastrointestinal tract and induce diseases. Regarding CRC, mucosa-associated

Escherichia coli, which are more frequently identified in colon tissue from patients with adenocarcinomas than in controls, can have pro-carcinogenic features^{15,16} by promoting down-regulation of DNA mismatch repair proteins¹⁷. In addition, various toxins produced by pathogenic *E. coli* have carcinogenic potential such as cytotoxic necrotising factor (CNF), cycle-inhibiting factor (CIF) and cytolethal distending toxins (CDT)^{9,18}. Certain *Escherichia coli* strains harbour the *pks* island, which encodes enzymes responsible for the synthesis of a genotoxin called colibactin¹⁹. Colibactin induces double-strand DNA breaks, cell cycle arrest and megalocytosis¹⁹. Colibactin-induced DNA damage was associated with mutations and chromosomal instability²⁰; lesions frequently observed in cancers. Interestingly, *pks*-harbouring *E. coli* (*pks*+ *E. coli*) strains were very recently reported to be found in 50 to 60% of human colorectal tumours *versus* 20% in controls^{21,22} and, as shown using an AOM/II10^{-/-} mouse model, *pks*+ *E. coli* have a carcinogenic effect independently of the inflammation²².

In the present study, we deciphered the mechanism of *pks*+ *E. coli*-induced carcinogenesis. We report that *pks*+ *E. coli* infection induces cellular senescence, marked by high increase in production of growth factors that promote proliferation of uninfected cells and subsequently tumour growth. We also show that colibactin-induced cellular senescence is a direct consequence of protein SUMOylation alteration.

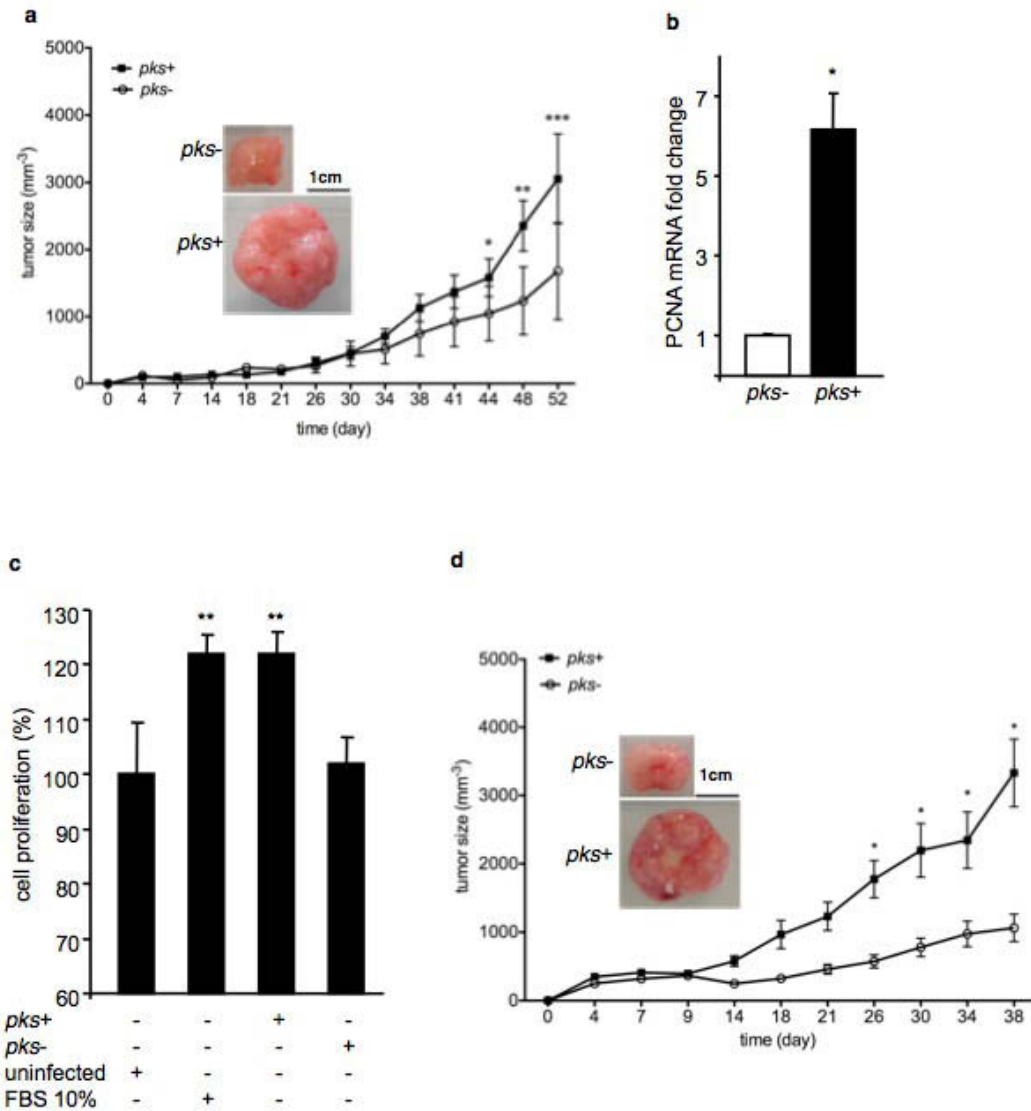


Figure 1. Infection with *pks+* *E. coli* promotes tumour growth and cell proliferation. **a**, 2×10^6 HCT116 cells were subcutaneously injected to nude mice and immediately infected for 3 hours with *pks+* *E. coli* or *pks-* *E. coli*. Tumour size was monitored during the indicated period ($n = 5/\text{group}$; $*P < 0.05$; $**P < 0.01$, $***P < 0.001$). Representative photographs of tumours are presented (bar scale = 1 cm). **b**, qRT-PCR analyses of PCNA expression levels in tumours 52-days post-infection ($n = 5/\text{group}$; $*P < 0.05$). **c**, Conditioned media derived from *pks+* *E. coli*- or *pks-* *E. coli*-infected HCT116 cells were collected 5-days post-infection and were used to stimulate the growth of non-infected HCT116 cells. Cellular growth was assessed by a MTT-based cell proliferation assay ($n = 6/\text{group}$; $**P < 0.01$). **d**, 0.4×10^6 HCT116 cells collected 5-days post-*pks+* or *pks-* *E. coli* infection, were mixed with 1.6×10^6 non-infected HCT116 cells and subcutaneously injected to nude mice. Tumour size was monitored during the indicated period ($n = 10/\text{group}$; $*P < 0.05$). Representative photographs of tumours are presented (bar scale = 1 cm).

Results

Colibactin-producing E. coli infection increases tumour growth

Colibactin-producing *E. coli* induce DNA damage and chromosomal rearrangements^{19,20}, we therefore investigated whether and how such bacteria may play a role in CRC. Intestinal epithelial cells HCT116 were subcutaneously injected to nude mice and immediately infected for 3 hours with DH10B *E. coli* strain producing colibactin (*pks+*) or not (*pks-*). A single transient 3 hours exposure to *pks+* *E. coli* significantly increased tumour growth (Figure 1a). Accordingly, an increased level of proliferating cell nuclear antigen (PCNA)-encoding mRNA was observed in tumours infected with *pks+* *E. coli* (Figure 1b). To further investigate whether *pks+* *E. coli* infection was able to directly or indirectly promote cellular proliferation, HCT116 cells were infected *in vitro* with *pks+* or *pks-* *E. coli* for 3 hours and cellular proliferation was assessed at 24 hours post-infection. We observed a ~20% drop in cell proliferation of *pks+* *E. coli*-infected cells compared to *pks-* *E. coli*-infected cells (supplementary Figure 1). This decrease could be related to colibactin-induced cell cycle arrest¹⁹. The decrease in cell proliferation observed *in vitro* and the pro-proliferation effect observed in the xenograft mouse model in response to *pks+* *E. coli* infection suggested an indirect effect of colibactin on tumour growth *in vivo*. We therefore assessed the ability of conditioned medium (CM) derived from HCT116 cells infected with *pks+* *E. coli* to promote cell proliferation. The CMs were obtained by collecting serum-free supernatant of HCT116 cells at days 1, 3 or 5 post-infection. As shown in Figure 1c and supplementary Figure 2, cells incubated with CM derived from *pks+* *E. coli*-infected cell 5-days post-infection proliferate as fast as cells incubated with a medium containing 10% FBS. Similar results were obtained with four other intestinal epithelial cell lines (Caco-2, HT-29, HCT8 and Int-407; supplementary Figure 3). However, CMs derived from *pks-* *E. coli*-infected cells did not promote cellular proliferation (Figure 1c and supplementary Figure 2).

To investigate whether a few *pks+* *E. coli*-infected cells were able to promote tumour growth *in vivo*, nude mice were subcutaneously injected with a combination of 20% HCT116

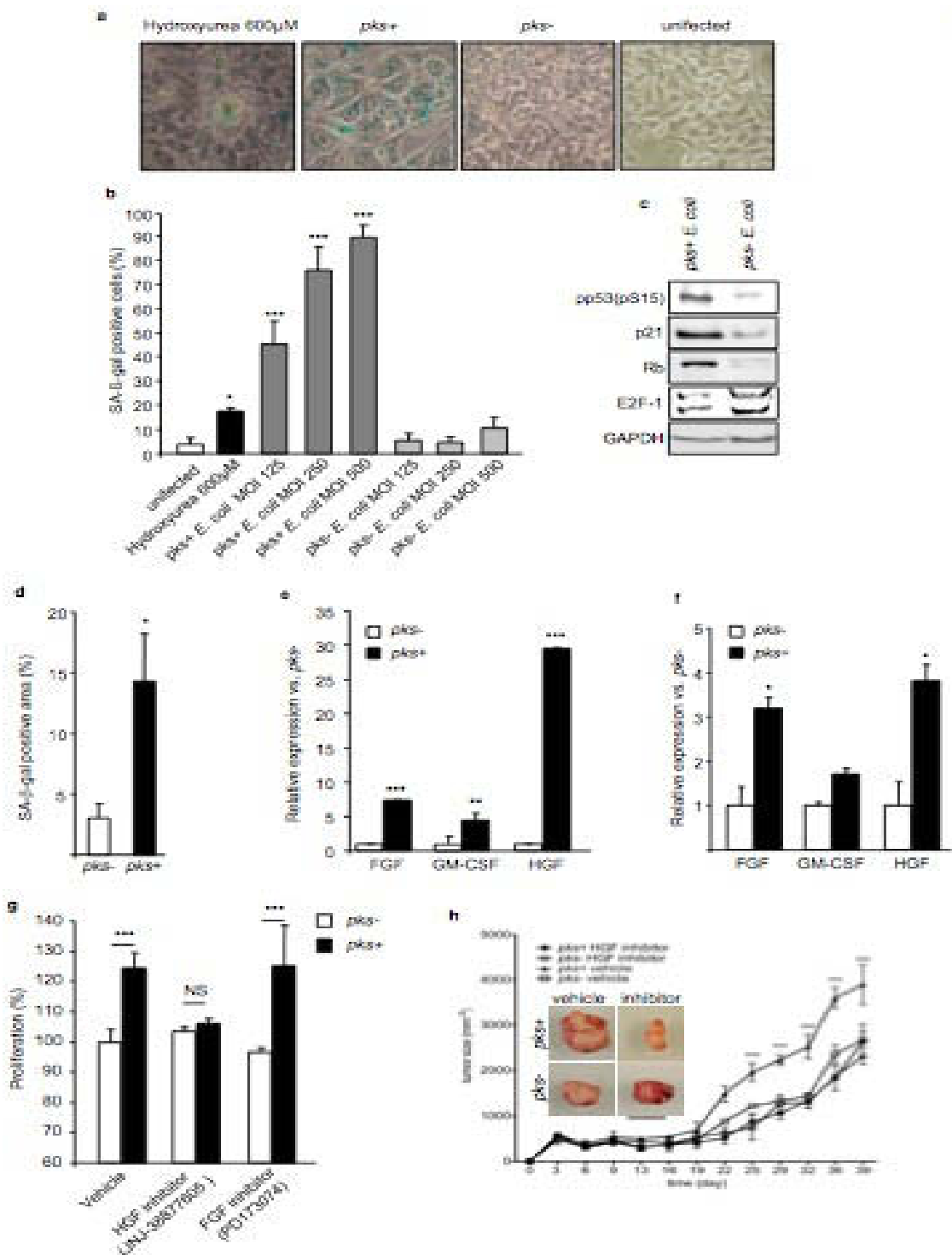


Figure 2. Colibactin-producing *E. coli* induce cellular senescence. **a**, HCT116 cells were infected with *pks+* *E. coli* or *pks-* *E. coli* and SA-β-gal positive cells were assessed 5 days post-infection. 600µM hydroxyurea was used as a positive control. **b**, Percentage of SA-β-gal positive cells determined at 5-days post-infection using different MOI (Multiplicity Of Infection). Values represent means ± S.E.M. (n = 6/group; **P*<0.05, ****P*<0.001 vs uninfected cells). **c**, Whole-cell lysates were analyzed by immunoblotting experiments 5 days post-infection. **d**, 2 × 10⁶ HCT116 cells were subcutaneously injected to nude mice and immediately infected for 3 hours with *pks+* *E. coli* or *pks-* *E. coli*. Senescent cells were revealed by detection of SA-β-gal positive cells 10 days post-infection (n = 5/group; **P*<0.05). **e and f**, Growth factors mRNA expression levels were assessed by qRT-PCR in HCT116 cells 5-days post-infection (**e**) or in mice tumours (**f**) (n = 5/group; **P*<0.05; ***P*<0.01; ****P*<0.001). **g**, HCT116 cells were cultured in presence of conditioned media derived from *pks+* *E. coli* or *pks-* *E. coli*-infected HCT116 cells and supplemented with growth factor inhibitors. Cellular growth was assessed by a MTT-based proliferation assay (n = 8/group; ****P*<0.001; NS, not significant vs *pks-*). Data are presented as means ± S.E.M. **h**, 0.4 × 10⁶ HCT116 cells collected 5-days post-*pks+* or *pks-* *E. coli* infection, were mixed with 1.6 × 10⁶ non-infected HCT116 cells and subcutaneously injected to nude mice. Mice received twice a week 3.4 µg/kg HGF inhibitor or vehicle. Tumours size was monitored during the indicated period (n=5/group; ****P*<0.001 vs other groups). Representative photographs of tumours are presented (bar scale = 1 cm).

cells collected at 5-days post-*pks+* or *pks-* *E. coli* infection and 80% of uninfected HCT116 cells. Xenografts containing *pks+* *E. coli*-infected cells developed more rapidly than those containing *pks-* *E. coli*-infected cells (Figure 1d). These results indicate that few cells infected with *pks+* *E. coli* promote tumour growth.

Colibactin-producing E. coli induce senescence of intestinal epithelial cells

Colibactin produced by *pks+* *E. coli* is known to induce megalocytosis and cell cycle arrest¹⁹, which are features of cellular senescence²³. Senescence-associated β -galactosidase (SA- β -gal) activity was therefore assessed in *pks+* *E. coli*-infected cells. We found that *pks+* *E. coli* infection induced a time-dependent increase in the number of SA- β -Gal positive megalocytes (Figure 2a and supplementary Figure 4a). Similar results were obtained with various intestinal epithelial cell lines (supplementary Figure 4b). Quantification of SA- β -gal positive cells showed that *pks+* *E. coli* induced cellular senescence in an infection dose-dependent manner (Figure 2b). In contrast, the number of SA- β -Gal positive cells induced by *pks-* *E. coli* was similar to that of non-infected cells, indicating that the induction of senescence was due to colibactin synthesis (Figure 2b). We also observed that *pks+* *E. coli*-infected cells exhibited at day 5 post-infection an accumulation of phospho-p53, p21^{Cip}, and Rb as well as a decrease in E2F-1 expression (Figure 2c); a protein expression profile characteristic of senescent cells²⁴, confirming that *pks+* *E. coli* induce senescence of human intestinal epithelial cells. To investigate whether *pks+* *E. coli* promote cellular senescence *in vivo* at early time points in tumour development, nude mice were subcutaneously injected with HCT116 cells infected with *pks+* or *pks-* *E. coli* and tumours were analyzed at day 10 post-implantation. Tumour sizes were comparable at this early time point in animals injected with cells infected by *pks+* and *pks-* *E. coli* (Figure 1a). However, tumours derived from *pks+* *E. coli*-infected HCT116 cells exhibited a ~3-fold increase in the percentage of senescent cells compared to tumours derived from *pks-* *E. coli*-infected cells (Figure 2d). Senescent cells are metabolically active and

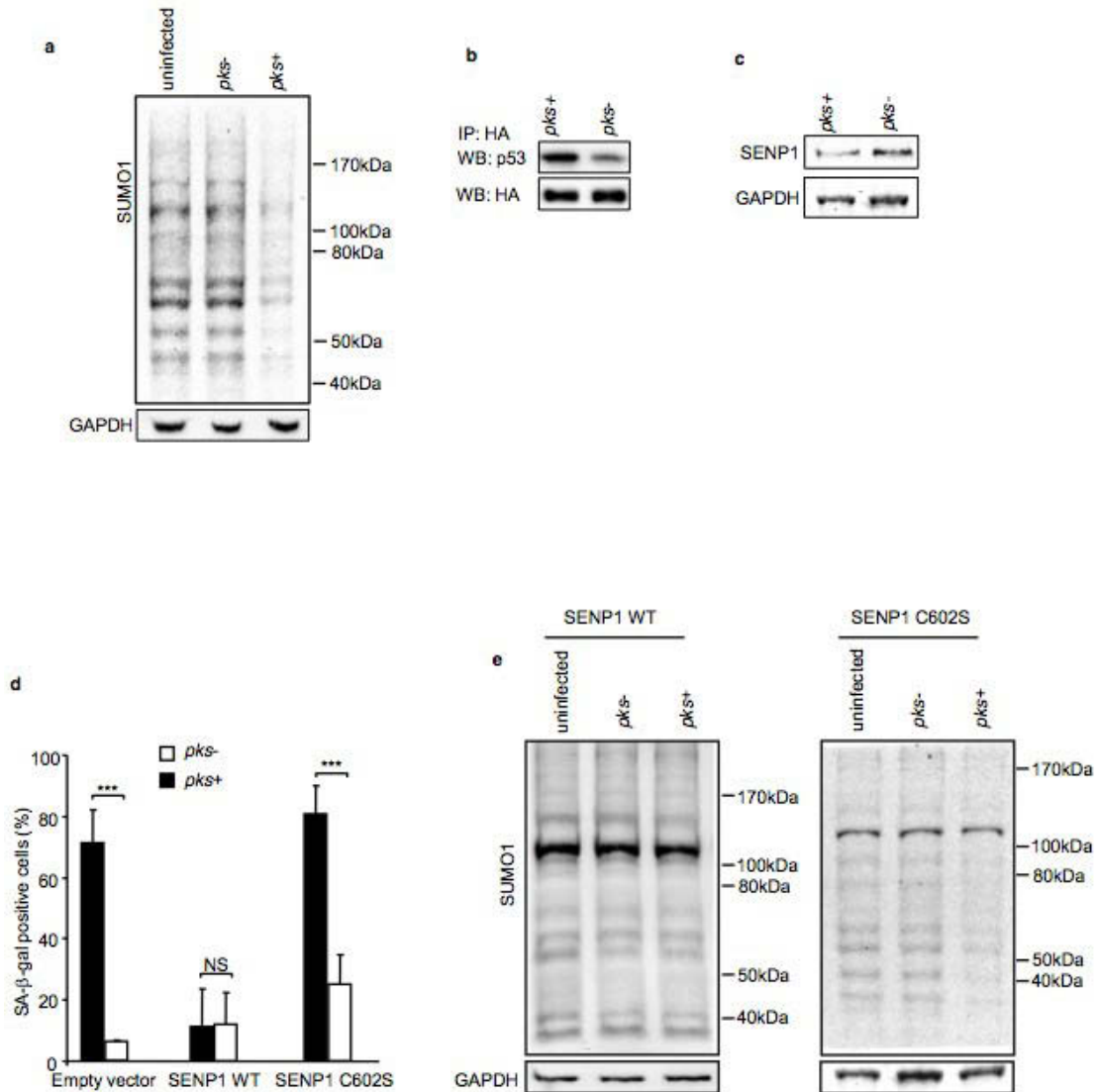


Figure 3. SUMOylation controls senescence upon infection with *pks*⁺ *E. coli*. **a**, Comparison of SUMO-conjugated protein pattern from HCT116 cells 5-days post-infection with *pks*⁺ *E. coli* or *pks*⁻ *E. coli*. **b**, HCT116 cells transiently transfected with a vector encoding SUMO1 HA-tagged protein, were lysed 5-days post-infection and then immunoprecipitated using an antibody specific for HA. The amount of co-immunoprecipitated p53 was examined by Western blot. **c**, Immunoblot analysis of SENP1 expression in HCT116 cells 5-days post-infection. **d** and **e**, HCT116 cells transiently transfected with an empty vector or a vector encoding SENP1 (SENP1 WT) or an inactive mutant (SENP1 C602S), were infected with *pks*⁺ *E. coli* or *pks*⁻ *E. coli*. **d**, Bar graph represents the percentage of SA-β-gal positive cells 5-days post-infection. Data are presented as means ± S.E.M (n = 6/group; NS, not significant; ****P*<0.001), **e**, SUMO1-conjugated protein patterns were assessed 5-days post-infection.

usually acquire a senescence-associated secretory phenotype, which can be characterized by the production of growth factors^{25,23}. FGF, GM-CSF and HGF expression levels were therefore investigated at day 5 post-infection. We observed that *pks+* *E. coli*-infected cells expressed significantly higher levels of growth factors, in particular HGF, than *pks-* *E. coli*-infected HCT116 cells (Figure 2e). Similar results were obtained with all intestinal epithelial cell lines tested (supplementary Figure 5). Interestingly these growth factors presented also an increased expression in tumours of animals in which the xenograft was infected with *pks+* *E. coli* (figure 2f).

To identify the relevant secreted mediators involved in cell proliferation, HCT116 cells were cultured with conditioned media supplemented with growth factors inhibitors. Interestingly, HGF pathway inhibitor abrogated *in vitro* the pro-proliferative activity of conditioned media derived from *pks+* *E. coli*-infected cells, in contrast to vehicle or FGF inhibitor (Figure 2g). Similar results were obtained using neutralizing antibodies (Supplementary Figure 6). As observed *in vitro*, HGF inhibitor significantly blocked the growth of xenograft composed of *pks+* *E. coli*-infected cell combined with uninfected cells (Figure 2h). These data showed that the pro-proliferative effect of *pks+* *E. coli*-induced senescent cells was mainly mediated by HGF.

SENP1 down-expression and protein SUMOylation modification are key features of pks+ E. coli-induced senescence

Protein SUMOylation has recently emerged as a key regulator of cellular senescence by stimulating the transcriptional activity of p53²⁶. Interestingly, in our study *pks+* *E. coli*-infected HCT116 cells displayed a modified pattern of SUMO-conjugated proteins, compared with cells infected with *pks-* *E. coli* or uninfected cells (Figure 3a). In addition, *pks+* *E. coli* induced an increase in p53 SUMOylation (Figure 3b), which is known to drive cellular senescence²⁷, and a significant decreased expression of SENP1, a key protein involved in the control of the SUMOylation process²⁷ (Figure 3c). To confirm the role of SENP1 in *pks+* *E. coli*-induced

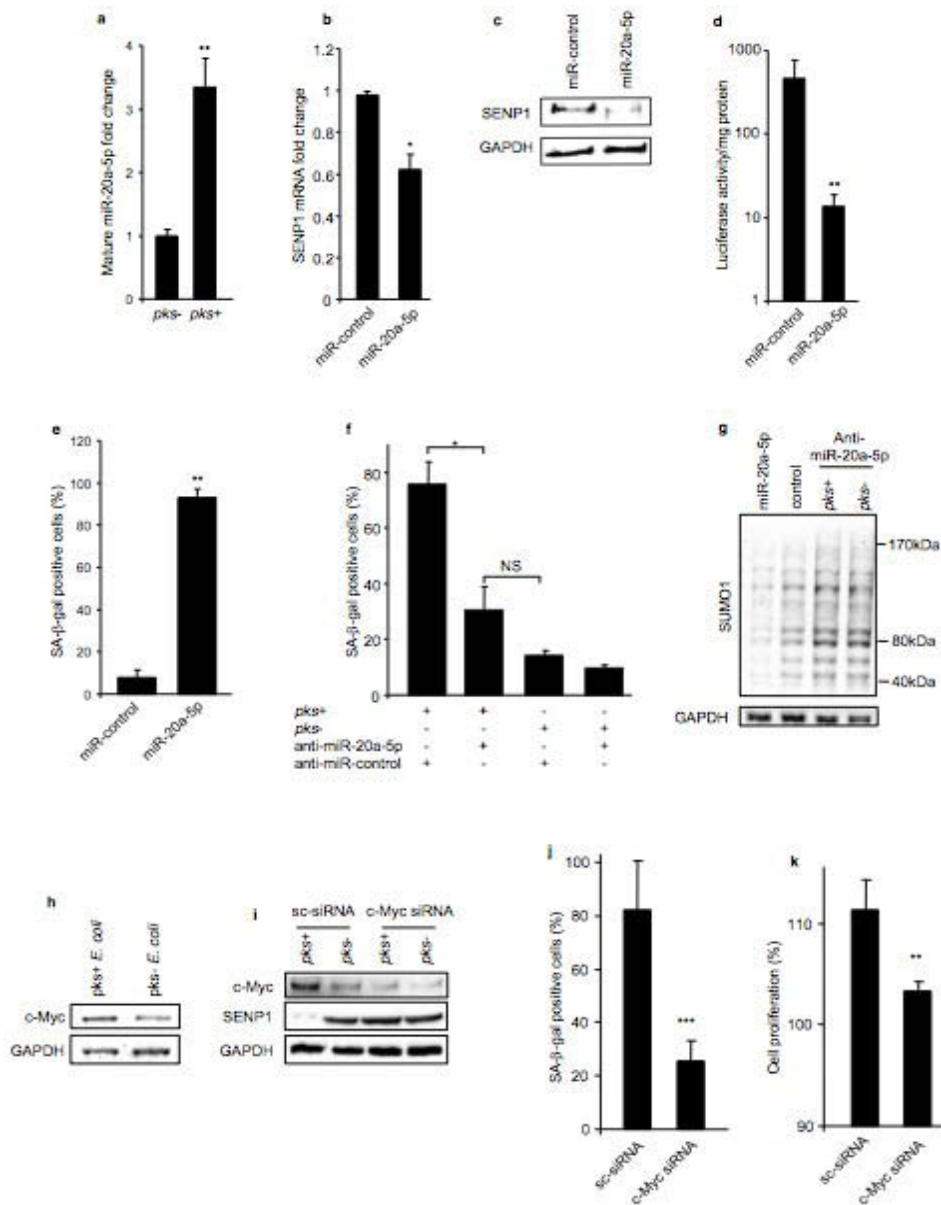


Figure 4. MiR-20a-5p represses SENP1 expression and controls the senescence induced by *pks+* *E. coli*. **a**, Mature forms of miR-20a-5p were quantified by qRT-PCR in HCT116 cells 5-days post-infection. **b**, **c**, HCT116 cells were transfected with precursors of miR-20a-5p or miR-control, and SENP1 expression was assessed by **b**, qRT-PCR and **c**, Western blot. **d**, HCT116 cells were transfected with a luciferase vector encoding the *SENP1* mRNA 3'-UTR in the presence of miR-20a-5p or miR-control. Luciferase activity was determined. **e**, HCT116 cells were transfected with precursor of miR-20a-5p or miR-control and senescence was assessed at day 5 by detecting SA-β-gal positive cells. **f**, HCT116 cells were transfected with anti-miR-20a-5p or anti-miR-control and infected with *pks+* *E. coli* or *pks-* *E. coli*. 5-days post-infection senescent cells were assessed by detecting SA-β-gal positive cells. **g**, HCT116 cells transiently transfected with anti-miR-20a-5p were infected with *pks+* *E. coli* or *pks-* *E. coli* and SUMO-conjugated protein patterns were assessed 5-days post-infection. Effect of miR-20a-5p on SUMO-conjugated protein patterns is also presented. **h**, HCT116 cells were infected with *pks+* or *pks-* *E. coli* and c-Myc expression was analyzed by Western blot at day 5 post-infection. **i**, HCT116 cells transfected with c-Myc siRNA or scramble (sc) siRNA were infected with *pks+* or *pks-* *E. coli* and c-Myc and SENP1 expression levels were assessed by Western blot. **j**, HCT116 cells transfected with c-Myc siRNA or scramble (sc) siRNA were infected with *pks+* *E. coli* and SA-β-gal positive cells were determined. **k**, Conditioned media were derived from HCT116 cells transfected with c-Myc siRNA or scramble (sc) siRNA and infected with *pks+* *E. coli*. Conditioned media were used to culture HCT116 cells and cellular growth was assessed by MTT assay. All data are presented as means ± S.E.M. n = 6/group; NS, not significant; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

senescence, HCT116 cells were transfected with a vector encoding SENP1²⁸. As shown in Figure 3d, over-expression of SENP1 significantly decreased the number of senescent cells induced by *pks+* *E. coli* infection to a level similar to that observed with *pks-* *E. coli*-infected cells. This effect was dependent on SENP1 enzymatic activity since the expression of an inactive mutant of SENP1 did not protect against *pks+* *E. coli*-induced senescence (Figure 3d). In addition, in cells with over-expression of wild-type SENP1, unlike in cells over-expressing an inactive SENP1, decrease in SUMO-conjugated proteins upon infection with *pks+* *E. coli* was no longer observed (Figure 3e). During these experiments, the genotoxic effect of *pks+* *E. coli* was not affected, as assessed by S139-phosphorylation of H2Ax, highlighting the key role of SENP1 in senescent process (Supplementary Figure 7). Of note, conditioned media derived from *pks+* *E. coli*-infected cells over-expressing functional SENP1 did not promote HCT116 cell proliferation (Supplementary Figure 8). All together, these data show that SENP1 down-expression and subsequent SUMOylation modifications are key features in *pks+* *E. coli*-induced cell senescence.

Colibactin-producing E. coli decrease SENP1 expression via miR-20a-5p and c-Myc

Among the microRNAs deregulated during senescence²⁹, *in silico* predictions revealed that miR-20a-5p potentially targets SENP1 (Supplementary Figure 9). Interestingly, miR-20a-5p expression was significantly up-regulated in *pks+* *E. coli* infected cells compare to *pks-* *E. coli* infected cells (Figure 4a). Furthermore, transfection of cells with mature miR-20a-5p decreased SENP1 expression at both mRNA and protein levels (Figure 4b,c). The *SENP1* mRNA 3'-UTR was cloned downstream of the luciferase gene and this construct was transfected into cells in the presence of miR-20a-5p or miR-control. As shown in Figure 4d, miR-20a-5p significantly repressed luciferase activity, confirming the binding of miR-20a-5p to the *SENP1* mRNA 3'-UTR. Altogether these results indicate that *pks+* *E. coli* are able to up-regulate the expression of miR-20a-5p, which in turn down-regulates SENP1 expression by directly targeting *SENP1* mRNA 3'-

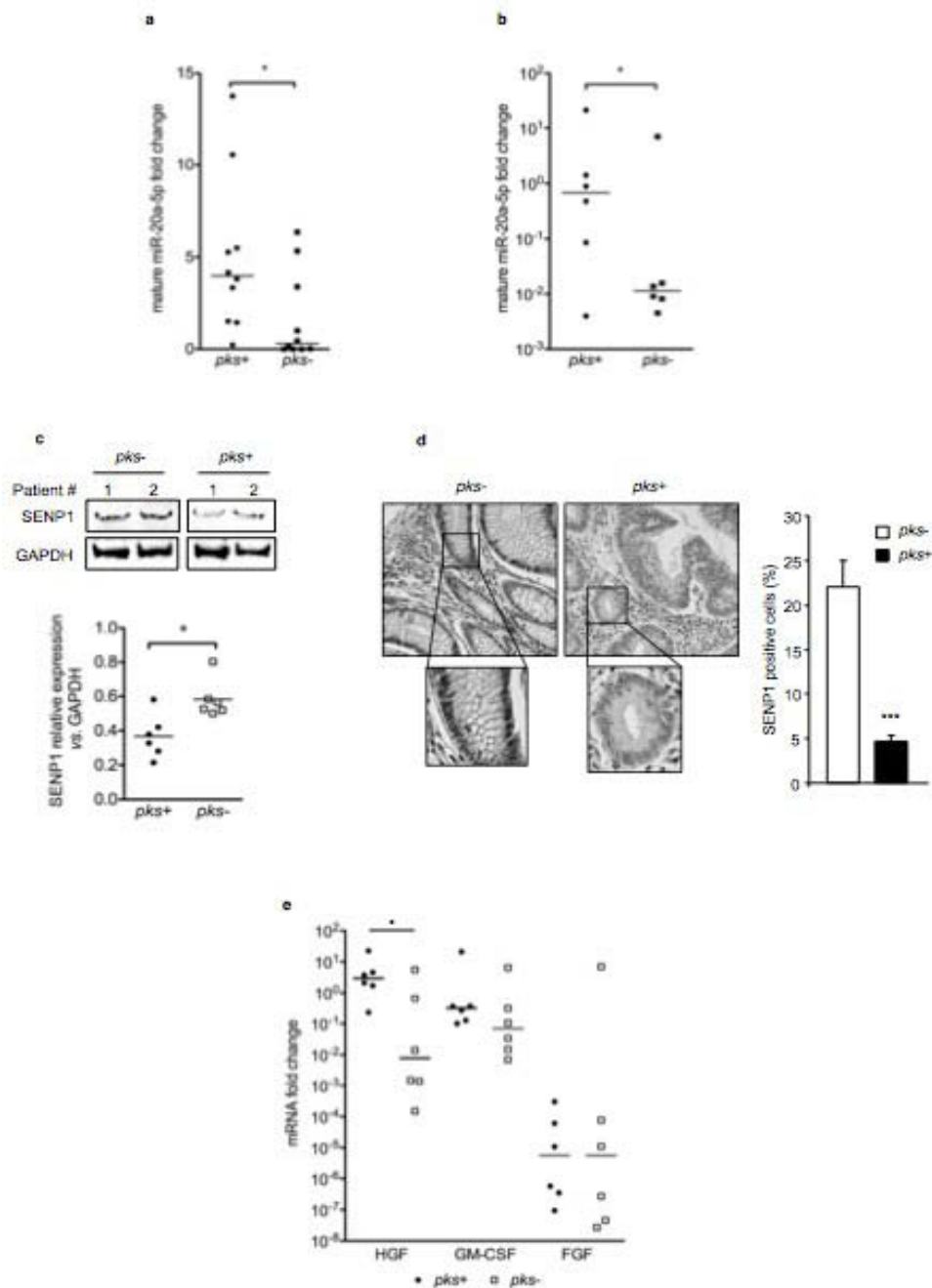


Figure 5. Modified expression pattern of miR-20a-5p, SENP1 and growth factors in human *pks*⁺ *E. coli*-colonized colon tumours. **a**, HCT116 cells were infected with clinical strains of *E. coli* which do not possess any genotoxin-encoding genes or which possess only the *pks* island. Mature miR-20a-5p expression levels were assessed by qRT-PCR 5 days post-infection. (n = 10/group; **P*<0.05). **b**, **c**, **d**, and **e**, miR-20a-5p, SENP1 and growth factors expression levels assessed in human colonic tumours colonized by a single clone of *E. coli* which does not possess any genotoxin-encoding genes or which possess only the *pks* island (n= 6 patients/group, **P*<0.05, ****P*<0.001). All data are presented as means ± S.E.M. (**b**) Mature miR-20a-5p expression levels were assessed by qRT-PCR. (**c**) Western blot is representative of SENP1 expression in tissues isolated from colorectal biopsies and the bar graph on the bottom represents the densitometric quantification of SENP1 expression in the entire tissue collection. (**d**) Immunohistochemical staining of SENP1 expression in human biopsies. Bar graph represents the quantification of SENP1 positive cells (**e**) Growth factor (HGF, GM-CSF and FGF) mRNA expression levels were assessed by qRT-PCR from human biopsies.

UTR.

We next investigated the role of miR-20a-5p in senescence and cell proliferation. HCT116 cells were transfected with mature miR-20a-5p or a miR-control and SA- β -gal positive cells were counted 5-days post-transfection. In cells transfected with miR-20a-5p, a significant increase in the number of senescent cells was observed compared to cells transfected with miR-control (Figure 4e). In addition, the transfection of an anti-miR-20a-5p in HCT116 cells significantly decreased the percentage of senescent cells emerging upon *pks+* *E. coli* infection in contrast to the transfection of an anti-miR control (Figure 4f). As observed after *pks+* *E. coli* infection, miR-20a-5p induced a significant increase of HGF expression and cell proliferation (supplementary Figure 10a,b); while anti-miR-20a-5p abolished *pks+* *E. coli*-induced HGF expression and the increase in cell proliferation (supplementary Figure 10c,d). These results show that *pks+* *E. coli* induce senescence *via* miR-20a-5p and subsequently, can stimulate cell proliferation.

Since miR-20a-5p controls the expression of SENP1, we investigated its impact on protein SUMOylation. Transfection of mature miR-20a-5p induced modifications in the SUMOylation profile similar to those observed after *pks+* *E. coli* infection (Figure 4g). In contrast, transfection of anti-miR-20a-5p abrogated modifications in the SUMOylation profile induced by *pks+* *E. coli* (Figure 4g). *Pks+* *E. coli* therefore increase miR-20a-5p expression, leading to a decrease in SENP1 expression level that results in modification of proteins SUMOylation.

MiR-20a-5p is known to be regulated by the transcription factor c-Myc⁴¹ and we observed that infection of HCT116 cells with *pks+* *E. coli* led to an increase in c-Myc expression (Figure 4h) and in its binding to miR-20a-5p promoter (Supplementary Figure 11). The transfection of *pks+* *E. coli*-infected cells with c-Myc siRNA strongly reduced the expression of c-Myc and miR20a-5p, increased the expression of SENP1, decreased the number of senescent cells and consequently abolished pro-proliferative effect (Figure 4i,j,k), confirming that *pks+* bacteria regulate miR-20a-5p expression *via* c-Myc.

Modified expression of miR-20a-5p, SENP1 and growth factors in human colon tumours colonized by pks+ E. coli

Clinical *pks+* and *pks-* *E. coli* strains were isolated from human CRC biopsies²¹. The *pks+* *E. coli* strains induced in HCT116 cells a significant increase in miR-20a-5p levels compared to *pks-* *E. coli* strains (Figure 5a). To confirm the clinical relevance of our findings, we also assessed miR-20a-5p expression level in colorectal tumours colonized by a single clone of *E. coli* producing or not colibactin as the only genotoxin. Overall, the expression level of miR-20a-5p was significantly higher in tumours colonized by *pks+* *E. coli* than in those colonized by *pks-* *E. coli* (Figure 5b). In addition, SENP1 expression was significantly decreased in tumours colonized with *pks+* *E. coli* compared to tumours colonized with *pks-* *E. coli*, as shown by Western blot (Figure 5c) or by enumerating SENP1 positive cells after immunohistochemical staining (figure 5d). Finally, analysis of the expression of growth factors indicated that HGF was significantly increased in tumours colonized with *pks+* *E. coli* compared to tumours harbouring *pks-* *E. coli* (Figure 5e), indicating that colibactin-producing bacteria can promote tumour growth by inducing the expression of HGF, as observed *in vitro*.

Discussion

Cancer research is predominantly focused on the identification of multistep genetic changes, which are intimately associated with the induction and progression of tumours, however recent emerging evidences have raised intestinal microbiota as a promoter of colonic tumour formation⁵. *E. coli* are consistent members of human microbiota and are classified in four major phylogroups (A, B1, B2 and D), each group being more specifically associated with certain ecological niches. The *pks* island encoding colibactin is hosted by B2 *E. coli* commonly found the human microbiota³⁰. The incidence of *E. coli* belonging to the phylogroup B2 has risen during the last decade in colonic microbiota of human beings living in high-income countries, likewise the incidence of CRC, suggesting a link between those bacteria and CRC^{6,31}. In addition, *pks+* *E. coli* were recently shown to be overrepresented in human colorectal tumours^{21,22} and to have a carcinogenic effect²².

In the present study, we show that colibactin-producing *E. coli* contribute to the emergence of a senescent microenvironment which improves tumour progression *via* growth factors production. Infection of tumours with colibactin-producing *E. coli* increases tumour growth in mouse. Interestingly, 3 hours contact between tumours and colibactin-producing *E. coli* is sufficient to stimulate tumour growth, suggesting that even a transient colonization might affect the fate of tumours. This new paradigm for bacteria-enhanced tumour progression may be highly relevant for CRC, since *pks+* *E. coli* colonize 50 to 60% of colorectal tumours^{21,22}.

A link between tumour persistence and senescence has been reported in human melanoma³², and is compatible with the hierarchical cellular organization of malignancy^{33,34}. The presence of senescent cells in the tumour microenvironment, which provide fuel for other cells to proliferate, is in agreement with the hierarchical model of cancer development. We observed that the colibactin-producing *E. coli*-induced senescent cells are metabolically active and have acquired a senescence-associated secretory phenotype characterized by the production of growth factors. Interestingly, the production of HGF, which is recognized as a key determinant of

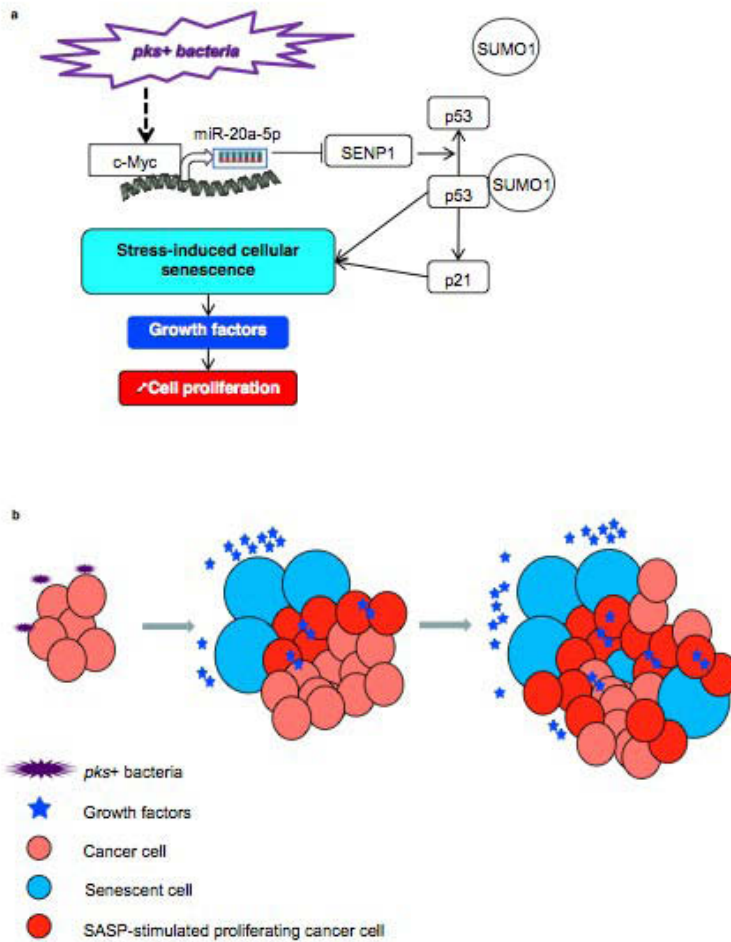


Figure 6: Proposed model for the contribution of colibactin-producing bacteria in colorectal cancer. **a**, Signalling pathway associated with *pks*⁺ bacteria-induced cellular senescence. **b**, *pks*⁺ bacteria contribute to the emergence of a senescent microenvironment, which fuels tumour growth via the secretion of growth factors. SASP, senescence-associated secretory phenotype.

colon cancer progression, a marker of poor prognosis and a target for treatment^{35,36}, was significantly increased in CRC colonized by *pks+* *E. coli* compared to those harbouring *pks-* *E. coli*. Furthermore, inhibition of HGF significantly decreased cellular proliferation *in vitro* as well as in a mouse xenograft model, demonstrating that HGF was the main mechanistic link between *pks+* *E. coli*-induced senescence and cell proliferation.

We also deciphered the mechanism underlying *pks+* *E. coli*-induced senescence and found that it involves c-Myc, miR-20a-5p and SENP1. As shown in the model presented in figure 6a, colibactin-producing *pks+* *E. coli* increased miR-20a-5p expression in infected cells *via* c-Myc transcription factor. We confirmed that this miRNA bound the 3'UTR of *SENP1*-encoding mRNA, resulting in its translational silencing. All together, our results indicate that colibactin-producing *E. coli* regulate SENP1 expression by promoting miR-20a-5p accumulation. Interestingly, we observed a similar deregulation of SENP1 and miR-20a-5p in human colon tumours colonized by *pks+* *E. coli*.

SENP1 is a protein involved in the regulation of SUMO conjugation^{27,37}. Accordingly, SENP1 down-regulation induced by colibactin-producing *E. coli* or transfection of miR-20a-5p triggered a modification of SUMO-conjugated proteins. Interestingly, SENP1 over-expression inhibited cellular senescence induced by colibactin. These results highlight the role of SENP1 and protein SUMOylation in colibactin-induced senescence. To date, few bacteria (*Yersinia pseudotuberculosis*, *Xanthomonas campestris* and *Listeria monocytogenes*) have been reported to impair SUMOylation of host cell proteins³⁸⁻⁴⁰. SUMOylation of p53 is known as a key step in the regulating framework that triggers cellular senescence^{26,27}. The *pks+* *E. coli*-infected cells showed increased amounts of SUMO-conjugated p53, which is consistent with the predominant enzymatic activity of SENP1³⁷.

In conclusion, our data support the hypothesis that colibactin-producing bacteria play a predominant role in CRC. To our knowledge, colibactin-producing *E. coli* is the first infectious agent shown to induce senescence, by subverting the SUMOylation process *via* SENP1 down-

regulation under the control of a miRNA. Our findings reveal a new paradigm for colorectal carcinogenesis in which colibactin-induced senescence plays a critical role. We propose that colibactin-producing *E. coli*, or any other colibactin-producing bacteria, can modulate the tumour microenvironment and favour the emergence of senescent cells, which promote tumour growth by the secretion of growth factors, notably HGF (figure 6b). Targeting colibactin production may be therefore an efficient strategy to restrain the production of pro-tumourigenic factors from the tumour microenvironment.

References

1. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.* **8**, 286–298 (2007).
2. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).
3. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
4. de Martel, C. *et al.* Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol.* **13**, 607–615 (2012).
5. Tjalsma, H., Boleij, A., Marchesi, J. R. & Dutilh, B. E. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575–582 (2012).
6. Ferlay, J. *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int. J. Cancer* **127**, 2893–2917 (2010).
7. Dove, W. F. *et al.* Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status. *Cancer Res.* **57**, 812–814 (1997).
8. Uronis, J. M. *et al.* Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE* **4**, e6026 (2009).
9. Lax, A. J. Opinion: Bacterial toxins and cancer—a case to answer? *Nat. Rev. Microbiol.* **3**, 343–349 (2005).
10. Goodwin, A. C. *et al.* Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15354–15359 (2011).
11. Wang, X. & Huycke, M. M. Extracellular superoxide production by *Enterococcus faecalis* promotes chromosomal instability in mammalian cells. *Gastroenterology* **132**, 551–561 (2007).
12. Wu, S., Rhee, K.-J., Zhang, M., Franco, A. & Sears, C. L. *Bacteroides fragilis* toxin stimulates intestinal epithelial cell shedding and gamma-secretase-dependent E-cadherin cleavage. *J. Cell. Sci.* **120**, 1944–1952 (2007).
13. Wu, S. *et al.* A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* **15**, 1016–1022 (2009).
14. Abdulmir, A. S., Hafidh, R. R. & Bakar, F. A. Molecular detection, quantification, and isolation of *Streptococcus gallolyticus* bacteria colonizing colorectal tumors: inflammation-driven potential of carcinogenesis via IL-1, COX-2, and IL-8. *Mol. Cancer* **9**, 249 (2010).
15. Swidsinski, A. *et al.* Association between intraepithelial *Escherichia coli* and colorectal cancer. *Gastroenterology* **115**, 281–286 (1998).
16. Martin, H. M. *et al.* Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* **127**, 80–93 (2004).
17. Maddocks, O. D. K., Short, A. J., Sonnenberg, M. S., Bader, S. & Harrison, D. J. Attaching and effacing *Escherichia coli* downregulate DNA mismatch repair protein in vitro and are associated with colorectal adenocarcinomas in humans. *PLoS ONE* **4**, e5517 (2009).
18. Nougayrède, J.-P., Taieb, F., De Rycke, J. & Oswald, E. Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol.* **13**, 103–110 (2005).
19. Nougayrède, J.-P. *et al.* *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848–851 (2006).
20. Cuevas-Ramos, G. *et al.* *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537–11542 (2010).

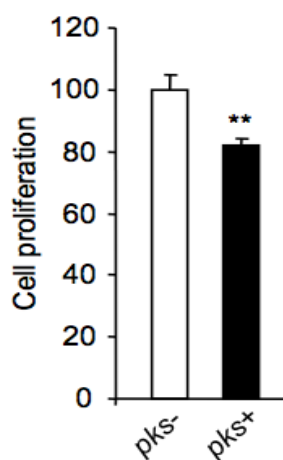
21. Dubois, D. *et al.* High prevalence of cyclomodulin- and genotoxin-producing *E. coli* associated with colon tumors. *In submission* (2012).
22. Arthur, J. C. *et al.* Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science* (2012).doi:10.1126/science.1224820
23. Rodier, F. & Campisi, J. Four faces of cellular senescence. *J. Cell Biol.* **192**, 547–556 (2011).
24. Ben-Porath, I. & Weinberg, R. A. The signals and pathways activating cellular senescence. *Int. J. Biochem. Cell Biol.* **37**, 961–976 (2005).
25. Kuilman, T. & Peeper, D. S. Senescence-messaging secretome: SMS-ing cellular stress. *Nat. Rev. Cancer* **9**, 81–94 (2009).
26. Bischof, O. & Dejean, A. SUMO is growing senescent. *Cell Cycle* **6**, 677–681 (2007).
27. Yates, K. E., Korb, G. A., Shtutman, M., Roninson, I. B. & DiMaio, D. Repression of the SUMO-specific protease Senp1 induces p53-dependent premature senescence in normal human fibroblasts. *Aging Cell* **7**, 609–621 (2008).
28. Witty, J., Aguilar-Martinez, E. & Sharrocks, A. D. SENP1 participates in the dynamic regulation of Elk-1 SUMOylation. *Biochem. J.* **428**, 247–254 (2010).
29. Bonifacio, L. N. & Jarstfer, M. B. MiRNA profile associated with replicative senescence, extended cell culture, and ectopic telomerase expression in human foreskin fibroblasts. *PLoS ONE* **5**, (2010).
30. Nowrouzian, F. L. & Oswald, E. Escherichia coli strains with the capacity for long-term persistence in the bowel microbiota carry the potentially genotoxic pks island. *Microb. Pathog.* **53**, 180–182 (2012).
31. Escobar-Páramo, P. *et al.* Large-scale population structure of human commensal Escherichia coli isolates. *Appl. Environ. Microbiol.* **70**, 5698–5700 (2004).
32. La Porta, C. A. M., Zapperi, S. & Sethna, J. P. Senescent cells in growing tumors: population dynamics and cancer stem cells. *PLoS Comput. Biol.* **8**, e1002316 (2012).
33. Vermeulen, L., de Sousa e Melo, F., Richel, D. J. & Medema, J. P. The developing cancer stem-cell model: clinical challenges and opportunities. *Lancet Oncol.* **13**, e83–89 (2012).
34. Bonnet, D. & Dick, J. E. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nat. Med.* **3**, 730–737 (1997).
35. Stein, U. *et al.* MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat. Med.* **15**, 59–67 (2009).
36. Matsui, S. *et al.* Clinical significance of aggressive hepatectomy for colorectal liver metastasis, evaluated from the HGF/c-Met pathway. *Int. J. Oncol.* **37**, 289–297 (2010).
37. Mikolajczyk, J. *et al.* Small ubiquitin-related modifier (SUMO)-specific proteases: profiling the specificities and activities of human SENPs. *J. Biol. Chem.* **282**, 26217–26224 (2007).
38. Orth, K. *et al.* Disruption of signaling by Yersinia effector YopJ, a ubiquitin-like protein protease. *Science* **290**, 1594–1597 (2000).
39. Ribet, D. *et al.* Listeria monocytogenes impairs SUMOylation for efficient infection. *Nature* **464**, 1192–1195 (2010).
40. Hotson, A., Chosed, R., Shu, H., Orth, K. & Mudgett, M. B. Xanthomonas type III effector XopD targets SUMO-conjugated proteins in planta. *Mol. Microbiol.* **50**, 377–389 (2003).

Acknowledgements

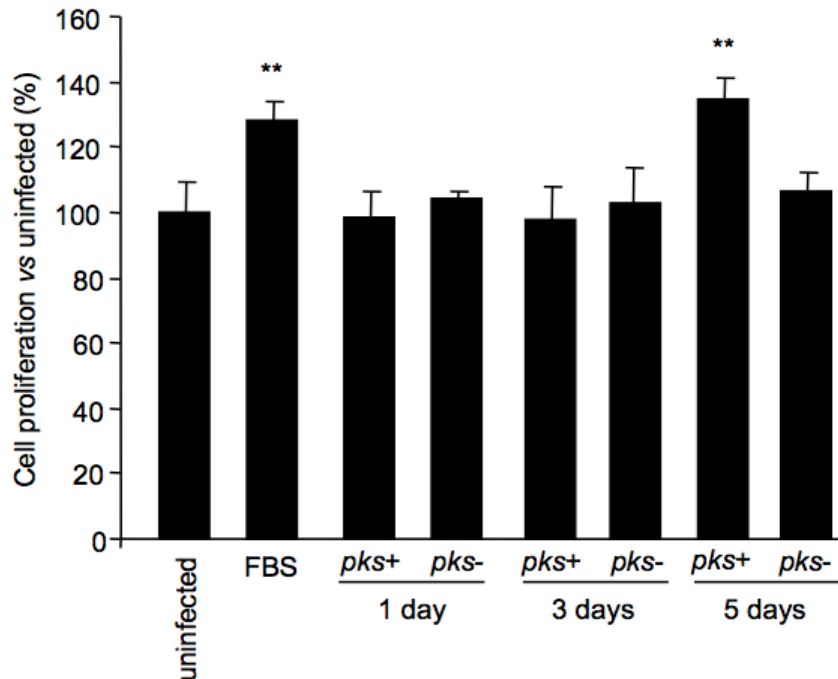
We thank Professor Andrew D Sharrocks (Faculty of Life Sciences, University of Manchester, U.K.) for kindly providing us vector encoding WT or mutated SENP1. We thank Marlène Jan for her excellent technical assistance. This work was supported by the Ministère de la Recherche et de la Technologie, the Institut national de la santé et de la recherche médicale (UMR Inserm U1071) the Institut National de la Recherche Agronomique (USC-2018), the Ligue contre le cancer and the Centre Hospitalier Régional Universitaire de Clermont-Ferrand, France.

Supplementary data

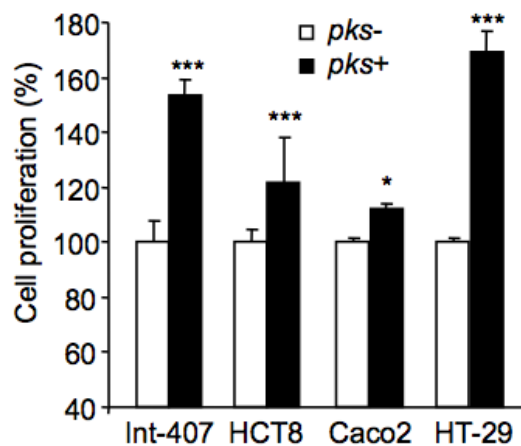
Supplementary figure 1: *pks+* *E. coli* decrease proliferation of HCT116 infected cells. HCT116 cells were infected with *pks+* *E. coli*- or *pks-* *E. coli* and cellular growth was assessed 24 hours post-infection by a MTT-based cell proliferation assay (n = 6/group, ** $P < 0.01$ vs uninfected). Data are presented as means \pm S.E.M.



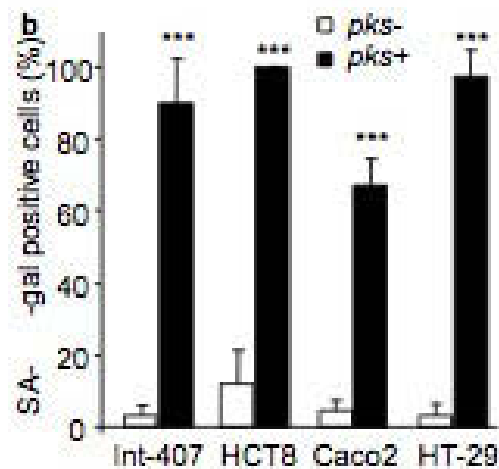
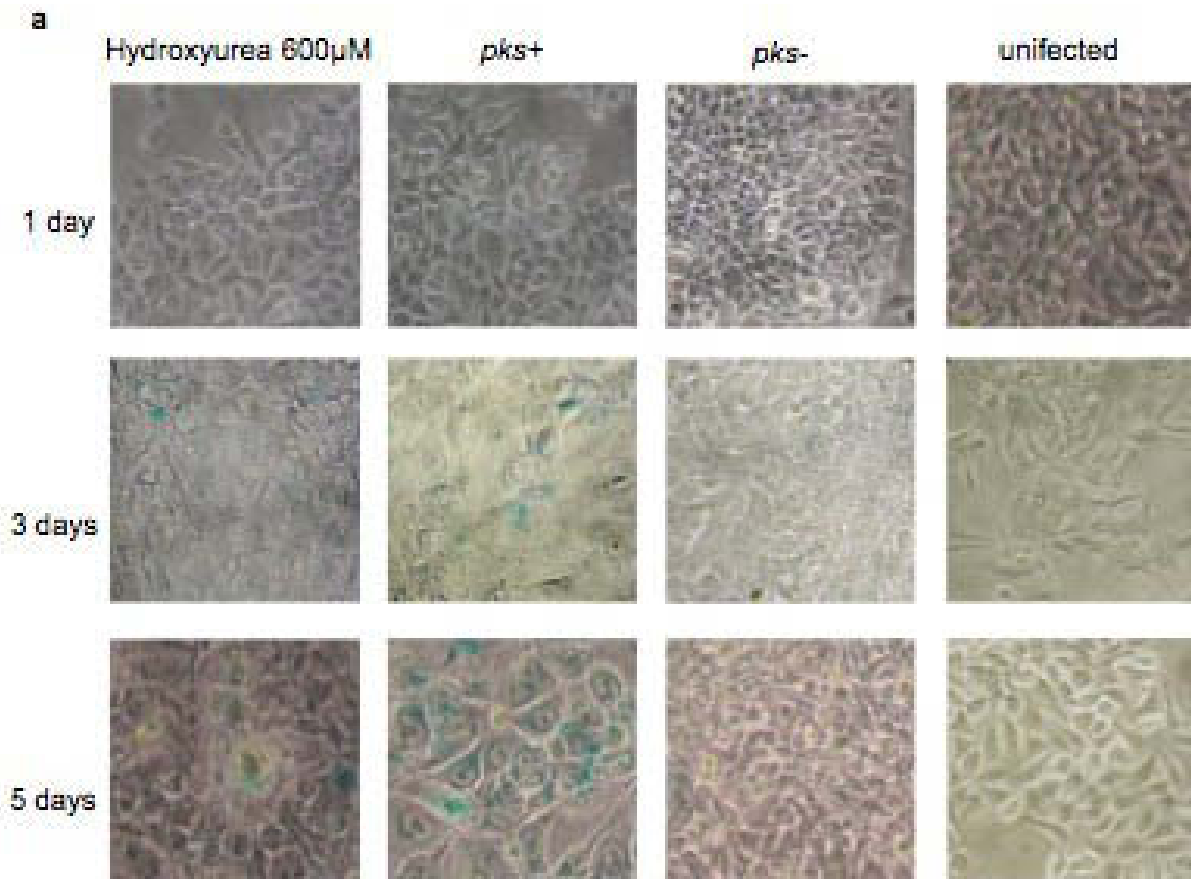
Supplementary figure 2: Conditioned media derived from *pks+* *E. coli*-infected HCT116 cells collected at day 5 post-infection stimulate cell proliferation. Conditioned media derived from *pks+* or *pks-* *E. coli*-infected HCT116 cells were collected at day 1, 3 and 5 post-infection and were used to stimulate the growth of non-infected HCT116 cells. Cellular growth was assessed by a MTT-based cell proliferation assay (n = 6/group, ***P*<0.01 vs uninfected). Data are presented as means ± S.E.M.



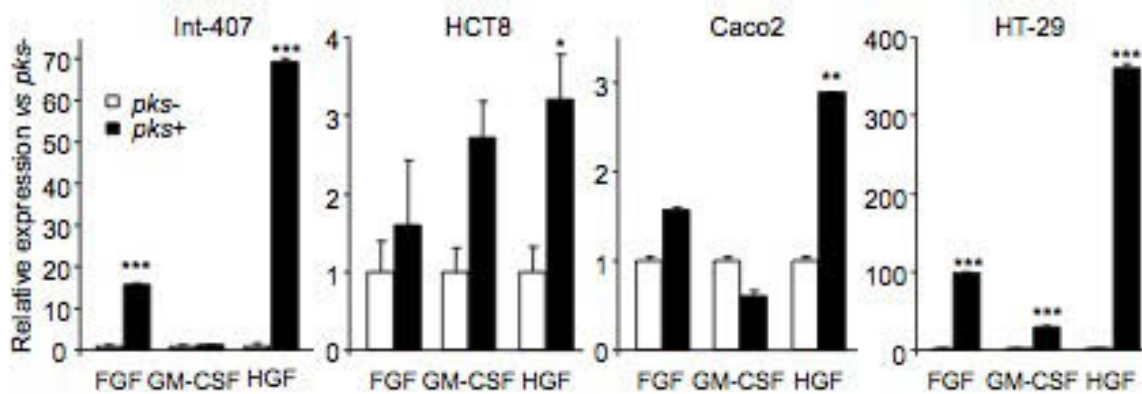
Supplementary figure 3: Conditioned media derived from *pks+* *E. coli*-infected cells stimulate proliferation of uninfected cells. Conditioned media derived from *pks+* or *pks-* *E. coli*-infected intestinal epithelial cell lines were collected at day 5 post-infection and were used to stimulate the growth of non-infected cells. Cellular growth was assessed by a MTT-based cell proliferation assay (n = 8/group; **P*<0.05; ****P*<0.001 vs *pks-*). Data are presented as means ± S.E.M.



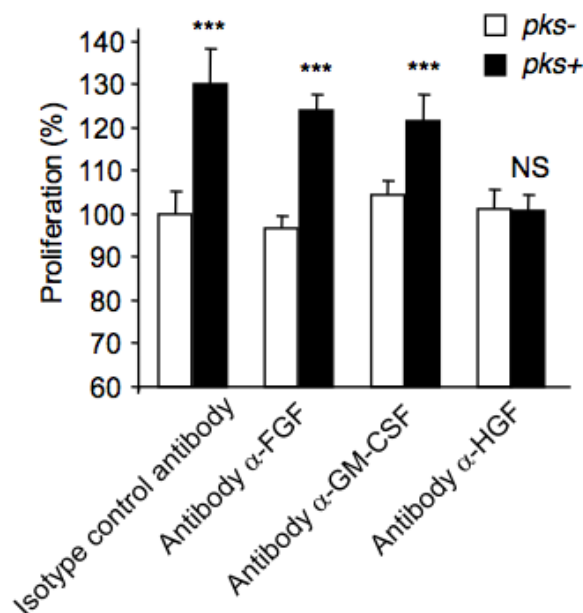
Supplementary figure 4: *pks+* *E. coli* induce senescence of intestinal epithelial cells. a, HCT116 cells were infected with *pks+* or *pks-* *E. coli* and SA-β-gal positive cells were assessed 1, 3 or 5 days post-infection. Hydroxyurea was used as a positive control. b, Various human intestinal epithelial cells were infected with *pks+* or *pks-* *E. coli* and SA-β-gal positive cells were counted 5 days post-infection. Data are presented as means ± S.E.M. (n = 8/group; ****P*<0.001 vs *pks-*).



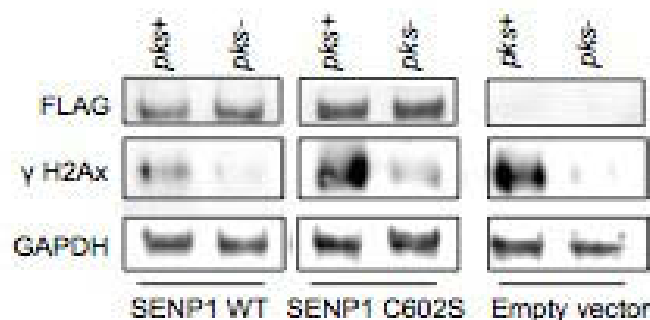
Supplementary figure 5: Senescent cells produce growth factors. Various intestinal epithelial cells were infected with *pks+* or *pks-* *E. coli* and growth factors mRNA expression levels were assessed by qRT-PCR 5-days post-infection (n = 8/group; * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$ vs *pks-*).



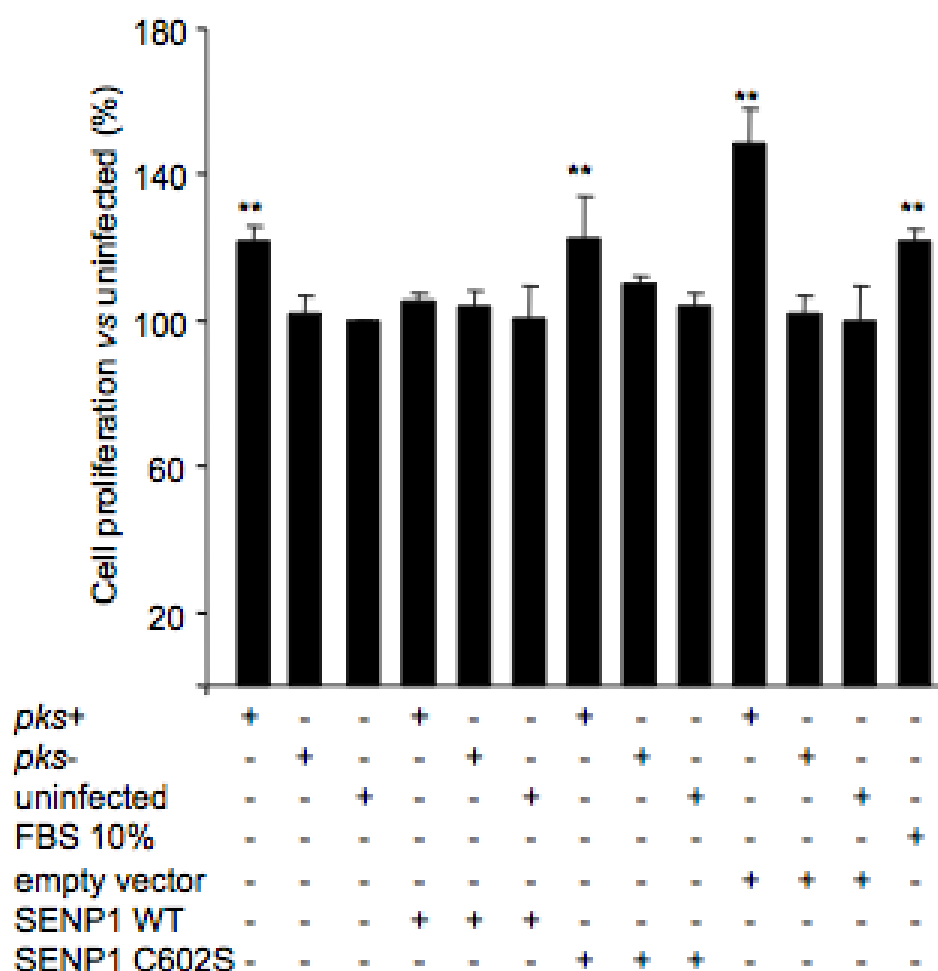
Supplementary figure 6: HGF mediated proliferation of HCT116 cells. Conditioned media derived from *pks+* *E. coli*- or *pks-* *E. coli*-infected HCT116 cells collected at day 5 post-infection were incubated 1 hour at room temperature with growth factor neutralizing antibodies or control IgG (300 $\mu\text{g/mL}$). Uninfected HCT116 cells were then cultured in the presence of this conditioned media and cellular growth was assessed by a MTT-based cell proliferation assay (n = 8/group; *** $P < 0.001$; NS, not significant vs isotype control antibody). Data are presented as means \pm S.E.M.



Supplementary Figure 7: *pks+* *E. coli* induced DNA damage in cells over-expressing functional or nonfunctional SENP1 protein. HCT116 cells transiently transfected with an empty vector, or a vector encoding SENP1 (SENP1 WT) or an inactive mutant (SENP1 C602S), were infected with *pks+* or *pks-* *E. coli*. SENP1 over-expression was assessed using anti-FLAG antibody. S39-phosphorylation level of H2Ax was assessed at 24 hours post-infection.



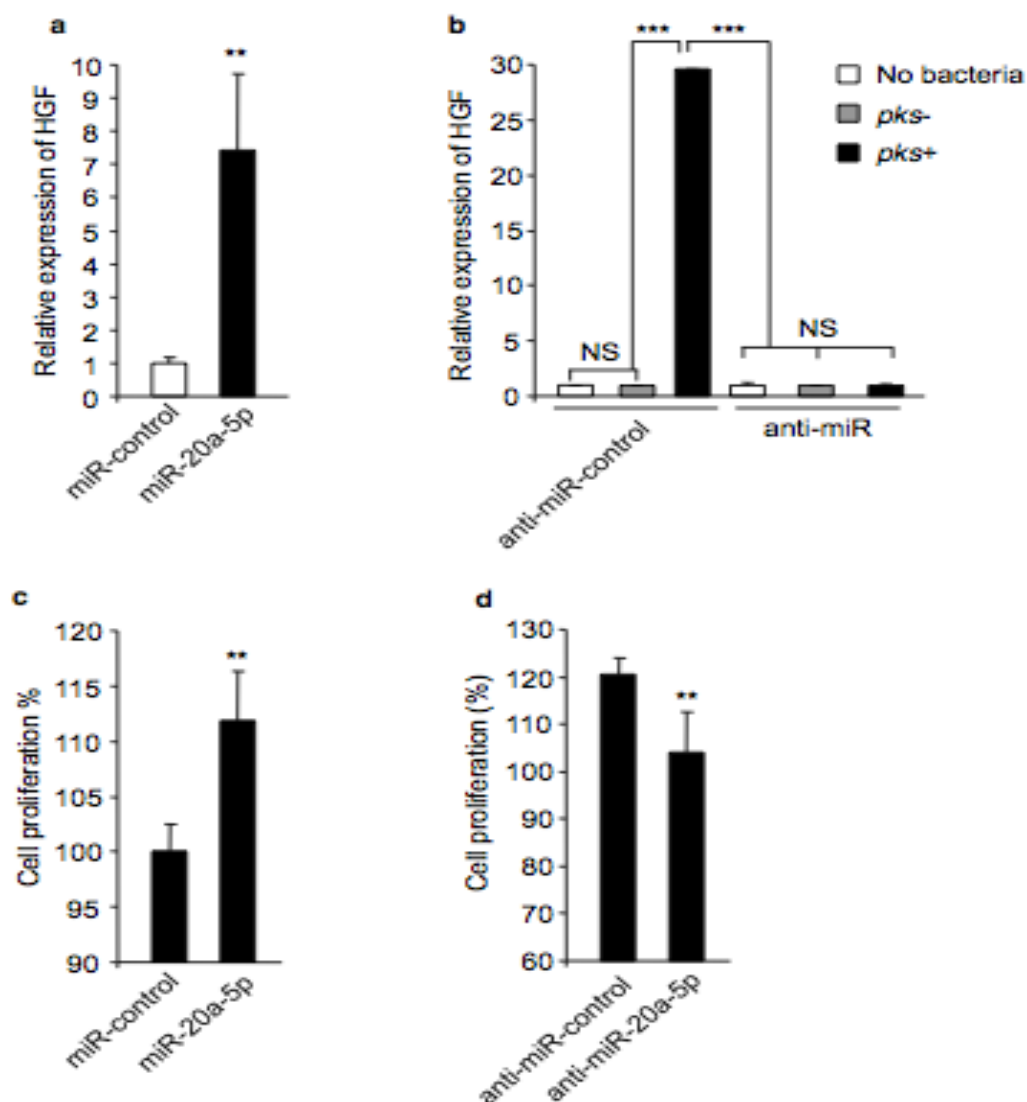
Supplementary figure 8: SENP1 over-expression abolished the pro-proliferative effect of the conditioned media derived from *pks+* *E. coli*-infected cells. HCT116 cells transiently transfected with an empty vector, or a vector encoding SENP1 (SENP1 WT) or an inactive mutant (SENP1 C602S), were infected with *pks+* or *pks-* *E. coli*. Conditioned media collected at day 5 post-infection were used to stimulate uninfected cells and cellular growth was assessed by MTT assay ($n=6$ / group, $**P<0.01$ vs uninfected). Data are presented as means \pm S.E.M.



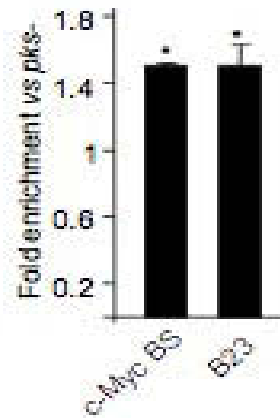
Supplementary Figure 9: Sequence alignment of *SENP1* mRNA 3'-UTR miR-20a-5p target site.



Supplementary figure 10: MiR-20a-5p controls senescence-associated secretory phenotype. **a**, HCT116 cells were transfected with miR-20a-5p or miR-control and HGF mRNA expression levels were assessed by qRT-PCR (n=8/group, $**P<0.001$). Data are presented as means \pm S.E.M. **b**, HCT116 cells transfected with anti-miR-20a-5p or an anti-miR-control were infected with *pks+* or *pks-* *E. coli* and HGF mRNA expression levels were assessed by qRT-PCR (n=8/group). Data are presented as means \pm S.E.M. **c**, HCT116 cells were cultivated in the presence of conditioned media derived from HCT116 cell transfected with mature miR-20a-5p, or miR-control and cellular growth was assessed by MTT assay (n=8/group, $**P<0.01$). Data are presented as means \pm S.E.M. **d**, Conditioned media were derived from HCT116 transfected with anti-miR-20a-5p or anti-miR-control and then infected with *pks+* *E. coli*. Uninfected HCT116 cells were cultured in the presence of this conditioned media and cellular growth was assessed by MTT assay (n=8/group, $**P<0.01$). Data are presented as means \pm S.E.M.



Supplementary Figure 11: c-Myc binds miR-20a-5p promoter. HCT116 cells were infected by *pks+* or *pks-* *E. coli* and chromatin immunoprecipitation assay was performed. Data represent a qRT-PCR analysis of c-Myc bound to miR-20a-5p promoter in infected cells at a conserved c-Myc binding site (BS) located 1480 nucleotides upstream of miR-20a-5p gene. Amplification of a validated c-Myc-binding site in intron 1 of B23 gene served as a positive control (* $P < 0.05$ vs *pks-*). Data are presented as means \pm S.E.M.



Online Methods

Cell culture and infection: HCT116, I407, HCT8, Caco2 and HT29 cells were grown following ATCC's guidelines. Cells were kept at 37°C in a 5% CO₂ atmosphere and 90% humidity. Cells were infected at a multiplicity of infection (MOI) of 500. After 3 h of infection, cells were washed with PBS, and culture medium supplemented with 200 µg/mL gentamicin was added. Culture medium was then changed every 2 days. To prepare conditioned medium, cells were infected as described above and 1-, 2- or 4-day post-infection, the serum was removed. 24 h later, medium was collected and used as conditioned medium.

Bacterial strains: The laboratory *E. coli* strain DH10B hosting the pBeloBAC 11 bearing the *pks* island (*pks*⁺ *E. coli*), or hosting an empty pBeloBAC 11 (*pks*⁻ *E. coli*) were used¹. Clinical *pks*⁺ *E. coli* strains were previously isolated from human colon tumors². Bacteria were grown at 37°C in Luria Bertani broth (LB) or on LB agar plates.

Clinical studies: All human studies were performed in accordance with an approved University of Clermont-Ferrand Review Board protocol after legal informed consent from patients. We included in our study the tumor biopsies colonized by a single *pks*⁺ or *pks*⁻ *E. coli* clone devoided of genes encoding haemolysin as well as other known genotoxins or cyclomodulins.

Western blot: Total proteins were extracted using RIPA buffer (150 mM NaCl, 50 mM Tris-HCl [pH 8.0], 0.1% SDS, 0.1% Nonidet P-40, 2 mM Na₃VO₄, 10 mM NaF, supplemented with protease inhibitor cocktail [Roche]), resolved on polyacrylamide gels and transferred to PVDF membranes (Amersham Biosciences). Membranes were blocked and then probed overnight at 4°C with anti-SENP1 (Sigma), anti-p21 (Cell signaling), anti-Rb (Santa cruz), anti-E2F1 (Santa cruz), anti-HA (Sigma), anti-c-Myc (Biolegend), anti-SUMO1 (Sigma), anti-p53 (Sigma), anti-

pp53(S15) (Cell signaling) or anti-GAPDH antibody (Sigma). After several washes, membranes were incubated with the appropriate HRP conjugated secondary antibodies (Sigma). Immunoreactive bands were detected using the enhanced chemiluminescence detection system (Amersham Biosciences).

Transfection experiments: All the transfection were performed using Lipofectamine 2000 (Invitrogen). We used 30 nM of scramble siRNA or c-Myc siRNA (Mision^o siRNA, Sigma), 40 nM of miR-20a-5p precursor (Ambion), pre-miR negative control (Ambion), anti-miR-20a-5p (Ambion) or anti-miR-control (Ambion).

Quantitative real-time PCR (qRT-PCR): Total RNAs isolated using the TRIzol reagent (Invitrogen) were reversely transcribed using the NCodeTM miRNA first-strand cDNA synthesis kit (Invitrogen) to quantify mature miRNA expression, or using the first-strand cDNA synthesis kit (Euromedex) to quantify mRNA expression according to the manufacturer's instruction. qRT-PCR was performed using SYBR Green qPCR Master Mix (Roche) on a Mastercycler Realplex⁴ (Eppendorf) using specific primers (Supplementary Table 1). Fold-induction was calculated using the Ct method: $\Delta\Delta Ct = (Ct_{\text{Target gene}} - Ct_{\text{Housekeeping gene}})_{\text{group1}} - (Ct_{\text{Target gene}} - Ct_{\text{Housekeeping gene}})_{\text{group2}}$, and the final data were derived from $2^{-\Delta\Delta Ct}$.

MiRNA target prediction: To determine the potential target genes of detected miRNAs, three different miRNA target prediction algorithms were used: PicTar (<http://pictar.mdc-berlin.de>)³, miRDB (<http://mirdb.org/miRDB/>)⁴, TargetScan (<http://www.targetscan.org/>)⁵ and MicroCosm (<http://www.ebi.ac.uk/enright-srv/microcosm/htdocs/targets/v5/>)⁶. Genes that were identified by at least two algorithms were selected for further investigations.

Luciferase repression experiment: *SENP1* mRNA 3'-UTR was cloned into *SacI/HindIII* sites of the pMIR-REPORTTM vector (Ambion) using the following primers: *SENP1 SacI*: 5'-**GAG CTC** CTG GTC ATC AAA GCC TAC CC-6'; *SENP1 HindIII*: 5'-**AAG CTT** GTG GCA ATA TGG TGT CGA TTT G-3'. For luciferase assay, HCT116 cells cultured on 24-well plastic plates were cotransfected with 1 µg of the *SENP1* 3'-UTR-luciferase construct and 40 nM of miR-20a-5p precursor (Ambion) or pre-miR negative control (Ambion) using Lipofectamine 2000 (Invitrogen). Luciferase activity was measured at 1 day post-transfection using the Dual-Luciferase Reporter Assay system (Promega) and a Luminoskan Ascent luminometer (Thermo Electron Corp). Values were normalized to lysate protein concentration.

SUMOylation assay: Cells were transfected with a pcDNA-HA vector (Invitrogen) encoding SUMO1. After infection, total cell lysate was immunoprecipitated using an anti-HA antibody and Western blot was then performed.

Senescence-Associated β -Galactosidase Staining: Senescent cells were visualized using the Senescence Cells Histochemical Staining Kit (Sigma) according to manufacturer's instructions. *In vivo* β -galactosidase staining was performed as previously described⁷. Briefly, tumors were embedded into OCT (Thermo Fisher), freeze dried and immediately processed. 5 µm sections were fixed 1 min with 1% paraformaldehyde solution and colored overnight using X-gal staining solution (Sigma).

Growth factor inhibition: Conditioned media prepared as described above were supplemented with 4 nM HGF inhibitor (JNJ-38877605, Selleckchem), FGF inhibitor (PD173074, Selleckchem), vehicle or were incubated 1 hour at room temperature with 0.1 mg/ml of

neutralizing antibody (Sigma) or rat IgG isotype control (Biolegend). Conditioned media were then used to culture HCT116 cells.

Tumorigenesis assay: 2×10^6 HCT116 cells were subcutaneously injected (200 μ l) to 5 week-old nude mice (Charles River) and immediately infected with *pks+* *E. coli* or *pks-* *E. coli* (MOI = 20). Bacteria were killed 3 h post-infection with the use of 150 mg/kg of imipenem (MSD). In another set of experiments, nude mice were subcutaneously injected into the dorsal flap with a mix of 2×10^6 HCT116 cells composed of 1.6×10^6 HCT116 cells + 0.4×10^6 senescent HCT116 cells (5 days post-*pks+* *E. coli* infection) or 1.6×10^6 HCT116 cells + 0.4×10^6 non-senescent HCT116 cells (5 days post-*pks-* *E. coli* infection) as previously described ⁸. Cells were included into matrigel (BD). Tumors were measured twice a week with a caliper and their volumes were calculated by the formula: $(a \times b^2)/2$ where a and b are the larger and the smaller diameter respectively.

Immunohistochemistry: 5 μ m paraffin-embedded tissue sections were deparaffinized in xylene, incubated in 3% hydrogen peroxide in methanol for 30 minutes, rehydrated in ethanol gradient, and then treated with 10 mM sodium citrate buffer (pH 6.0) containing 0.05% Tween 20 at 120°C for 5 minutes in a microwave oven. Sections were blocked with 2% non-fat dry milk and 0.01% Tween 20 in PBS for 1 hour at 37°C, and then incubated with anti-SENP1 (Sigma) overnight at 4°C. After washes with PBS containing 0.01% Tween 20, sections were treated with appropriate biotinylated secondary antibodies for 30 minutes at 37°C, and color development was performed using the Vectastain ABC kit (Vector Laboratories). Sections were then counterstained with hematoxylin, dehydrated, and coverslipped. Images were acquired using a Zeiss Axioskop LSM510 microscope (Carl Zeiss MicroImaging).

Chromatin immunoprecipitation assay: c-Myc chromatin immunoprecipitation assay (ChIP) was performed using the Agarose ChIP kit (Pierce) following the manufacturer's instructions. The c-Myc promoter elements in the immunoprecipitates were detected by qRT-PCR using specific primers as described earlier⁹. Amplification of a validated c-Myc-binding site in intron 1 of the B23 gene served as a positive control⁹.

Statistical analysis: Values were expressed as mean \pm S.E.M. Statistical analysis was performed using Mann-Whitney by R software. $P < 0.05$ was considered statistically significant.

Bibliography

1. Dubois, D. *et al.* ClbP is a prototype of a peptidase subgroup involved in biosynthesis of nonribosomal peptides. *J. Biol. Chem.* **286**, 35562–35570 (2011).
2. Dubois, D. *et al.* High prevalence of cyclomodulin- and genotoxin-producing *E. coli* associated with colon tumors. *In preparation* (2012).
3. Krek, A. *et al.* Combinatorial microRNA target predictions. *Nat. Genet.* **37**, 495–500 (2005).
4. Wang, X. & El Naqa, I. M. Prediction of both conserved and nonconserved microRNA targets in animals. *Bioinformatics* **24**, 325–332 (2008).
5. Lewis, B. P., Burge, C. B. & Bartel, D. P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15–20 (2005).
6. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res.* **32**, D109–111 (2004).
7. Debacq-Chainiaux, F., Erusalimsky, J. D., Campisi, J. & Toussaint, O. Protocols to detect senescence-associated beta-galactosidase (SA-beta-gal) activity, a biomarker of senescent cells in culture and in vivo. *Nat Protoc* **4**, 1798–1806 (2009).
8. Krtolica, A., Parrinello, S., Lockett, S., Desprez, P. Y. & Campisi, J. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12072–12077 (2001).
9. O'Donnell, K. A., Wentzel, E. A., Zeller, K. I., Dang, C. V. & Mendell, J. T. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839–843 (2005).

2. Exploration structurale et fonctionnelle de la protéine ClbP

Publications n°2 et 3:

ClbP is the prototype of a peptidase subgroup involved in the biosynthesis of nonribosomal peptides.

Damien Dubois, Olivier Baron, **Antony Cougnoux**, Julien Delmas, Nathalie Pradel, Michèle Boury, Bernadette Bouchon, Marie-Agnès Bringer, Jean-Philippe Nougayrède, Eric Oswald and Richard Bonnet.

Journal of Biological Chemistry 2011 Oct 14;286(41):35562-70.

Analysis of structure-function relationships in the colibactin-maturing enzyme ClbP.

Antony Cougnoux, Lucie Gibold, Frédéric Robin, Damien Dubois, Nathalie Pradel, Alette Darfeuille-Michaud, Guillaume Dalmasso, Julien Delmas, and Richard Bonnet.

Journal of Molecular Biology. 2012 Oct 2. pii: S0022-2836(12)00782-6.

Le gène *clbP* est l'un des 16 gènes de l'îlot *pks* nécessaire à son activité génotoxique [26]. L'analyse *in silico* de ClbP montre qu'elle serait constituée de trois domaines :

1. une séquence signal clivable N-terminale dépendant de la voie de translocation Sec (ALA-QE 30-31),
2. un domaine principal extra-cellulaire (résidus 31-375), pouvant appartenir à la famille des enzymes S12 de la classification MEROPS,
3. trois hélices transmembranaires en C-terminal (résidus 390-412, 433-455, 465-485).

Ce type de protéines n'avait pas été identifié jusqu'alors comme intervenant dans la biosynthèse de PK ou de NRP, alors qu'il apparaît comme étant essentiel à l'effet génotoxique de la colibactine [26]. Nous avons donc décidé d'explorer plus avant la structure et la fonction de ClbP.

L'originalité de ClbP nous a amené à rechercher des orthologues. Les orthologues découverts ont été analysés sur le plan phylogénétique, topologique et génomique. L'arbre phylogénétique construit à partir des séquences peptidiques a révélé trois branches majeures. La première portait ClbP et des homologues appelés ClbP-like qui présentaient une topologie

similaire. Ces protéines étaient dans 31% des cas codées par des clusters de gènes codant des NRPS ou PKS-NRPS. Elles étaient principalement issues de bactéries appartenant au phylum bactérien des Firmicutes (*Bacillus* et *Clostridium*). La deuxième branche phylogénétique incluait les protéines FmtA-like présentes chez les bactéries du genre *Staphylococcus*. Ces protéines de fonction inconnue présentaient une topologie similaire à celle de ClbP, mais leurs gènes n'étaient jamais observés à proximité de gènes codant des mégasynthèses. La troisième branche phylogénétique s'écartait distinctement des deux autres et regroupait des membres de la famille S12 de la classification MEROPS [378]. Ces données suggéraient que ClbP et les protéines ClbP-like d'une part, et les protéines FmtA-like d'autre part constituent des sous-groupes distincts de la famille S12.

Afin d'évaluer la proximité fonctionnelle des protéines ClbP, ClbP-like et FmtA-like, nous avons analysé la capacité de ClbP-like représentatives (provenant de différents genres bactériens appartenant à des phyla différents et associées génétiquement ou non à des îlots codant des mégasynthèses) et d'une FmtA-like à xéno-complémenter l'activité cytopathique d'une souche de *E. coli pks* défective pour ClbP. Une xéno-complémentation a été observée pour toutes les protéines ClbP-like testées mais pas pour la protéine FmtA-like. Ces données soutiennent la promiscuité fonctionnelle des protéines ClbP et ClbP-like. Elles révèlent aussi leur tolérance au substrat car elles sont interchangeable alors qu'elles sont issues d'îlots génomiques codant des NRPS ou PKS-NRPS sans homologie significative.

Afin d'identifier la fonction de ClbP et de ses homologues fonctionnels, nous avons déterminé la structure cristallographique du domaine central de ClbP par diffraction aux rayons X à une résolution de 2,4 Å. La structure de ce domaine est similaire à celle des enzymes S12, notamment des β -lactamases bien connues telles que AmpC [379] et surtout la peptidase Pab87 [380] (**Figure 1, Publication 2**). La comparaison de la structure de ClbP avec ces dernières a permis d'identifier la poche active. Elle forme un sillon ouvert sur le solvant, profond et inhabituellement étendu, grâce au positionnement atypique d'une boucle (résidus 295 à 315) et de deux brins β (résidus 310 à 311 et 316 à 318 respectivement) à une extrémité de la poche catalytique. Cette caractéristique de ClbP suggère une adaptation de la poche active de ClbP à un substrat filiforme pouvant s'étendre vers l'extrémité N-terminale du brin β 11, alors que les substrats de AmpC sont accommodés vers l'extrémité C-terminale de cet élément structural.

Au fond de cette poche, nous avons observé des résidus correspondant aux motifs

critiques du site actif des enzymes de la famille S12 , à savoir les motifs ⁹⁵SxxK et ¹⁸⁶Yxx (**Figure 2, Publication 2**), qui assurent le processus catalytique [381]. Ces résidus ont été mutés et l'impact des mutations a été exploré dans une souche de *E. coli* K12 contenant un îlot *pks* défectif pour le gène *clbP*. Les résidus ⁹⁵S, ⁹⁸K et ¹⁸⁶Y des motifs décrits ci-dessus se sont révélés essentiels pour l'effet génotoxique de l'îlot, montrant ainsi que ClbP agit selon un processus catalytique basé sur le mécanisme généralement accepté pour des enzymes à sérine active S12 [382]. Par ailleurs, l'analyse du positionnement de ces résidus dans le site actif suggérait une activité peptidasique (**Figure 2, Publication 2**), qui a été confirmée à l'aide de substrats chromogéniques. Par contre, ClbP ne présentait pas d'activité β -lactamase détectable et était incapable de conférer une résistance aux β -lactamines antibiotiques, contrairement à la β -lactamase AmpC.

Les acides aminés situés à proximité des résidus catalytiques sont connus pour jouer un rôle dans la fixation du substrat chez AmpC ou des enzymes proches. Dans ClbP, la mutation des résidus correspondant n'affectait pas l'activité biologique de ClbP. Par ailleurs, les protéines ClbP-like présentent peu de résidus à la fois conservés et spécifiques. La plupart des résidus conservés chez ClbP et les protéines ClbP-like sont observés dans le plus proche paralogue, la protéine FmtA-like et ne sont donc pas liés à leur spécificité fonctionnelle. Pour comprendre plus avant la spécificité de ClbP pour son substrat, nous avons donc déterminé la structure tridimensionnelle de FmtA-like, puis comparé la structure globale et les charges de surface de ClbP, FmtA-like et AmpC . Nous avons observé chez ClbP une large ouverture de la poche catalytique à l'extrémité N-terminale du brin β 11 par rapport à AmpC (**Figure 2, Publication 3**). FmtA-like présentait une conformation intermédiaire. De même, l'analyse du potentiel électrostatique de surface a montré un site actif chargé positivement chez AmpC, et au contraire, des charges négatives focalisées dans le sillon catalytique chez ClbP (**Figure 3, Publication 3**). Des résultats comparables ont été observés pour les protéines ClbP-like (surface électronégative) à partir de structures modélisées par homologie (données non publiées). Chez FmtA-like, les charges négatives du site actif étaient plus diffuses et moins marquées vers l'extrémité N-terminale du brin β 11. Le potentiel électrostatique de surface de la poche catalytique de ClbP et des protéines ClbP-like est ainsi corrélé à la capacité à restaurer la bioactivité d'un îlot *pks* déficient pour ClbP, et semble donc être un déterminant majeur de leur activité. Ce résultat était soutenu par le calcul du potentiel de surface des ligands connus ou putatifs de ces enzymes, qui était complémentaire de celui des enzymes. Le potentiel électrostatique de surface du composé PK-NRP Zwittermicine A (possédant la protéine ClbP-

like ZmaM dans son îlot de biosynthèse [383]) était positif alors que celui des substrats des β -lactamases ou des carboxypeptidases était négatif (β -lactamines et précurseurs du peptidoglycane). Par des approches de modélisation moléculaire, nous avons identifié deux résidus acides (acide aspartique) pouvant être responsables d'une partie importante des charges de surface du sillon catalytique de ClbP. Ces deux résidus (position 365 et 367), situés à distance des résidus catalytiques, vers l'extrémité N-terminale du brin β 11, ont été mutés par mutagenèse dirigée. La substitution de ces deux résidus par des acides aminés non chargés ou présentant des charges opposées abolissait l'activité de ClbP (**Figure 4, Publication 3**). Ces données confirment le rôle clé de l'extrémité N-terminale du brin β 11 et des charges électro-négatives du site actif dans la bioactivité de ClbP.

Par ailleurs, nous avons confirmé par fractionnement cellulaire que ClbP était associée à la membrane cytoplasmique bactérienne (**Figure 1, Publication 3**). En utilisant des protéines de fusion GFP ou PhoA, et des *E. coli* mutants isogéniques du système Sec, nous avons aussi montré que la séquence N-terminale était une séquence clivable adressant le domaine enzymatique au périplasme *via* le système de translocation SRP/Sec/YidC (**Figure 1, Publication 3**). Cette translocation était nécessaire à la bioactivité de l'îlot *pks* (**Figure 1, Publication 3**). Nous avons également montré que ClbP restait ancré à la membrane cytoplasmique grâce aux 3 hélices C-terminales qui étaient également indispensables à la bioactivation de la colibactine (**Figure 1, Publication 3**).

En conclusion, nous avons proposé de regrouper ClbP et les ClbP-like dans une nouvelle sous-famille de peptidases à sérine active, les peptidases associées aux NRP ; des enzymes probablement impliqués dans la maturation de NRP bioactifs chargés positivement à l'extérieur du cytoplasme bactérien. Depuis Reimer et collaborateurs [384] ont montré la bioactivation de la pré-xenocoumarine en xénocoumarine par la protéine ClbP-like XcnG de *Xenorhabdus nematophila*. La maturation de la colibactine par ClbP à distance de l'ADN bactérien, au niveau du périplasme, peut contribuer à protéger le matériel génétique de l'action génotoxique de la colibactine (**Figure 5, Publication 3**), et faisait de ClbP une cible d'intérêt pour des molécules visant à contrôler la bioactivation de la colibactine.

ClbP IS THE PROTOTYPE OF A PEPTIDASE SUBGROUP INVOLVED IN THE BIOSYNTHESIS OF NONRIBOSOMAL PEPTIDES

Damien Dubois^{1,2}, Olivier Baron^{3,4,5,6,7†}, Antony Cougnoux², Julien Delmas^{1,2}, Nathalie Pradel², Michèle Boury^{3,4,5,6}, Bernadette Bouchon⁸, Marie-Agnès Bringer², Jean-Philippe Nougayrède^{3,4,5,6}, Eric Oswald^{3,4,5,6,7} and Richard Bonnet^{1,2}

From ¹CHU Clermont-Ferrand, Laboratoire de bactériologie clinique, ²Clermont Université, JE 2526, USC-INRA 2018, ³Inserm U 990, Clermont-Ferrand, France, and ⁴Inserm U1043, ⁵CNRS U5282, ⁶INRA USC, ⁷Université de Toulouse, UPS, CPTP, ⁸CHU Toulouse, Hôpital Purpan, Service de Bactériologie-Hygiène, Toulouse, France

Running head: Structure and functionality of ClbP peptidase

Corresponding authors: Richard Bonnet; rbonnet@chu-clermontferrand.fr; Phone (+33) 4 73 75 49 20; Fax (+33) 4 73 75 49 22 or Eric Oswald; e.oswald@envt.fr; Phone (+33) 5 61 19 39 91; Fax (+33) 5 61 19 39 75

† in memoriam

The *pks* genomic island of *Escherichia coli* encodes polyketide (PK) and nonribosomal peptide (NRP) synthases that allow assembly of a putative hybrid PK-NRP compound named colibactin that induces DNA double-strand breaks in eukaryotic cells. The *pks*-encoded machinery harbors an atypical essential protein, ClbP. ClbP crystal structure and mutagenesis experiments revealed a serine-active site and original structural features compatible with peptidase activity, which was detected by biochemical assays. Ten ClbP homologs were identified *in silico* in NRP genomic islands of close- and distant-related bacterial species. All tested ClbP homologs were able to complement a *clbP*-deficient *E. coli* mutant. ClbP is therefore a prototype of a new subfamily of extracytoplasmic peptidases probably involved in the maturation of NRP compounds. Such peptidases will be powerful tools for the manipulation of NRP biosynthetic pathways.

Polyketide (PK), nonribosomal peptide (NRP) and hybrid NRP-PK natural products are very interesting compounds in drug development strategies (1). Their biosynthesis occurs on multimodular enzymatic assembly lines, called megasynthases (2). The conformations that

establish biological activity are attained by tailoring and editing enzymes such as acylases, alkylases, glycosylases, and oxidoreductases during chain growth and after thioesterase-mediated release. The dedicated tailoring enzymes are encoded by genes clustered with the assembly line genes for coordinated regulation.

Recently, Nougayrède *et al.* identified in *Escherichia coli* strains and other *Enterobacteriaceae* a large (54 kb) genomic island named *pks*, which encodes a typical cluster of NRP synthases (NRPS), PK synthases (PKS) and hybrid NRPS/PKS (3,4). The *pks* island was detected in *E. coli* strains involved in extraintestinal infections such as urinary tract infection and septicemia, but also in commensal strains from healthy people (3,5,6). The genetic and functional analyses of the *pks* island indicate that it encodes for synthesis of a PK-NRP hybrid compound, named colibactin. *E. coli* strains expressing this gene cluster cause DNA double-strand breaks in human eukaryotic cells during infection (3). This DNA damage induces phosphorylation of the H2AX histones (γ H2AX) and activates the ATM-Chk2 signaling pathway that leads to transient G₂/M cell cycle arrest and cell swelling (megalocytosis). Exposed cells exhibit signs of incomplete DNA repair, which leads to chromosomal instability, an increase in gene mutation frequency and

anchorage-independent colony formation, demonstrating the mutagenic and transforming potentials of colibactin (7).

In addition to eight NRP and PK megasynthases, the *pks* island encodes nine accessory, tailoring and editing enzymes (3). Eight of these proteins, including the protein ClbP, are required to induce DNA damage in infected cells. No protein similar to ClbP has been reported so far as an accessory or tailoring enzyme. Here, we determined the crystal structure of ClbP major domain, its biochemical activity, and we characterized distant ClbP homologous proteins that form a new subgroup of serine-reactive peptidases associated with NRP assembly lines.

Experimental procedures

Strains, plasmids, and media. The strains and plasmids used in this study are listed in Suppl. Tables 1 and 2, respectively. *E. coli* strains were routinely grown at 37°C in Lysogeny broth (LB) or on LB agar plates. Ampicillin (Amp) (100 µg/ml), kanamycin (Kan) (50 µg/ml), chloramphenicol (Cm) (25 µg/ml), or tetracyclin (200 ng/ml) were added as required to the medium. Non *E. coli* strains were grown as recommended by the DSMZ or CIP collections, or by the givers.

DNA cloning for ClbP investigations. The DNA sequences of *clbP* were amplified from the genome of *E. coli* strain IHE3034 (3). The DNA fragments encoding the entire ClbP and the catalytic domain of ClbP designated ClbPpep were amplified by PCR using primer pairs clbPMK-F and clbPBN-R, and clbP-NdeI2-F and clbP-NotI2-R (Suppl. Table 3), respectively. The PCR reactions for cloning experiments were performed using the High Fidelity Platinum® *Taq* polymerase or the Expand High Fidelity PCR System according to the manufacturer's instructions (Invitrogen, France). The amplified fragments were purified, double digested by *MfeI/BamHI* for cloning at the *EcoRI/BamHI* sites of pASK-IBA33plus to obtain ClbP-RGS6His (ClbP + C-terminal in-frame RGS6His tag), or double digested by and cloned at the *NdeI/NotI* sites of

pET9am to obtain pClbPpep (ClbP catalytic domain). In other experiments, ClbP-6His was subcloned from pMB702 (3) into pBRSK vector (8) to obtain pOB901. The constructions were checked by double-stranded DNA sequencing (Genome Express, France). Plasmids were subsequently electroporated into competent *E. coli* strains: p33ClbP and pOB901 into DH10B/pBAC*pksΔclbP* strain for HeLa cellular challenges, and pClbPpep into BL21(DE3) strain for ClbPpep and derivative overexpression (original strains are listed in Suppl. Table 1).

Site-directed mutagenesis. The site-directed mutagenesis experiments were performed using Quick-Change™ Site-Directed Mutagenesis kit according to the manufacturer's instructions (Stratagene, CA). The substitutions E159K, S188A, H257A, G328S, F316R+G328S, N331A and C337A were individually introduced into ClbP-RGS6His cloned into the p33ClbP plasmid and the substitution Y186E was introduced into ClbPpep cloned into the pClbPpep plasmid. In other experiments, ClbP-6His S95A, K98T and Y186G mutants were generated by 'reverse PCR' using pOB901 as template. Reverse PCR products were purified, ligated and transformed into DH5α. Successful mutations were confirmed by double DNA sequencing and the final constructs were used to transform *E. coli* DH10B/pBAC*pksΔclbP* for p33*clbP* and pOB901 derivative vectors, or *E. coli* BL21(DE3) for pClbPpep derivative vector. The mutagenic primers are listed in Suppl. Table 4.

ClbPpep and ClbPpepY186E production and purification. pClbPpep plasmid encodes the peptidase domain of ClbP without the first 102 nucleotides and the last 390 nucleotides of the *clbP* gene. These sequences encode the signal peptide and three putative anchor hydrophobic transmembrane helices that were removed to facilitate cytoplasmic expression and protein purification. *E. coli* BL21(DE3) transformed with pClbPpep plasmid was cultured in Terrific Broth medium supplemented with 50 µg/ml kanamycin, 0.4 M D-sorbitol, and

2.5 mM β -betaine. The cell culture was grown at 37°C with shaking to an $A_{600\text{ nm}} \sim 0.7$. Overexpression of the *clbPpep*-encoding gene was induced with 0.2 mM isopropyl β -D-thiogalactopyranoside overnight at 25°C with shaking. Cells were harvested by centrifugation (10,000 g for ten minutes at 4°C), resuspended in 20 mM Tris-HCl buffer (pH 9.0) (4 mL/g of pellet). After addition of 5 μ g lysozyme (2 hours at 4°C), cells were disrupted by ultrasonic treatment (four times for 30 seconds, each time at 20 W). The extract was clarified by centrifugation at 14,000g for 60 minutes at 4°C. After addition of 2 mg DNase I (15 min at room temperature; Roche), the supernatant (10,000 g for 60 minutes at 4°C) was dialyzed overnight against 20 mM Tris-HCl buffer (pH 9.0) and filtrated (0.22 μ m filter; Millex-GS, Millipore, Carrigtwohill, Ireland). Purification was carried out by ion-exchange chromatography onto a HiTrap™ Q Sepharose™ High Performance column (10 ml; Amersham Pharmacia Biotech) equilibrated with 20 mM Tris-HCl (pH 7.0) and eluted with a linear NaCl gradient (0 to 500 mM). The ClbPpep-containing elution peak was then purified by gel filtration chromatography onto a Superose™ 12 column (3.2 by 30 cm; Amersham Pharmacia Biotech), which had been equilibrated and eluted with 100 mM NaCl 5 mM phosphate buffer (pH 8.0). ClbPpep-containing elution peak was extensively dialyzed against 50mM NaCl, 5 mM phosphate buffer (pH 8.0), concentrated by ultrafiltration to 10 mg/ml for crystallization. The enzyme was more than 95% homogeneous as determined by Coomassie blue staining after SDS-PAGE. The N-terminal sequencing of the isolated protein was achieved by sequential Edman reaction (Plate-forme de Protéomique Analytique et Fonctionnelle, INRA, Nouzilly, France). Production and purification of ClbPpepY186E were performed in the same way using the plasmid pClbPpepY186E.

Crystallization and structure determination. ClbP crystals were grown in hanging drops over a well solution of 0.8 M

monosodium dipotassium phosphate buffer (pH 7.0). Crystals reached a size of 40 μ m X 40 μ m X 70 μ m within four weeks. Before data collection, crystals were immersed for about 30 s in a cryoprotectant solution (20% sucrose, 1.0 M monosodium dipotassium phosphate pH 7.0) and were flashcooled in liquid nitrogen. Data were collected using a Q315r ADSC-CCD detector on ESRF beamline 14-4 at the European Synchrotron Radiation Facility (Grenoble, France). Reflexions were indexed, integrated, and scaled using CCP4 package (9). The initial model was obtained by molecular replacement with the program PHASER (10) and the X-ray structures 1E15, 1RGY and 1FR1 as a search model. The structure was automatically and manually refined with REFMAC5 and COOT programs, respectively (11,12). Cross-validation was used throughout and 5 % of the data were used for the R_{free} calculation. The stereochemical quality of the models was monitored with the PROCHECK program (13). Ramachandran plots were calculated by RAMPAGE (14). Processing and crystallographic refinement statistics for ClbPpep crystal structure are listed in Suppl. Table 5. The coordinates and structural factors of ClbPpep have been deposited in the RCSB Protein Data Bank, entry 3O3V.

Enzyme assays. D-aminopeptidase activity was detected by monitoring the formation of *p*-nitroaniline from Gly-, D-Ala-, L-Ala-, and L-Ala-L-Ala-*p*-nitroanilide (Bachem, Germany) at 405 nm in a 100 mM Tris, 75 mM NaCl buffer (pH 7.5) at 25°C. β -lactamase activity was tested by monitoring the hydrolysis of nitrocefim (Calbiochem, Germany) at 482 nm in a 100 mM Tris, 75 mM NaCl buffer (pH 7.5) at 25°C.

Mass spectrometry. ClbPpep (20 μ M) was incubated at 20°C for 24 h in a 100 mM Tris, 75 mM NaCl buffer (pH 7.5) alone, or with 20 mM imipenem (Sigma), in a 1 mL final volume. The reaction mixtures were desalted by dialysis against a 5 mM ammonium bicarbonate buffer, before analysis by matrix-assisted laser desorption/ionisation-time of flight (MALDI-TOF)-mass spectrometry. The

sample was spotted onto the stainless steel target (Applied Biosystems) and an equal volume of matrix solution (14 mg/mL sinapinic acid in 50% acetonitrile in water containing 0.1% trifluoroacetic acid) was deposited into the sample drop. MALDI-TOF mass spectra were acquired on a Voyager DE-PRO mass spectrometer (Applied Biosystems, MA, USA) in linear positive mode, using an accelerating voltage of 25kV, a grid voltage of 93%, a pulse delay extraction of 700ns, in the 15,000–50,000 m/z mass range. Calibration was performed in a close external mode using a protein calibration mixture (LaserBio Labs, France).

Western-blot experiments. Protein samples (20–25 µg for each sample) were resolved by SDS-PAGE separation electrophoresis on 12.5% polyacrylamide gels, immobilized onto polyvinylidene difluoride (PVDF) membranes and analyzed by immunoblot using the ECL blotting system (Amersham Biosciences, UK), and monoclonal anti-RGS6His or anti-5His antisera (Qiagen, France), according to the manufacturer's instructions.

DNA cloning of ClbP homolog-encoding genes. The DNA fragments encoding ClbP homologues (the N-terminal signal sequence, the large extracytoplasmic and catalytic domain, and the three C-terminal transmembrane helices, without the putative 'half-size' ABC-type exporter domain if present) of *H. chejuensis* KCTC 2396^T (locus tag *HCH_03446*), *B. cereus* UW85 (locus tag *zmaM*), *B. weihenstephanensis* KBAB4 (locus tag *BcerKBAB4_1783*), *C. cellulolyticum* H10 (locus tag *Ccel_2389*), *B. mycoides* DSM 2048 (locus tag *bmyco0001_51990*), *B. pseudomycoides* DSM 12442 (locus tag *bpmyx0001_2150*), and *S. aureus* Mu50 (locus tag *SAV2441*) were amplified by PCR using primers listed in Suppl. Table 3, purified, double digested by *MfeI* or *EcoRI* (according to primer name) and *BamHI* or *XhoI* for cloning at the *EcoRI/BamHI* or *XhoI* sites of pASK-IBA33plus (C-terminal in-frame with RGS6His tag) resulting in plasmids listed in Suppl. Table 2. The constructions were

checked by double-stranded DNA sequencing (Genome Express, France). Plasmids were electroporated into competent DH10B/pBAC*pksΔclbP*.

Eukaryotic cell culture, bacterial infections, γH2AX staining and cell cycle analysis. Cell assays were performed as previously described, with minor modifications (3). For bacterial infections, overnight LB-Amp cultures of bacteria were diluted in interaction medium (DMEM, 5 % fetal calf serum (FCS), 25 mM HEPES, Amp). For experiments with ClbP homologs, tetracyclin was added to overnight cultures and interacting medium in order to induce their expression. ~50% confluent HeLa cell cultures (ATCC CCL2) were then infected with a multiplicity of infection (number of bacteria per eukaryotic cell) of 100. Cells were washed 4 h after inoculation and incubated in DMEM 10% FCS, 200 µg/ml gentamicin until analysis. Cell morphology was examined by Giemsa staining at 72 hours after infection. For γH2AX staining, cells were fixed at 4 or 20 hours after infection in 4% formaldehyde, and incubated with anti-phospho(Ser139)-H2AX antibodies (JBW301, Upstate) followed by FITC- or rhodamine-conjugated secondary antibodies. DNA was stained with DAPI or TO-PRO-3 (Invitrogen) and images were acquired with a Zeiss LSM 510 Meta or Olympus IX70 laser scanning confocal microscope, the confocal aperture being set to achieve a z optical thickness of ~0.5 µm. For cell cycle analysis, cellular suspensions were fixed and permeabilized with ethanol 70% at 48 hours after infection, then treated with 50 µg/ml propidium iodide and 250 µg/ml RNase. DNA-content data were acquired with a SLR II or FACScalibur flow cytometer (Becton Dickinson) and analyzed with FlowJo software (Tree Star).

Bioinformatics. The search for ClbP homologous proteins and analysis of their topology were performed using BLAST (15) and InterProScan server (<http://www.ebi.ac.uk/Tools/InterProScan/>). Multiple sequence alignment of peptidase domains and phylogenetic tree were performed using COBALT and TOPALI

v2.5 softwares (16,17). Analysis of the genetic context of ClbP- and FmtA-like-encoding genes (more or less 30 kb of both sides of the genes) was performed manually from NCBI genome database. The analysis of PK and NRP megasynthase domains and their substrate specificity was performed using the SEARCH NRPS-PKS program (<http://www.nii.res.in/searchall.html>).

RESULTS

***In silico* analysis of ClbP.** The *clbP* open reading frame was predicted to encode a 504 a.a. periplasmic inner membrane-anchored protein comprising three parts: (i) a C-terminal domain of three transmembrane helices (residues 390-412, 433-455, 465-485), (ii) an N-terminal signal sequence (ALA-QE 30-31), and (iii) a large periplasmic domain (positions 31 to 375). According to primary structure comparisons, the last domain exhibits the two conserved motifs (⁹⁵SMS⁹⁸K versus SxxK and ¹⁸⁶YAS versus YxN) that characterize the catalytic center of the MEROPS S12 enzyme family (18), and is therefore designated hereafter ClbPpep. The S12 family is a heterogeneous group of active-site serine enzymes, of which the substrates are peptides or closely related compounds such as β -lactams. The activity and biological function of a few of these enzymes have been described, such as class C β -lactamases, R61 D-Ala-D-Ala carboxypeptidase B, DmpB aminopeptidase, Pab87 peptidase, D-amino acid amidase and EstB esterase (19-24). However, the activity and the function of numerous enzymes of the MEROPS S12 family remain unassigned, notably the FmtA-like proteins (25), which have been previously reported to be related to ClbP (3).

Structure determination of the ClbP peptidase domain. In order to investigate the function and activity of ClbP, the structure of the ClbPpep domain was determined by X-ray diffraction. The structure was refined against diffraction data extending to 2.4 Å resolution. The electron

density map (Suppl. Fig. 1) showed well-defined density throughout most of the structure, with the exception of the six N-terminal residues. Diffuse density indicated disorder in the loop harboring residues 241 to 249. The stereochemical parameters of the model were satisfactory (Suppl. Table 5); a Ramachandran plot showed no residues in disallowed regions of ϕ/ψ space. The final model included three ClbPpep monomers and 277 water molecules.

The 3D structure of ClbPpep consists of two structural regions, all- α and α/β (Fig. 1A). The α/β region is formed by residues 41-112 and 206-375 and the domain folds as a seven stranded anti-parallel β -sheet (b1, b2, b9, b10, b11, b12, b13) with six α -helices and three β -strands (b3, b4 and b7) packed on both faces of the sheet. The all-helical region (residues 113-205) contains four helices and loops. The conserved motifs ⁹⁵SMSK and ¹⁸⁶YAS, which correspond to S12 enzyme active site, are located in a large groove, between the two structural domains near the N-terminus of the first helix of the all- α domain.

As expected, the most pronounced homology using the DALI program was obtained with the structure of S12 family members (Suppl. Table 6), especially *E. coli* AmpC β -lactamase and Pab87 peptidase (26). Although their overall fold was clearly similar, there are significant differences affecting the accessibility of the active site from the surrounding solvent. In comparison to AmpC β -lactamases, the upper part of the catalytic pocket was largely open, because of the absence of two helices (h9 and h10) (Fig. 1A and 1E), as observed in Pab87 peptidase and R61 carboxypeptidase (21,27). Moreover, the remaining structural element of ClbP in this area, the loop between residues 295 to 315, b9 and b10 strands (residues 310 to 311 and 316 to 318 respectively), moved away from the catalytic pocket, resulting in a widening of the catalytic groove, which was unusually larger than the known structures of S12 enzymes. In the bottom part of the catalytic pocket, ClbPpep and Pab87 peptidases harbored the additional helix h4' (residues

166 to 170) in contrast with AmpC enzymes (Fig. 1A and 1E). This helix slightly obstructed accessibility to the ClbPep catalytic pocket. However, the entrance of the binding site was more open in ClbPep than in S12 peptidases, which harbor one, two or several bulky helices in this area. Of note, between C337 and C367, ClbP harbored a disulfide bond, which links b12 strand (residues 335 to 341) with the C-terminal helix h11. To our best knowledge, this disulfide bond is not observed in other known structures of S12 enzymes. Overall, the ClbPep structure was closely similar to that of S12 peptidases. However, ClbPep had noticeable conformational differences in the active site, especially in the two extremities of the catalytic groove, which was unusually large (Fig. 1B, 1C and 1D). This atypical shape suggests an adaptation of ClbP catalytic pocket to a specific substrate.

The catalytic residues and implications for catalysis. Striking structural similarities were observed between the active site of ClbPep, class C β -lactamases and other S12 enzymes such as Pab87 peptidase (21,28,29), in particular, for the motif ⁹⁵SxxK and the Y186 residue. The positioning of this last residue and the presence of H327 make ClbPep active site very similar to that of Pab87 peptidase (Fig. 2A and 2B). In addition, the location and geometry of the residues S95, K98 and Y186 suggest a direct role in the catalytic mechanism mediated by ClbP (31,32).

To confirm the importance of these residues in ClbP, substitutions S95A, K98T and Y186G were introduced by site-directed mutagenesis residues into His-tagged ClbP. The resulting mutants were assessed for their ability to restore the cytopathic activity of the *pks*-positive *E. coli* Δ *clbP* isogenic mutant. Infected HeLa cells were analyzed for megalocytosis, G₂/M cell cycle arrest, and phosphorylation of histone H2AX (3). None of these mutants were able to restore the cytopathic activity of the *pks*-positive *E. coli* Δ *clbP* isogenic mutant in contrast with the wild-type ClbP protein (Fig. 2C and 2D).

These results confirm the identification of the active site and the importance of ClbP main domain for the functionality of *pks* island. They also suggest that these enzymes and other S12 serine active-site enzymes share a catalytic mechanism.

Residues surrounding the catalytic residues and substrate binding.

The motif YxN, which is observed in most S12 enzymes was replaced by the ¹⁸⁶YAS motif in ClbP (Fig. 2B). S188 in ClbP was hydrogen bonded to the K98 residue of motif SxxK (Fig. 2A). In class C β -lactamases, the structural analog of S188 is N152, which interacts with both the substrates and K residue of motif SxxK (19). H257 of ClbP, which points at the vicinity, of S188, may also play a role in substrate binding. Opposite the ¹⁸⁶YAS motif, the catalytic pocket of ClbP exhibits the HGG motif. The alignment of ClbP with class C β -lactamases revealed that HGG motif is a structural analog of KTG-box (Fig. 2B), which is involved in substrate binding because of residues K and T (29). In β -lactamases, K can be replaced by the closely related residues H or R, and T may be replaced by S. A positively charged side chain followed by one bearing a hydroxyl group has therefore been suggested to be universally conserved. The second residue in ClbP motif is an atypical residue G328 instead of the canonical residues S or T, which are usually directly involved in substrate recognition *via* a hydrogen bond (29). However, in the vicinity of G328 (Fig. 2A and Fig. 1E), positions 330 and 331 of ClbP harbor the unusual residues Q and N respectively, which point towards the catalytic pocket and may establish hydrogen bonds with substrates.

The residues E159 and F316 of ClbPep are structural analogs of positions involved in the binding of the N-terminal and C-terminal parts of peptides in S12 amino- and carboxypeptidases, respectively. These positions determine the carboxypeptidase or aminopeptidase activity of S12 peptidases (30). Previous studies suggest that the residues E159 and F316 of

ClbP (Fig. 2A) should favor aminopeptidase activity because the negative-charged residue E159 may accommodate the positive-charged N-terminal part of peptides, and the absence of positive-charged residue in position 316 may alter the accommodation of the negative-charged C-terminal part of peptides.

Overall, the six positions 159, 188, 257, 316, 328 and 331 can therefore participate in substrate binding in ClbP, because the site chain of these residues pointed towards the catalytic pocket (Fig. 2A). The hydrogen bond donors S188 and H257 of ClbP-His were replaced by the residue A, which is not able to accept or give hydrogen bonds. The substitutions E159K and F316R were introduced into ClbP-His to modify the charges in the critical areas of the catalytic pocket. We introduced the substitution G328S to check whether a small residue is required in the second position of the motif HGG to accommodate ClbP substrate. Substitution C337A was also introduced into ClbP-His to check the importance of the disulfide bond C337-C367 for stability or folding. No substitution altered the trans-complementing activity of these ClbP variants in the *pks*-positive *E. coli* Δ *clbP* isogenic mutant. ClbP activity may have been only weakly modified by these substitutions. Unfortunately, efficient biochemical tests are not yet available to monitor ClbP activity. However, cellular tests are a good representation of the physiological context and the results are evidence of substrate tolerance of ClbP.

ClbP is a peptidase and a NRP-recognizing protein. As ClbP presents a S12 enzyme scaffold and catalytic residue positioning close to Pab87 peptidase, peptidase activity was investigated using the reporters Gly-, D-Ala-, L-Ala-, and L-Ala-L-Ala-*p*-nitroanilides. The catalytic efficiency k_{cat}/K_m of ClbPpep against these substrates were 0.005, 0.018, 0.020, and 0.040 $\mu\text{M}^{-1} \text{min}^{-1}$ respectively. This peptidase activity is weak in comparison to that of enzymes really adapted to these

substrates such as S12 aminopeptidases (30). However, in a ClbPpep mutant harboring the substitution Y186E, there was no activity, which confirms the involvement of the ClbP binding site in the peptidase activity.

Imipenem is a NRP-type compound belonging to the chemical group of β -lactams, which efficiently bind members of the S12 enzyme family, such as carboxypeptidase and AmpC β -lactamases (33,34). ClbPpep incubated with imipenem was analyzed by mass spectrometry (Suppl. Fig. 2). Imipenem induced a mass shift consistent with the formation of imipenem acyl-enzyme complex of ClbPpep, as observed for β -lactam-binding proteins (35). Despite a high concentration of imipenem (ratio imipenem/ClbPpep >1/1000), part of ClbPpep exhibited a mass corresponding to the native protein. The results show that ClbP is able to bind NRP-type compounds, has a moderate β -lactam-recognizing activity, and is devoid of significant β -lactamase activity. The k_{cat} value ($< 0.001 \mu\text{M}^{-1} \text{s}^{-1}$) was around 1,000,000-fold lower than typical values of AmpC enzymes (36). Consequently ClbP, in contrast to AmpC, was not able to increase the resistance level of *E. coli* to β -lactam antibiotics.

Identification of ClbP homologs across prokaryotic lineages. We therefore screened public protein and genomic databases to identify proteins homologous to ClbP peptidase domain. 77 homologs were obtained including the protein ZmaM, which may be involved in the synthesis of the PK-NRP antibiotic zwittermicin A (ZmA) (37). These proteins only shared 29.3% (ZP_04059476.1) to 42.7% (ZP_04215168.1) identity and all belonged to the MEROPS S12 enzyme family. The amino acid sequences and the predicted topology of both ClbP and the homologs were compared with those of representative members of the S12 family that have been assigned to a biological function and/or a biochemical activity. The resulting phylogenetic tree and the topologies revealed that the proteins clustered in three

major branches supported by high bootstrap values ($\geq 64\%$) (Suppl. Fig. 3). One branch harbored the known representative members of the S12 family, which were sub-clustered according to their biochemical activity (β -lactamases, peptidases, esterases and amidases). The second harbored the FmtA-like proteins (n=17) of *Staphylococcus* species (25) that were phylogenetically closer to ClbP than the other members of the S12 family and that exhibited a predicted topology identical to that of ClbP in contrast with the other S12 family members. The third branch exhibited the longest phylogenetic distance, which clearly defined a subgroup containing ClbP and most close-related proteins (n=60). These proteins, designated ClbP-like, harbored a predicted topology identical to that of ClbP. Some (n=14), such as ZmaM, possessed in addition a putative 'half-size' ABC-type exporter domain in C-terminal. Except ClbP and its closest homolog (ClbP-like of *Hahella chejuensis*) produced by different γ -proteobacteria, all the other ClbP-like were identified in taxons of the Firmicute phylum (*Bacillus* and *Clostridium* geni), which is the main bacterial source of PK, NRP and hybrid PK-NRP compounds with the Actinobacteria phylum (38).

ClbP-like proteins display functional promiscuity. No clues for the presence of megasynthase catalytic unit-encoding genes were revealed in the vicinity of any *fmtA*-like and 17 *clbP*-like genes. In contrast, putative NRPS- and PKS-encoding genes were observed in the vicinity of 18 *clbP*-like genes (Suppl. Fig. 4). Moreover, putative NRPS-encoding genes were present in all of these genomic islands unlike PKS-encoding genes, suggesting that the cognate products contain peptide moiety that may be related to the peptidase activity of ClbP-like proteins. To test if these enzymes are functionally equivalent, we analyzed the ability of six representative ClbP-like proteins (originating from different phyla, different geni, and encoded by genes inserted or not in NRP biosynthesis clusters) and one FmtA-like to xeno-complement the

cytopathic activity of the *pks*-positive *E. coli* $\Delta clbP$ isogenic mutant. Megalocytosis, G₂/M-cell cycle arrest and histone H2AX phosphorylation were restored by all ClbP-like proteins, whose gene is located in a NRPS-encoding genomic island (n=4) or is not (n=2) (Fig. 3). In contrast, the FmtA-like protein did not restore cytopathic activity despite its full expression (Fig. 3). The efficient xeno-complementations thus support the functional promiscuity of ClbP-like proteins. Moreover, these results are evidence of substrate tolerance of NRP-associated peptidases, and confirm data obtained in experiments that have investigated residues of the ClbP catalytic pocket potentially involved in substrate binding.

DISCUSSION

The aim of this study was to identify the structure and function of the ClbP protein, which is essential to the genotoxic activity of the *pks* island. The crystal structure of the ClbP enzymatic domain together with mutagenesis experiments and biochemical assays revealed the same α/β fold observed in the members of the MEROPS S12 family and a serine-reactive active site conferring NRP-binding and peptidase activities necessary to colibactin bioactivity. ClbP presented an unusually wide and deep catalytic groove and an active site with original structural features. We extended investigations to ClbP-like proteins which are ClbP closest related proteins and are clustered on a distinct phylogenetic branch in the MEROPS S12 enzyme family. ClbP and some ClbP-like-encoding genes belong to NRPS-encoding gene clusters. The ClbP-like proteins displayed functional promiscuity, and shared a particular cellular localization and topology according to *in silico* analyses, namely a large N-terminal extracytoplasmic peptidase domain anchored to the inner membrane by three C-terminal transmembrane helices. Hence, ClbP and the closest related ClbP-like proteins form a functional peptidase subgroup of S12 serine enzymes, which is probably involved in

NRP or PK-NRP biosynthesis. We propose to designate this subgroup the 'NRP-associated peptidases'.

Bioinformatics and biochemical analyses of the ZmA biosynthesis enzymes performed by Kevany *et al.* suggest that the compound is initially biosynthesized as part of a large metabolite that is processed by a peptidase, resulting in the formation of ZmA (37). The only candidate enzyme encoded by the ZmA biosynthesis gene cluster to catalyze such a peptidase activity is ZmaM, a ClbP-like protein. This observation and our results suggest that 'NRP-associated peptidases' can process NRP or PK-NRP compounds. ClbP encoded by the *pks* island and the other NRP-associated peptidases may process a pro-colibactin and other NRP compounds into biologically active molecules by cleaving a nonribosomal-borne peptide bond. The ClbP-like peptidase domain of ZmaM had been suggested to cleave a 'pre-ZmA' metabolite between the D-Asn and D-Ser, releasing ZmA and a fatty acyl-D-Asn (37). Interestingly, *in silico* analysis of NRPS substrate specificity showed that all gene clusters producing NRP-associated peptidases encode NRPS able to insert the close related residues Asn and/or Gln into putative NRP compounds, suggesting that these residues may be involved in substrate recognition by NRP-associated peptidases. Of note, ClbP-like proteins were the only proteins encoded by these islands that harbored a predicted translocation signal sequence and a predicted periplasmic catalytic domain, in contrast to other enzymes involved in biosynthesis of PK, NRP, and mixed compounds which are cytoplasmic proteins. Accordingly, NRP-associated peptidases may act outside the cytoplasm.

Several NRP-associated peptidases harbor in C-term a typical 'half-size' ABC exporter domain, such as ZmaM (Suppl. Fig. 3 and 4). These ABC export systems are involved in the extrusion of noxious substances and the export of extracellular toxins (39). It has already been suggested that these ABC exporters transport NRP

compounds (39). ClbP and other NRP-associated peptidases did not harbor such an export domain, but their cognate gene clusters encode a MATE transporter (like ClbM in the *pks* island) or other efflux pumps. After the translocation of an inactive NRP compound from cytoplasm to periplasm *via* efflux pumps, the periplasmic location of NRP-associated peptidase domain may allow the maturation of the fully active toxin away from a potential cytoplasmic target of its producer cell. NRP-associated peptidases and efflux pumps can be therefore seen as a two-component resistance mechanism against the possible toxic activity generated by the cognate NRPS-encoding gene cluster.

NRP-associated peptidases genetically linked to NRPS gene clusters display a functional promiscuity with those genetically unlinked to NRPS gene clusters. These latter peptidases may also participate in the synthesis of a bioactive NRP compound, because we found one to three putative NRP biosynthetic clusters in 15 of 16 genomes possessing a NRP-associated peptidase genetically unlinked to a NRPS gene cluster. It has already been suggested that enzymes required for NRP and PK biosynthesis like phosphopantetheinyl transferases are encoded in a distinct region of the main biosynthetic gene cluster in bacterial genomes (40).

In conclusion, ClbP is a prototype of a new subfamily of extracytoplasmic serine-reactive peptidases probably acting as maturing enzymes for NRP compounds. The essential role of ClbP in the production of the biological effect make this protein a key target for the control of the bioactivity of the PK-NRP-producing *pks* gene cluster, which may affect commensalism and/or pathogenicity and constitute a predisposing factor for the development of colorectal cancer (7). The engineering of NRP compounds should take into account these new catalytic units, which could be interesting tools for modifying the structures and bioactivity of NRP and PK-NRP compounds, and be a target for identifying bioactive molecules.

REFERENCES

1. Li, J. W., and Vederas, J. C. (2009) *Science* **325**(5937), 161-165
2. Sattely, E. S., Fischbach, M. A., and Walsh, C. T. (2008) *Nat Prod Rep* **25**(4), 757-793
3. Nougayrede, J. P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., Buchrieser, C., Hacker, J., Dobrindt, U., and Oswald, E. (2006) *Science* **313**(5788), 848-851
4. Putze, J., Hennequin, C., Nougayrede, J. P., Zhang, W., Homburg, S., Karch, H., Bringer, M. A., Fayolle, C., Carniel, E., Rabsch, W., Oelschlaeger, T. A., Oswald, E., Forestier, C., Hacker, J., and Dobrindt, U. (2009) *Infect Immun* **77**(11), 4696-4703
5. Dubois, D., Delmas, J., Cady, A., Robin, F., Sivignon, A., Oswald, E., and Bonnet, R. (2010) *J Clin Microbiol*
6. Johnson, J. R., Johnston, B., Kuskowski, M. A., Nougayrede, J. P., and Oswald, E. (2008) *J Clin Microbiol* **46**(12), 3906-3911
7. Cuevas-Ramos, G., Petit, C. R., Marcq, I., Boury, M., Oswald, E., and Nougayrede, J. P. (2010) *Proc Natl Acad Sci U S A* **107**(25), 11537-11542
8. Schmitt, C. K., Darnell, S. C., Tesh, V. L., Stocker, B. A., and O'Brien, A. D. (1994) *J Bacteriol* **176**(2), 368-377
9. Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K., and Noble, M. (2002) *Acta Crystallogr D Biol Crystallogr* **58**(Pt 11), 1955-1957
10. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) *J Appl Crystallogr* **40**(Pt 4), 658-674
11. Steiner, R. A., Lebedev, A. A., and Murshudov, G. N. (2003) *Acta Crystallogr D Biol Crystallogr* **59**(Pt 12), 2114-2124
12. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1), 2126-2132
13. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J App Cryst* **26**, 283-291
14. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) *Proteins* **50**(3), 437-450
15. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res* **25**(17), 3389-3402
16. Papadopoulos, J. S., and Agarwala, R. (2007) *Bioinformatics* **23**(9), 1073-1079
17. Milne, I., Lindner, D., Bayer, M., Husmeier, D., McGuire, G., Marshall, D. F., and Wright, F. (2009) *Bioinformatics* **25**(1), 126-127
18. Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J., and Barrett, A. J. (2008) *Nucleic Acids Res* **36**(Database issue), D320-325
19. Beadle, B. M., Trehan, I., Focia, P. J., and Shoichet, B. K. (2002) *Structure* **10**(3), 413-424
20. Pratt, R. F. (2008) *Cell Mol Life Sci* **65**(14), 2138-2155
21. Delfosse, V., Girard, E., Birck, C., Delmarcelle, M., Delarue, M., Poch, O., Schultz, P., and Mayer, C. (2009) *PLoS One* **4**(3), e4712
22. Okazaki, S., Suzuki, A., Komeda, H., Asano, Y., and Yamane, T. (2008) *J Synchrotron Radiat* **15**(Pt 3), 250-253
23. Wagner, U. G., Petersen, E. I., Schwab, H., and Kratky, C. (2002) *Protein Sci* **11**(3), 467-478
24. Bompard-Gilles, C., Remaut, H., Villeret, V., Prange, T., Fanuel, L., Delmarcelle, M., Joris, B., Frere, J., and Van Beeumen, J. (2000) *Structure* **8**(9), 971-980
25. Komatsuzawa, H., Choi, G. H., Fujiwara, T., Huang, Y., Ohta, K., Sugai, M., and Suginaka, H. (2000) *FEMS Microbiol Lett* **188**(1), 35-39
26. Holm, L., and Sander, C. (1996) *Science* **273**(5275), 595-603

27. Meroueh, S. O., Minasov, G., Lee, W., Shoichet, B. K., and Mobashery, S. (2003) *J Am Chem Soc* **125**(32), 9612-9618
28. Dodson, G., and Wlodawer, A. (1998) *Trends Biochem Sci* **23**(9), 347-352
29. Matagne, A., Lamotte-Brasseur, J., and Frere, J. M. (1998) *Biochem J* **330** (Pt 2), 581-598
30. Delmarcelle, M., Boursoit, M. C., Filee, P., Baurin, S. L., Frere, J. M., and Joris, B. (2005) *Protein Sci* **14**(9), 2296-2303
31. Chen, Y., Minasov, G., Roth, T. A., Prati, F., and Shoichet, B. K. (2006) *J Am Chem Soc* **128**(9), 2970-2976
32. Galleni, M., Lamotte-Brasseur, J., Raquet, X., Dubus, A., Monnaie, D., Knox, J. R., and Frere, J. M. (1995) *Biochem Pharmacol* **49**(9), 1171-1178
33. Matagne, A., Ghuysen, M. F., and Frere, J. M. (1993) *Biochem J* **295** (Pt 3), 705-711
34. Beadle, B. M., and Shoichet, B. K. (2002) *Antimicrob Agents Chemother* **46**(12), 3978-3980
35. Drawz, S. M., Babic, M., Bethel, C. R., Taracila, M., Distler, A. M., Ori, C., Caselli, E., Prati, F., and Bonomo, R. A. (2010) *Biochemistry* **49**(2), 329-340
36. Bauvois, C., Ibuka, A. S., Celso, A., Alba, J., Ishii, Y., Frere, J. M., and Galleni, M. (2005) *Antimicrob Agents Chemother* **49**(10), 4240-4246
37. Kevany, B. M., Rasko, D. A., and Thomas, M. G. (2009) *Appl Environ Microbiol* **75**(4), 1144-1155
38. Donadio, S., Monciardini, P., and Sosio, M. (2007) *Nat Prod Rep* **24**(5), 1073-1109
39. Saurin, W., Hofnung, M., and Dassa, E. (1999) *J Mol Evol* **48**(1), 22-41
40. Lu, Y. W., San Roman, A. K., and Gehring, A. M. (2008) *J Bacteriol* **190**(20), 6903-6908

FOOTNOTES

This paper is dedicated to the memory of our recently deceased friend and colleague Dr. Olivier Baron. We thank R. Perroux, M. Jan and C. Habouzit for technical assistance, Alexei Sorokin (INRA, Jouy en Josas, France), Michael G. Thomas (University of Wisconsin, Madison, USA), and Jihyun F. Kim (Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea) for providing *B. weihenstephanensis* KBAB4 strain, *B. cereus* UW85 strain and fosmid clone containing *hch_03446* locus tag of *H. chejuensis* KCTC 2396^T, respectively. We also thank C. Blavignac and C. Szczepaniak (CICS, Université d'Auvergne, France) for technical assistance in flux cytometry. This work was supported by le Ministère Français de l'Education Nationale, de la Recherche et de la Technologie, grant number JE2526 and l'Institut National de la Recherche Agronomique, grant number USC-2018. Molecular graphics images were produced using the UCSF Chimera package (NIH P41 RR-01081).

FIGURE LEGENDS

Figure 1. Overall structure of ClbPpep, AmpC β -lactamase (1KVL) and Pab87 peptidase (2QMI). (A) Superimposition of ClbPpep (green), AmpC (grey) (19) and Pab87 peptidase (yellow) (21) presented as ribbons. The loops including helices h9 and h10 are circled in red. Helices h4' are circled in blue. The active serine is indicated by a red star. ClbPpep (B), AmpC (C) and Pab87 peptidase (D) surfaces. The catalytic grooves are circled in pink and the active serine surfaces are shown in red. (E) Sequence alignments with secondary structures of ClbPpep with Pab87 peptidase and *E. coli* AmpC β -lactamase. The alignment was performed based on the superposed structures. The motifs SxxK and YxN, which defined the MEROPS S12 peptidase family, are indicated by a diamond symbol.

Figure 2. Key residues in ClbP active site.

(A) Residues of ClbP located in close vicinity of the binding site. Hydrogen bonds are indicated by pink dash lines and the water molecules by the letter W. Carbon atoms are in green, oxygen atoms in red and nitrogen atoms in blue. (B) Overlay of penicillin-recognizing motifs of ClbP, Pab87 peptidase and AmpC β -lactamase. Carbon atoms of ClbP, Pab87 peptidase and AmpC β -lactamase are in green, yellow and grey respectively. The motifs SxxK, Yx(N/S/T) and (K/H)(S/T/G)G are numbered in red, blue and green respectively and residue numbering is according to ClbP sequence. (C, D) Phenotypic analysis of *pks* island activity in *E. coli* DH10B/pBAC*pks* deleted or not for the *clbP* gene and trans-complemented with ClbP wild-type (WT) protein or the S95A, K98T or Y186G active site ClbP mutants. HeLa cells were infected 4 hours with 100 *E. coli* per cell, and then the cells were washed and incubated with gentamicin for 4-72 hours. Panel C: anti-5His western-blot analysis of the expression of His-tagged WT and active site ClbP mutants in *E. coli* DH10B/pBAC*pks* Δ *clbP* after infection. Panels D: Host histone H2AX Ser139 phosphorylation (γ H2AX) indicative of DNA double strand breaks was assayed by confocal immunofluorescence 4 hours after infection. DNA and γ H2AX are pseudocolored in blue and red respectively (Bars = 20 μ m). Cell swelling (megalocytosis) was observed following Giemsa staining 72 hours after infection (Bars = 50 μ m). G2-cell cycle arrest following DNA damage was assessed by flow cytometric DNA-content analysis, 48 hours after infection.

Figure 3. Functional promiscuity displayed by ClbP-like proteins.

(A) Phenotypic analysis of *pks* island activity in *E. coli* strain DH10B/pBAC*pks* Δ *clbP* complemented with empty vector, ClbP wild-type (ClbP WT), four ClbP-like proteins genetically linked to NRPS gene clusters from *H. chejuensis* (Hche), *B. weihenstephanensis* (Bwei), *C. cellulolyticum* (Ccel), and *B. cereus* (ZmaM), two ClbP-like proteins genetically unlinked to NRPS gene clusters from *B. mycoides* (Bmyc) and *B. pseudomycooides* (Bpse), or one FmtA-like protein from *S. aureus*. ~50% confluent HeLa cells were infected 4 hours with 100 *E. coli* per cell. Left panels: The cell morphology was observed following Giemsa staining at 72 hours after infection (Bars = 50 μ m). Middle panels: Histone H2AX Ser139 phosphorylation (γ H2AX) indicative of DNA double strand breaks was assayed by confocal immunofluorescence at 20 hours after infection. DNA and γ H2AX are pseudocolored in blue and green, respectively (Bars = 50 μ m). Right panels: G2-cell cycle arrest following DNA damage was assayed by flow cytometry 48 hours after infection. (B) Anti-RGS6His western-blot analysis of the expression of RGS6His-tagged ClbP wild-type protein, ClbP-like proteins, and FmtA-like protein in *E. coli* DH10B/pBAC*pks* Δ *clbP*.

Figure 1.

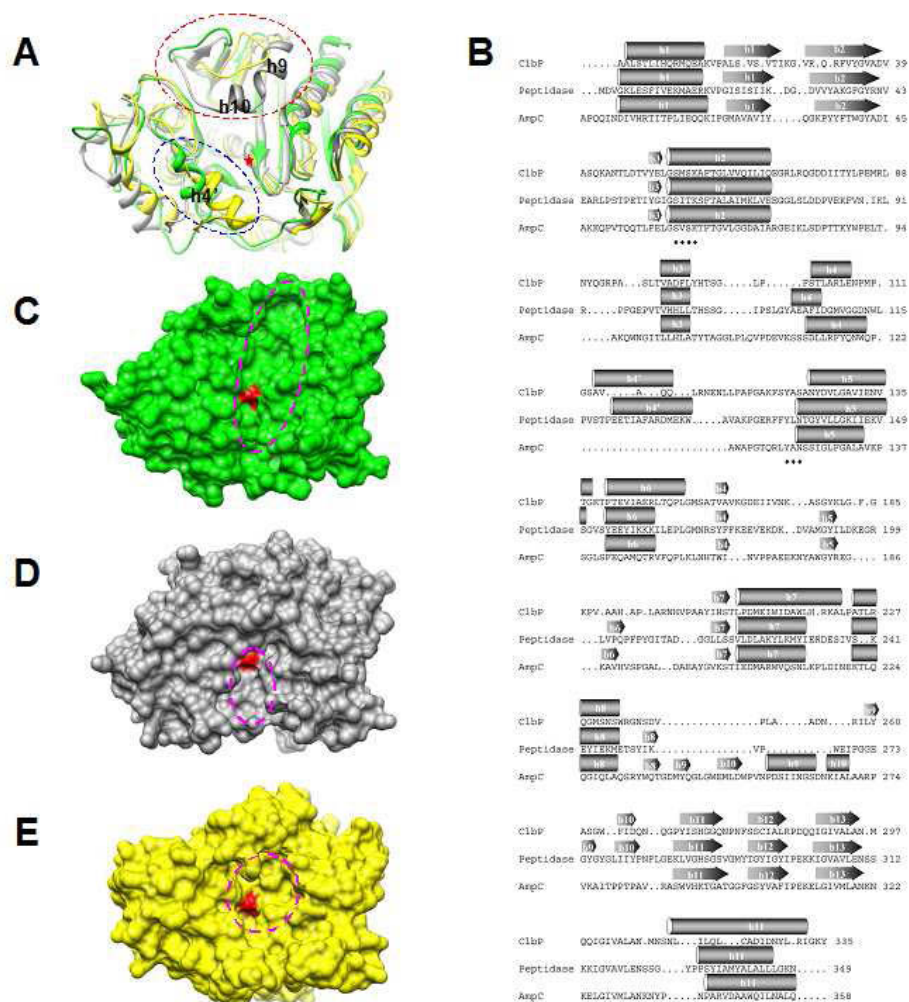


Figure 2.

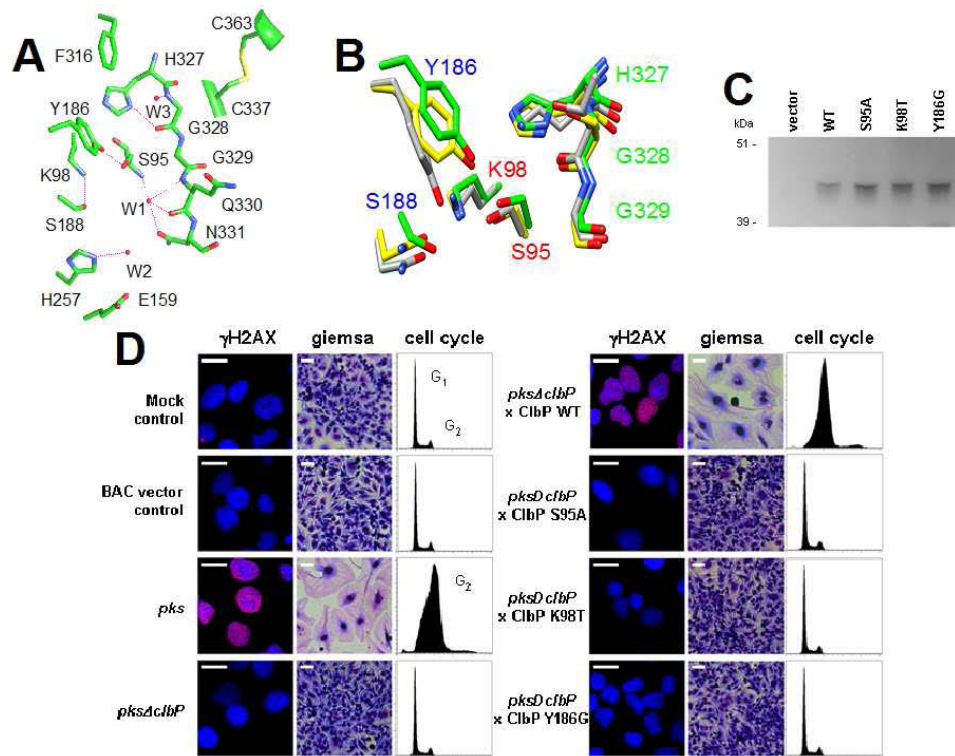
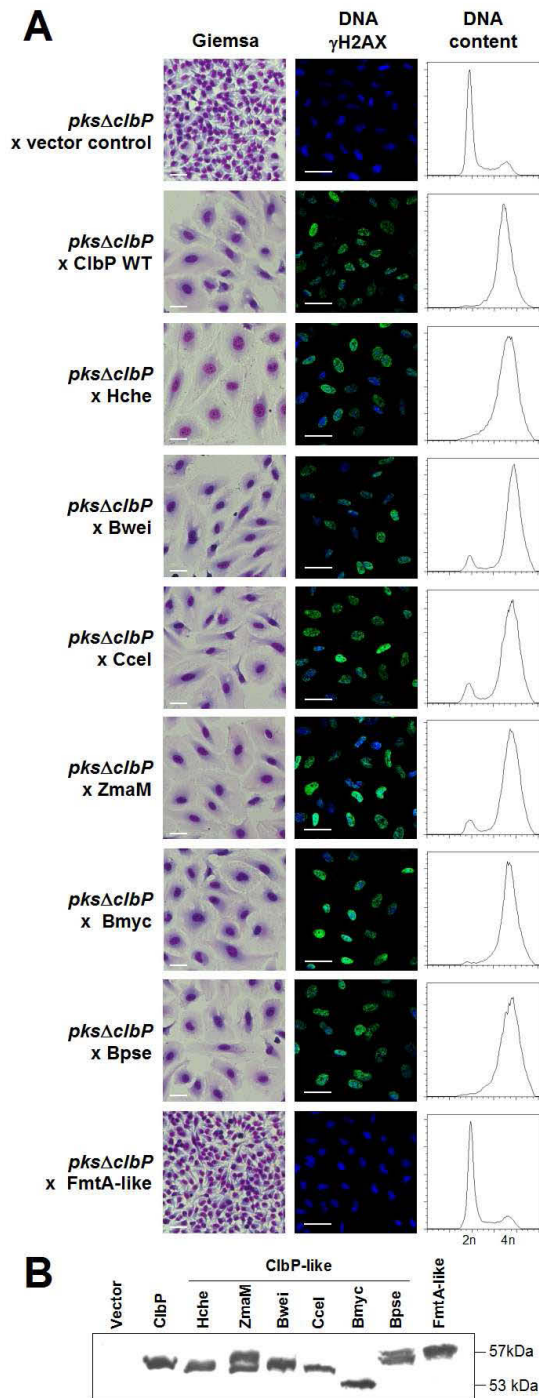


Figure 3.



Analysis of Structure–Function Relationships in the Colibactin-Maturing Enzyme ClbP

Antony Cougnoux¹, Lucie Gibold^{1,2}, Frederic Robin^{1,2}, Damien Dubois^{3,4}, Nathalie Pradel⁵, Arlette Darfeuille-Michaud¹, Guillaume Dalmasso¹, Julien Delmas^{1,2} and Richard Bonnet^{1,2}

1 - M2iSH, INSERM U1071, INRA USC2018, Clermont Université, Université d'Auvergne, 63000 Clermont-Ferrand, France

2 - Centre Hospitalier Universitaire, 63000 Clermont-Ferrand, France

3 - Université de Toulouse, ENVT, UMR 1225, Toulouse F-31076, France

4 - INSERM, UMR 1043, Toulouse F-31300, France

5 - IRD, UMR 180, IFR-BAIM, Universités de la Méditerranée, 13288 Marseille, France

Correspondence to Richard Bonnet: richard.bonnet@u-clermont1.fr
<http://dx.doi.org/10.1016/j.jmb.2012.09.017>

Edited by R. Huber

Abstract

pks genomic island of *Escherichia coli* is involved in the synthesis of the non-ribosomal peptide-type genotoxin colibactin, which has been suggesting as affecting the host immune response and having an impact on cancer development. The *pks*-encoded enzyme ClbP is an atypical peptidase that contributes to the synthesis of colibactin. In this work, we identified key features of ClbP. Bacterial fractionation and Western-blot analysis revealed the docking of ClbP to the bacterial inner membrane via a C-terminal domain harboring three predicted transmembrane helices. Whereas only one helix was necessary for the location in the inner membrane, the complete sequence of the C-terminal domain was necessary for ClbP bioactivity. In addition, the N-terminal sequence of ClbP allowed the SRP/Sec/YidC- and MreB-dependent translocation of the enzymatic domain in the periplasmic compartment, a feature also essential for ClbP bioactivity. Finally, the comparison of ClbP structure with that of the paralogs FmtA-like and AmpC revealed at an extremity of the catalytic groove a negative electrostatic potential surface characteristic of ClbP. Site-directed mutagenesis experiments identified in this zone two aspartic residues that were important for ClbP bioactivity. Overall, these results suggest a model for precolibactin activation by ClbP and pave a way for the design of inhibitors targeting colibactin production.

© 2012 Elsevier Ltd. All rights reserved.

Introduction

Nougayrède *et al.* recently reported a genomic island designated *pks* that is involved in the production of a genotoxin named colibactin.¹ The *pks* island is observed in both *Escherichia coli*, a versatile commensal bacterial inhabitant of human large intestine, and other *Enterobacteriaceae*.² *pks*-harboring bacteria (*pks*+) induce DNA double-strand breaks in eukaryotic cells both *in vitro* and *in vivo*.³ These lesions activate the ATM-CHK2 signaling pathway. However, the cells can present signs of incomplete DNA repair, which are marked by megalocytosis, G2/M cell cycle arrest, gene mutations, chromosomal instability and anchorage-

independent colony formation, demonstrating the carcinogenic potential of colibactin.^{1,3}

Colibactin has never been purified despite several attempts, and its precise structure remains unknown. However, genetic and functional analyses of the *pks* island indicate that it encodes non-ribosomal peptide synthetases (NRPSs) and polyketide synthetases (PKSs), which are required for the production of one or more polyketide (PK)–non-ribosomal peptide (NRP) hybrid compounds.⁴ In addition to NRPS and PKS, the *pks* island encodes nine accessory, tailoring and editing enzymes, among which ClbP is crucial in colibactin maturation.⁵ We recently determined the crystal structure of the ClbP major domain and identified

functional homologous proteins sharing about 30% identity with ClbP.⁵ These proteins, such as ZmaM and XcnG, which are involved in the synthesis of the PK-NRP compounds zwittermicin and xenocoumarin, respectively,^{6,7} belong to the S12 enzyme family of the MEROPS classification,⁸ share a similar putative topology and are able to transcomplement *pks* island defective for *clbP*. They form a new subgroup of enzymes associated with NRP assembly line and are required for the maturation and activation of NRP compounds.^{5,7} These enzymes, designated NRP-associated peptidases, could be an interesting target for molecular probes or molecules affecting the production of NRP bioactive molecules.

The protein FmtA-like, a protein of unknown function,⁹ and AmpC, a β -lactamase,^{10,11} also belong to the S12 enzyme family of the MEROPS classification and share 30% identity with ClbP. FmtA-like, in addition, exhibits a predicted topology similar to that of NRP-associated peptidases. However, these proteins are not able to transcomplement *pks* island defective for ClbP, indicating that they are paralogs of ClbP. By analogy with the well-known enzyme AmpC, Dubois *et al.* identified the catalytic triad required for ClbP activity.⁵ However, the structural elements, which explain the difference in activity among FmtA-like, AmpC and ClbP, remain unknown. In addition, the role of the N- and C-terminal sequences of ClbP has never been investigated.

In the present study, we investigated the role of ClbP C- and N-terminal sequences in the location and bioactivity of ClbP. By comparison with AmpC and FmtA-like crystal structures, we also identified key determinants of ClbP activity far from the catalytic site, suggesting an atypically large binding site.

Results

ClbP is an inner membrane protein

ClbP (504 amino acids) was predicted *in silico* to be an inner membrane protein harboring three C-terminal transmembrane helices (residues 390–412, 433–455 and 465–485). A C-terminal RGS6His-tagged ClbP protein was expressed in *E. coli*, and Western-blot analysis was undertaken from whole cell lysates and cellular fractions to establish the location of ClbP in bacteria. A reactive band with a molecular mass compatible with RGS6His-tagged ClbP mass was observed in the whole cell lysate of *E. coli* (Fig. 1a). In addition, RGS6His-tagged ClbP was predominantly found in the inner membrane fraction and the spheroplast fraction (a cell fraction containing the inner membrane), but not in the periplasmic or cytoplasmic fractions (Fig. 1a). These results show that ClbP is located in the bacterial inner membrane.

HA-tagged ClbP harboring three, two, one or no C-terminal helices were expressed in *E. coli*. Anti-HA Western-blot analysis was performed on cellular fractionation experiments. It revealed that only one helix was required to associate ClbP with the inner membrane (Fig. 1b). To investigate the role of helices in ClbP activity, we expressed constructs in *E. coli* containing a *pks* island defective for *clbP* (*pBACpks Δ clbP*) and we infected eukaryotic cells with the bacteria. ClbP derivatives harboring one, two or no helices did not induce megalocytosis and cell cycle arrest in response to the infection (Fig. 1c and d). In contrast, expression of the construct harboring the three putative helices induced megalocytosis and cell cycle arrest, which show restoration of a full functional *pks* island (Fig. 1c and d). In conclusion, the location of ClbP at the inner membrane is necessary but not sufficient to assume colibactin production and the complete ClbP C-terminal domain is required for ClbP bioactivity and colibactin production.

ClbP possesses an N-terminal cleavable signal sequence, which is SRP/Sec/YidC and MreB dependent

The N-terminal residues of ClbP were predicted *in silico* to be a transmembrane helix (residues 12–30) or a cleavable Sec-dependent signal sequence (ALA-QE 30-31). This putative signal sequence of ClbP (ClbPss) plus the first six amino acids of the mature part were fused with the GFP (green fluorescent protein) reporter to form the fusion protein ClbPss-GFP. Western-blot analysis using a GFP-specific antibody was performed against *E. coli* cells expressing ClbPss-GFP. It revealed the precursor (ClbPss-GFP) and mature (GFP) forms of the construct (Fig. 1e). Cellular fractionation of *E. coli* expressing ClbPss-GFP revealed only little ClbPss-GFP in cytoplasmic fraction and an accumulation of both the complete construct and its processed form in periplasmic fraction (Fig. 1d). In addition, we analyzed the bioactivity of ClbP harboring its N-terminal sequence or not as described above. The results showed that the N-terminal sequence of ClbP was required to transcomplement *E. coli* harboring a *pks* island defective for *clbP* (Fig. 1c). All together, these results confirm the presence of a cleavable signal sequence in ClbP N-terminal sequence, which mediate a translocation to the periplasm that was required for ClbP bioactivity.

ClbP signal sequence is SRP/Sec/YidC and MreB dependent

SRP, SecY and YidC are major components of the translocation/insertion pathway used by inner membrane proteins. This translocation system recognizes in the protein target an N-terminal signal

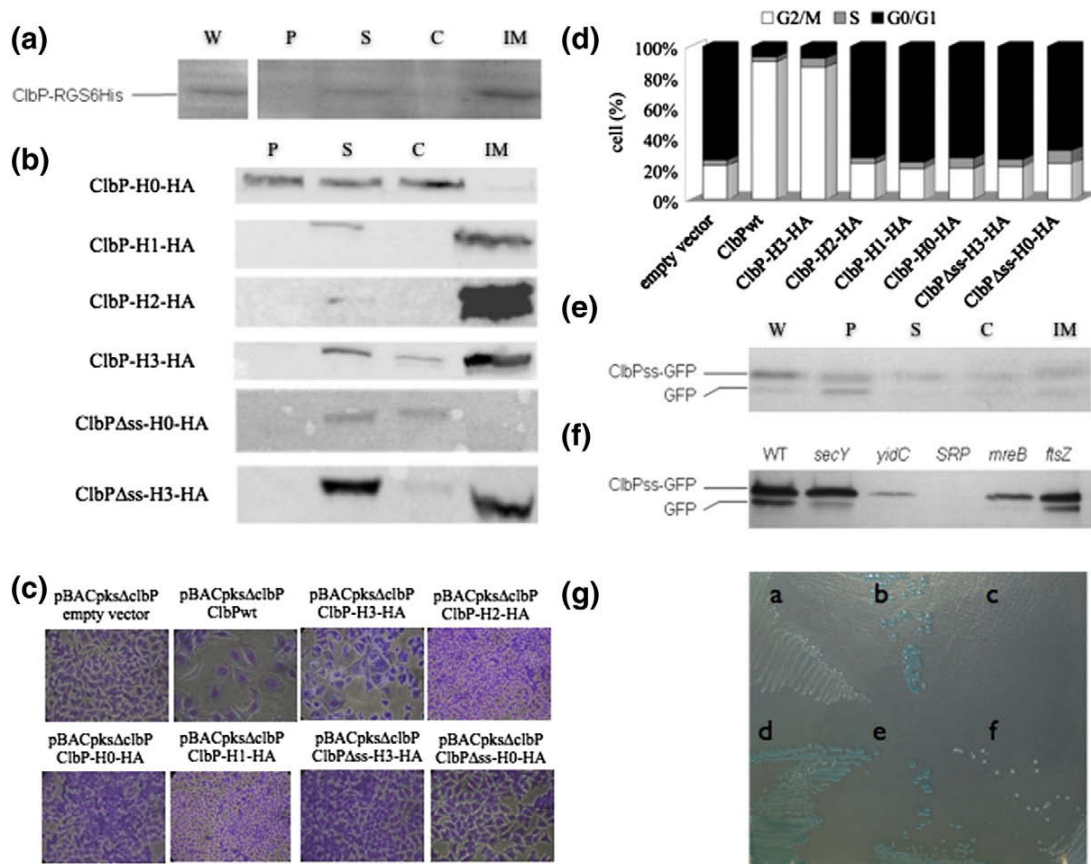


Fig. 1. The role of C- and N-terminal sequences in the location and bioactivity of ClbP. Each construction used was represented in Fig. S1. (a) Western-blot analysis of RGS6His-tagged ClbP location in *E. coli* cellular fractions. (b) Western-blot analysis of HA-tagged ClbP derivatives harboring one, two or three transmembrane helices in *E. coli* cellular fractions. (c and d) The genotoxicity of *E. coli* DH10B pBAC Δ clbP expressing ClbP derivatives. The genotoxicity was assessed by megalocytosis assays (c) and cell cycle analysis, which is reported as percentage of cells in phase G0/G1, S and G2/M (d). (e) Cellular location of ClbPss-GFP and its processed mature counterpart GFP in *E. coli* cellular fractions. (f) Cellular location of ClbPss-GFP and its processed mature counterpart GFP in wild-type *E. coli* (WT) and mutants for *secY*, *yidC*, *mreB*, *SRP* and *ftsZ*. (g) Alkaline phosphatase (PhoA) export monitored on Mueller Hinton agar containing the chromogenic substrate BCIP. (g-a) *E. coli* defective for PhoA; (g-b) *E. coli* defective for PhoA transcomplemented by a plasmid encoding PhoA; (g-c) no bacteria; (g-d, g-e and g-f) *E. coli* defective for PhoA transcomplemented by a plasmid encoding PhoA fused to ClbP-signal peptide (ClbPss-PhoA), (g-d) ClbP-signal peptide and the three transmembrane helices of ClbP (ClbPss-PhoA-ClbP3TM) (g-e) or the three transmembrane helices of ClbP (PhoA-ClbP3TM) (g-f). Blue staining occurs on *E. coli* expressing PhoA or PhoA-fusion proteins in the bacterial periplasm. W, whole cells; P, periplasm; S, spheroplast; C, cytoplasm; SM, solubilized inner membrane fraction. ClbP-H0-HA, ClbP derivative without transmembrane helix; ClbP-H1-HA, ClbP derivative harboring one transmembrane helix; ClbP-H2-HA, ClbP derivative harboring two helices; ClbP-H3-HA, ClbP harboring three helices; ClbP Δ ss-H0-HA, ClbP derivative devoid of both signal sequence and transmembrane helix; ClbP Δ ss-H3-HA, ClbP derivative devoid of signal sequence; ClbPss-GFP, ClbP signal sequence fused to GFP.

sequence, which was cleaved at the time of translocation into periplasm and/or insertion to the inner membrane.¹² MreB and FtsZ are proteins of bacterial cytoskeleton, and it has been suggested that MreB is involved in cytoplasmic trafficking of proteins toward the Sec translocation system.¹³ To further analyze the translocation system involved in ClbP maturation, we analyzed the cleavage of ClbPss-GFP fusion protein by Western blot using

SRP, *secY*, *yidC*, *mreB* and *ftsZ* *E. coli* knockout or conditional isogenic mutants. The processing of ClbPss-GFP was impaired in *secY*, *yidC* and *mreB* mutants in comparison with the wild-type strain, but not in *ftsZ* mutant (Fig. 1e). No signal was observed with the *SRP* mutant, suggesting an absence of translocation and a degradation of the nascent peptide in the cytoplasm. These data suggest that actin-like cytoskeleton MreB and the SRP/Sec/YidC

pathway perform the translocation of ClbP to the inner membrane by targeting ClbP signal sequence.

The N-terminal domain of ClbP is periplasmic

ClbP is associated with the inner membrane because of its C-terminal sequence and harbors a cleavable N-terminal signal sequence allowing the translocation to periplasm, and hence, we decided to investigate whether the central enzymatic domain associated with ClbP C-terminal sequence is located at the periplasmic side of the bacterial inner membrane. Alkaline phosphatase enzymatic domain of PhoA, which is only active in periplasm, was used as a reporter of the periplasmic location as previously reported.^{14,15} The PhoA domain was therefore fused to ClbP signal sequences, ClbP C-terminal sequence (380–504) or to both, and the constructs were expressed in *E. coli* DH5 α F'. The bacteria expressing PhoA fused to ClbPss and fused to ClbPss plus ClbP C-terminal sequence harbored in the presence of PhoA chromogene substrate a blue staining, which revealed the presence of a PhoA active domain. This activity indicated a location in the periplasmic compartment (Fig. 1g-d and g-e). In contrast, *E. coli* DH5 α F' expressing PhoA fused to ClbP C-terminal sequence alone remained white (Fig. 1g-f), although the construct was expressed in the bacteria, as assessed by Western-blot analysis (data not shown). The absence of signal sequence prevents the location of enzymatic domain at the periplasmic side (Fig. 1g-f) without affecting the inner membrane association (Fig. 1b). Overall, these data suggest that an enzymatic domain docked to the inner membrane but facing inward prevents ClbP bioactivity (Fig. 1c and d). In addition, these results show that the N-terminal cleavable signal sequence of ClbP is sufficient and necessary to translocate the partner enzymatic domain into periplasm and confirmed therefore the location of the ClbP enzymatic domain on the periplasmic side of the bacterial inner membrane.

FmtA-like has peptidase activity

FmtA-like is the most closely related paralog of ClbP.^{1,5} We therefore speculated that comparison of these enzymes could help in understanding the specificity of ClbP activity. We first compared their enzymatic activities against peptidase substrates formed by amino acids conjugated to *p*-nitroanilines. The catalytic efficacy k_{cat}/K_m of FmtA-like against Gly and L-Ala-*p*-nitroanilines was 154.4 and 78.9 $\mu\text{M}^{-1} \text{min}^{-1}$, respectively, which is more than 10,000-fold higher than that of ClbP (0.005 and 0.020 $\mu\text{M}^{-1} \text{min}^{-1}$). However, FmtA-like had no detectable activity against D-Ala-D-Ala- and D-Ala-*p*-nitroanilines ($k_{\text{cat}} < 10^{-4} \text{min}^{-1}$). These results sug-

gest that FmtA-like is not a DD-peptidase and confirm the differences in enzymatic activity between ClbP and FmtA-like.

Comparison of FmtA-like, AmpC and ClbP structures

FmtA-like is the most closely related paralog of ClbP. However, its structure is unknown. We therefore determined the crystal structure of FmtA-like enzymatic domain (33–369) by X-ray diffraction to compare its structure with those of ClbP and AmpC, another paralog of ClbP.^{10,11} The crystal structure was solved at 3.2 Å resolution by molecular replacement using AmpC and ClbP as search models. The stereochemical parameters of the model are reported in Table S3. The Ramachandran plot¹⁶ showed 97.6% in favored region and 2.4% in allowed region of ϕ/ψ space. The final model included 4 FmtA-like enzymatic domain monomers and 35 water molecules.

The overall structure of FmtA-like enzymatic domain is composed of two regions: one α/β region and one all-helical region (Fig. S2). The organization of the corresponding secondary structure elements was conserved in ClbP and FmtA-like except two additional small helices (H5 and H10) in ClbP, but differed from that of AmpC (Fig. S3). The α/β domain consisted of a main 10-stranded antiparallel β -sheet flanked by seven helices, and the all-helical region consisted of four helices. Between these two structural domains was a groove, which spreads over the β_3 strand and contains the catalytic triad, as observed for ClbP, AmpC and other S12 enzymes. The first motif of the catalytic triad (consensus: S-X-X-K) is strictly conserved in S12 MEROPS enzymes and consists of serine residue, acting as catalytic nucleophile during the acylation step of the catalytic reaction, plus a lysine residue one helix turn downstream. The second motif (consensus: Y-X-N/S) consists of Y190, which is usually responsible for the deacylation step, and N192, which is frequently involved in substrate binding in S12 enzymes. The third motif, which is less conserved (consensus: K/R/H-T/S/G-G) and usually also involved in substrate binding, consists of the HSG sequence in FmtA-like, KTG in AmpC and HGG in ClbP. The positioning of these catalytic motifs was strikingly conserved in the active sites of FmtA-like and ClbP [PDB (Protein Data Bank) ID: 3O3V] and AmpC (PDB ID: 1FSY) (Fig. S4), suggesting a common catalytic mechanism of acylation–deacylation.

The structural elements surrounding the catalytic motifs are known to determine the specificity of S12 MEROPS enzymes.¹⁷ We therefore investigated the positioning of these elements in ClbP and in the two paralogs FmtA-like and AmpC. The major differences observed between these enzymes were the positioning and the size of the loop, which connects helix H9 and the β_5 β -strand of FmtA-like. In ClbP

and FmtA-like, this loop was composed of 21 and 19 residues, respectively. The corresponding structural element comprised a loop (279–310) and the helices H13 (280–287) and H14 (289–292) in AmpC. In ClbP structure, it is positioned toward the solvent and, consequently, allows a large extension of the catalytic groove toward the N-terminal extremity of the $\beta 3$ strand. In AmpC and, at lesser extent, in FmtA-like, the loop shifted toward the catalytic groove in close contact with the helices H16 and H11 in AmpC and FmtA-like, respectively, and covered the N-terminal extremity of the $\beta 3$ strand (Fig. 2). These elements interacted in FmtA-like *via* one close contact between the residue 307 of the loop and the residue 363 of helix H11 (distance between the $O^{\delta 1}$ atom of Asn307 and the $O^{\epsilon 2}$ atom of Glu363, 4.4 Å; Fig. 2a), despite a 2.3- to 2.7-Å shift of helix H11 toward the solvent compared to AmpC and ClbP. Likewise, 26 close contacts (distances comprised between 3.2 and 4.5 Å; Fig. 2b) were observed in AmpC in this area, since the distances between the corresponding residues are higher than 10 Å in ClbP (Fig. 2c). Consequently, the long catalytic groove observed in ClbP is severely shortened in FmtA-like and particularly in AmpC at the N-terminal extremity of the $\beta 3$ strand.

The surface electrostatic potentials of ClbP, FmtA-like and AmpC

Electrostatic fields around the surface of proteins are known to play an important role in molecular recognition, owing to the attractive or repulsive effect of the coulombic potential at long-range distance.¹⁸ We therefore calculated the electrostatic potential molecular surface of ClbP, FmtA-like and AmpC enzymes (Fig. 3a–c) using APBS software.¹⁹ The results revealed very different electrostatic potentials in their catalytic grooves. On the side harboring the catalytic groove, we observed in ClbP a strong negative electrostatic potential surface restricted to the catalytic groove, especially at the two extremities (Fig. 3a). A diffuse negative electrostatic potential was observed in FmtA-like with strong charged environment at the C-terminal extremity of the $\beta 3$ strand, as in ClbP, but a weak or neutral charged environment at the N-terminal extremity of the $\beta 3$ strand (Fig. 3c). In contrast, the electrostatic potential surface of AmpC catalytic pocket was mainly positive (Fig. 3b) as observed with other closely related S12 enzymes such as carboxypeptidases or transpeptidases (data not shown). Overall, ClbP differed therefore from AmpC and FmtA-like

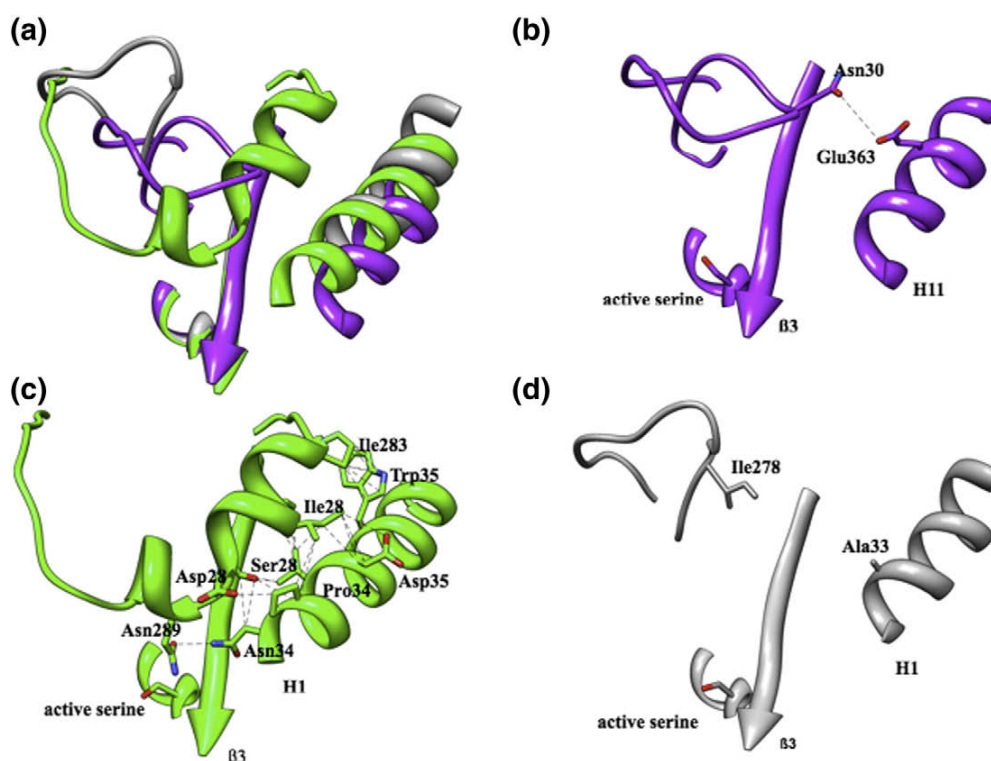


Fig. 2. Superimposed structures of ClbP, FmtA-like and AmpC (a) and close contact analysis in FmtA-like (b), AmpC (c) and ClbP (d) and structures at the vicinity of $\beta 3$ strand. FmtA-like is purple, AmpC is green and ClbP is gray. Close contacts between residues are indicated by gray broken lines. Carbon atoms are purple for FmtA-like, green for AmpC and gray for ClbP; oxygen atoms are red and nitrogen is blue.

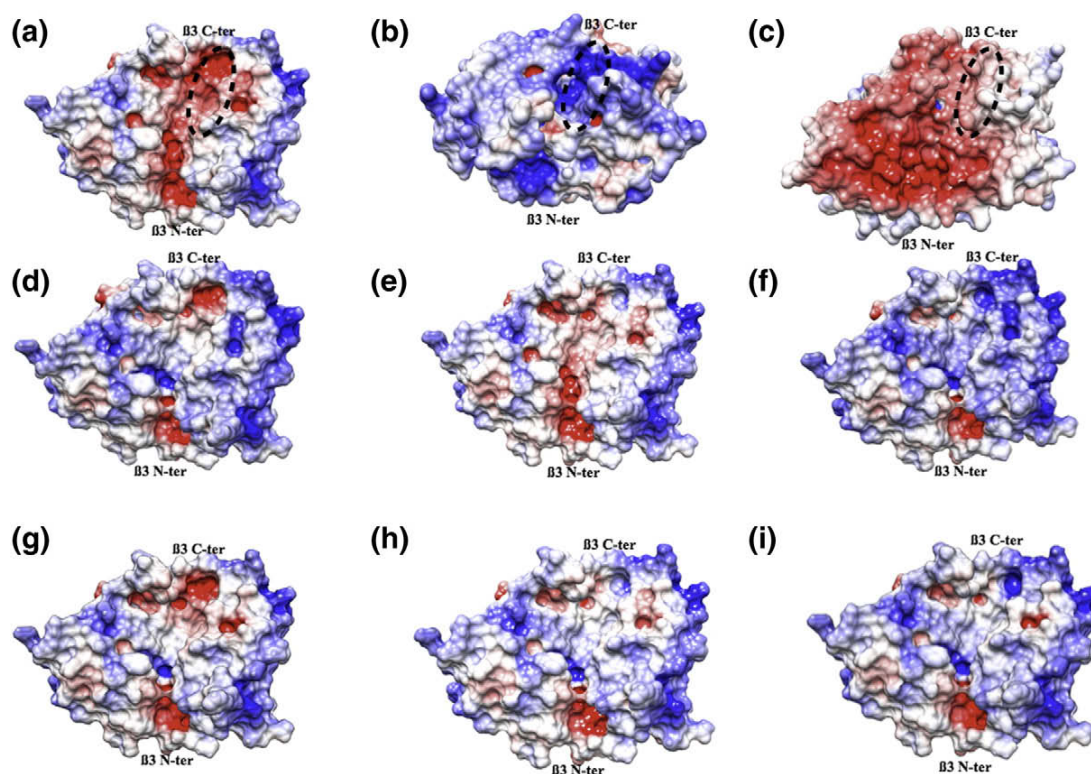


Fig. 3. Molecular surfaces strained according to the electrostatic potential for ClbP (a), AmpC (b), FmtA-like (c) and ClbP mutants harboring the substitutions D334Q (d), D336L (e), D334Q-D336L (f), D334E (g), D336N (h) and D334Q-D336N (i). The surfaces were colored red for negative charge, white for neutral and blue for positive. Helix H13 from ClbP and the corresponding helices in AmpC (H16) and FmtA-like (H11) were circled in black.

by a strong negative electrostatic potential at the N-terminal extremity of the $\beta 3$ strand in an area edged by the C-terminal part of helix H11.

The distribution of charges on protein surface was compared to that of known substrates or putative ligands. Benzylpenicillin and cephalothin, which are the substrates of AmpC, exhibited an overall negative electrostatic potential (Fig. S5). This can favor their binding into the positive-charged catalytic pocket of AmpC. Interestingly, the final bioactive PK-NRP compound zwittermicin and its putative preprocessing counterpart pre-zwittermicin, which Kevany *et al.* suggest is cleaved by the NRP-associated peptidase ZmaM,⁶ exhibited overall a positive electrostatic potential surface (Fig. S5). The positive potential surface of NRP compounds can therefore match with the negative potential surface of ClbP, which may be a key feature of this enzyme.

The residues, which may be responsible for the negative electrostatic potential surface in ClbP catalytic groove, were investigated using molecular modeling approaches. The negative-charged residues were replaced *in silico* by the corresponding residues in AmpC or FmtA-like, and the electrostatic potential was calculated to evaluate their contribu-

tion to the electrostatic properties of ClbP. We found at the C-terminal part of helix H11 two positions harboring negative-charged residues, D334 and D336. The substitutions D334Q and D336L, which introduce residues observed in AmpC, abrogated and decreased the negative electrostatic potential (Fig. 3d-f), respectively. In combination, this substitution conferred a positive electrostatic potential, as observed in AmpC. The substitutions D334E and D336N, which introduce residues observed in FmtA-like, slightly decreased and abolished the negative electrostatic potential observed at the N-terminal extremity of the $\beta 3$ strand (Fig. 3g-i). Overall, these results suggest that residues D334 and D336 may be key elements in ClbP bioactivity.

The positions 334 and 336 of ClbP are key residues for ClbP bioactivity

To evaluate the contribution of residues 334 and 336 to the bioactivity of ClbP, we introduced substitutions D334E (FmtA-like residue), D334Q (AmpC residue), D336N (FmtA-like residue) and D336L (AmpC residue) alone and in combination into RGS6His-tagged ClbP by site-directed mutagenesis.

These constructs and wide-type RGS6His-tagged ClbP were expressed in *E. coli* containing a *pks* island defective for ClbP. The constructs were expressed at a similar level, as assessed by Western-blot analysis (data not shown). The genotoxicity of bacteria was then investigated using comet assays, which show at early-time postinfection (24 h) the DNA damage induced by colibactin-producing *E. coli*. Megalocytosis assays and cell cycle analysis were also performed to reveal the presence of patent lesions at 72 h. RGS6His-tagged ClbP mutants at position 334 or 336 induced a significant decrease in mean comet tail moment in comparison with wide-type RGS6His-tagged ClbP (RGS6His-tagged ClbP *versus* RGS6His-tagged ClbP-D334E/Q or RGS6His-tagged ClbP D336L/N, $p < 0.05$; Fig. 4a). This decrease revealed an alteration in the genotoxic activity of bacteria expressing ClbP mutants. However, the genotoxicity was not enough impaired to affect the cell cycle arrest and megalocytosis induced by the bacteria at 72 h (Fig. 4b and c). In contrast, the expression of the mutants harboring the combinations of substitutions at positions 334 and 336 strongly impaired the genotoxicity of bacteria; DNA damage was almost abrogated at 24 h (95% and 92.8%, $p < 0.01$; Fig. 4a), cell cycle arrest was significantly affected ($p < 0.001$; Fig. 4b) and megalocytosis decreased at 72 h (Fig. 4c). All together, these results demonstrate that positions 334 and 336, which are located far from the catalytic triad

(distance between C $^{\alpha}$ atoms of active serine Ser64 and residues 334/336, 19.5/19.0 Å, respectively), are still important for ClbP bioactivity.

Discussion

The bacterial genomic island *pks* is involved in the production of enzymes, which synthesize a PK-NRP-type genotoxin named colibactin.¹ This type of toxin can have a major impact on host-microbe relationships.^{20,21} It has been suggested that colibactin affects the host immune response and has an impact on colorectal cancer development.^{1,3} The *pks*-encoded enzyme ClbP has been identified as a prototype of an atypical peptidase subgroup of the S12 MEROPS family. These enzymes, designated NRP-associated peptidases, contribute to the maturation of NRP compounds such as colibactin.⁵ In this work, we identified key features of ClbP bioactivity: (i) the association of ClbP with the bacterial inner membrane, (ii) the exposition of its enzymatic domain in the bacterial periplasm and (iii) key residues of the enzymatic domain involved in the negative electrostatic potential of the catalytic groove. These findings enlighten our understanding of colibactin synthesis and pave the way for the design of inhibitors targeting colibactin production.

We show that the N-terminal sequence is a cleavable signal sequence mediating a translocation

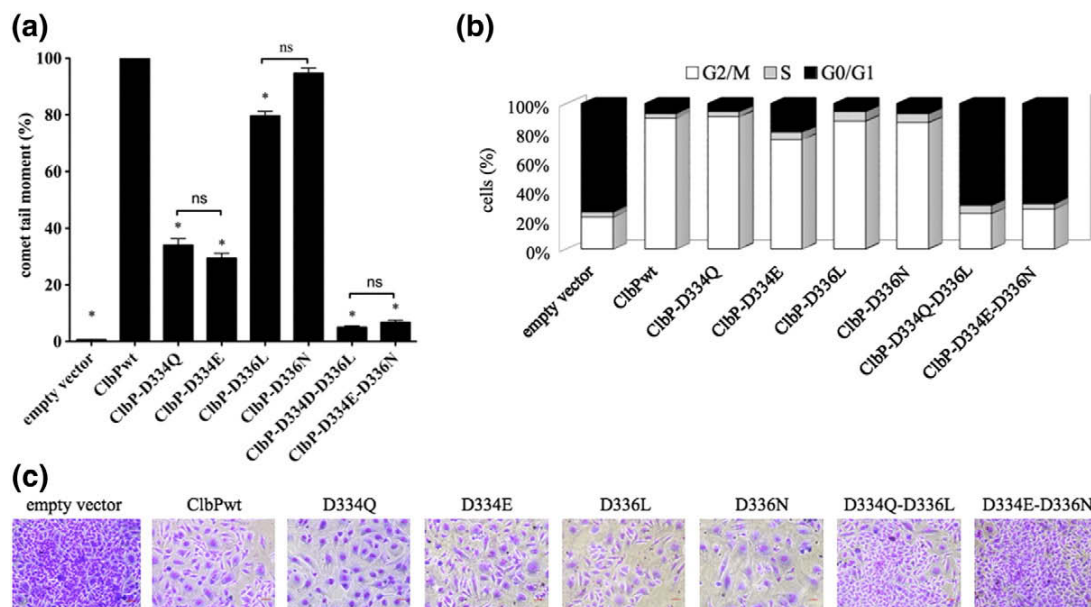


Fig. 4. The genotoxicity of *E. coli* pBACpks Δ clbP expressing ClbP or ClbP mutants at position 334 and/or 336 was investigated with HeLa cells. (a) DNA damage at 24 h postinfection assessed by comet assays. The results are expressed in percentage of comet tail moment (100% was attributed to bacteria expressing wide-type ClbP). (b) Cell cycle analysis at 72 h postinfection reported as percentage of cells in phase G0/G1, S and G2/M. (c) Cytopathic effect (megalocytosis) at 72 h postinfection.

of the passenger domain to periplasm *via* the actin-like cytoskeleton MreB and the SRP/Sec/YidC pathway. A similar putative signal sequence is observed in other NRP-associated peptidases,^{6,7} suggesting that the enzymatic domain of NRP-associated peptidases such as ClbP is located outside the cytoplasm. In contrast, the other colibactin-producing enzymes, essentially NRPS and PKS, are devoid of putative signal sequence and are known to be in the cytoplasm. ClbP acts therefore outside the cytoplasm, after assembly line chains of NRP biosynthesis, and thus is involved in the final steps of colibactin maturation. It allows maturation of the genotoxin colibactin outside the cytoplasm at a distance from the DNA of the producing bacteria probably in order to protect its genome. This location also makes ClbP and the other NRP-associated peptidases more accessible than the other enzymes of NRP assembly lines and therefore renders them key targets for the control of the bioactivity of PK–NRP-producing gene clusters.

We observed that ClbP is not free in the periplasm but associated with the inner membrane because of the C-terminal domain. Other S12 enzymes of MEROPS classification such as transpeptidases are associated with the inner membrane with their enzymatic domain in the periplasm.^{22,23} They are docked to the membrane *via* a single N-terminal transmembrane helix,¹² while the docking of ClbP to the membrane is dependent on the C-terminal domain, which contains three predicted transmembrane helices. Only one helix was sufficient to assure the docking of ClbP to the membrane. However, the complete C-terminal domain was required for the bioactivity of ClbP. The presence of three transmembrane helices is compatible with a periplasmic enzymatic domain in the periplasm and a C-terminal extremity in the cytoplasm. Interestingly, several NRP-associated peptidases are fused with ABC exporters belonging to the family of cyclic peptides.^{5–7} The role of ClbP C-terminal domain may also be important therefore for interactions with the cytoplasmic assembly line and/or efflux pumps carrying out the export of precolibactin from the cytoplasm to the periplasm.

We analyzed the phylogeny of the condensation domain in the NRPS enzymes encoded by *pks* island as previously reported²⁴ and found typical starter C-type domain in ClbN. This type of domain acylates the first amino acid of NRP compounds with typically β -hydroxyl fatty acids to form a lipopeptide.²⁵ Reimer *et al.* recently reported the maturation of the lipopeptides pre-xenocoumacin in an NRP compound devoid of fatty acid (xenocoumacin) by the NRP-associated peptidase XcnG.⁷ Kevany *et al.* previously proposed a similar biosynthesis pathway for zwittermicin A.⁶ Accordingly, precolibactin may therefore be a lipopeptide inserted in the bacterial inner membrane and actively

released from this membrane by the inner-membrane-located enzyme ClbP. This model (Fig. 5) can explain the importance of ClbP location at the inner membrane and why *pks*-harboring bacteria killed by heat or gentamicin are not genotoxic.¹

The enzymatic domain of ClbP was compared to those of the two paralogs FmtA-like and AmpC. AmpC is a class C β -lactamase, which is well known to confer a resistance to β -lactam antibiotics and to be widespread in Gram-negative bacteria. FmtA-like is encoded by the genome of *Staphylococcus* isolates. We found that this protein harbors a peptidase activity different from that of ClbP, but its biological function remains unknown. Phylogenetic analysis has shown that FmtA-like is the most closely related paralog of ClbP, while AmpC is more distantly related.⁵ The determination of FmtA-like crystal structure shows that this phylogenetic lineage is also apparent in the three-dimensional structure of these enzymes and helps to identify interesting characteristic of the ClbP catalytic groove. From AmpC to ClbP, we observed a progressive opening of the catalytic groove at the vicinity of the N- and C-terminal extremities of the β 3 strand and helix H11 (or their structural equivalent H16 and H13), respectively (Fig. 2). This opening is associated with an inversion of the electrostatic potential at the surface of the catalytic groove. Whereas the AmpC electrostatic potential surface of AmpC active site is overall positive, we observed a strong negative electrostatic potential focused in the catalytic groove of ClbP. These negative charges may be conserved around the NRP-associated peptidases and fit the positive charges of NRP compounds produced by assembly lines related to the *pks*-encoded assembly line.

In FmtA-like, the negative electrostatic potential is more diffuse than in ClbP and the charge is less marked especially in the vicinity of the N- and C-terminal extremities of the β 3 strand and helix

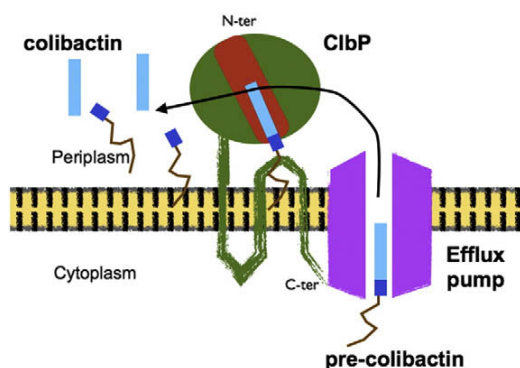


Fig. 5. ClbP role in the maturation of colibactin. Negatively charge active site of ClbP was colored red. The putative positively charged colibactin compound was colored blue.

H11. These data suggest that this latter zone is crucial for the specificity of ClbP bioactivity. Accordingly, we found two acidic residues important for ClbP bioactivity in this zone in helix H13 (positions 334 and 336). Their substitutions with the corresponding residues of FmtA-like and AmpC affected the bioactivity of ClbP in a synergistic manner and, consequently, eliminated the genotoxicity mediated by *pks*-harboring *E. coli*. This inhibition of ClbP bioactivity is probably due to a modification of the electrostatic potential, but we cannot exclude that ClbP activity is impaired by steric hindrance. Several residues have been mutated in the vicinity of the catalytic triad and the C-terminal part of the β 3 strand without affecting ClbP bioactivity,⁵ suggesting that the non-catalytic residues in this zone are less important for ClbP bioactivity than the extremity of the catalytic groove around the N-terminal part of the β 3 strand. These results also reveal that the ClbP catalytic groove is atypically long and is designed to accommodate a filiform compound.

In conclusion, the NRP-associated peptidase ClbP is a periplasmic enzyme that is docked to the bacterial inner membrane by its C-terminal domain. This topology allows maturation of the genotoxin colibactin outside the cytoplasm at a distance from the genomic material of the producing bacteria. It also makes ClbP an interesting drug target for control of the genotoxic effect of colibactin-producing bacteria. We also identified at a distance from the catalytic triad an electronegative area as a key determinant of ClbP bioactivity that may be targeted by drugs.

Materials and Methods

Strains, plasmids and media

Strains and plasmids used in this study are listed in Table S1, and the primers are in Table S2. *E. coli* strains and *Staphylococcus aureus* mu50 were grown at 37°C in Mueller Hinton medium or on Luria–Bertani (LB) plates. Amoxicillin (32 µg/ml) (Sigma), kanamycin (50 µg/ml) (Sigma), chloramphenicol (25 µg/ml), glucose (0.2%), arabinose (0.2%), lactose (0.2%), isopropyl β -D-1-thiogalactopyranosid (1 mM) (Euromedex) or tetracyclin (200 ng/ml) (Sigma) was added as required to the medium.

Cell induction assays of translocation pathway and fractionation

Cell induction assays of translocation pathway were performed as previously described.^{26–28} We used the cold-sensitive *secY-cs* and *yidC* conditional mutants CU164A and MYC-cs, in which protein export is partially slowed down at 37°C and is severely affected upon exposure to 23°C. The following strains harboring p61ClbPssGFP were incubated in LB–amoxicillin–glucose (0.2%) medium at permissive temperature (37°C for cold-

sensitive strains) with shaking until early exponential phase: MC4100A, MC1060 and MC1000 wild type; conditional translocation mutant CU164A (Δ *secY*) and MYC-cs (Δ *yidC*); HPT406A (SRP mutant), YLS-3 (Δ *mreB*) and YLS-24 (Δ *ftsZ*) strains. The cultures were incubated on ice for 3 min and then kept at 23°C (nonpermissive temperature) for 1 h for conditional mutant induction. The cells were harvested by centrifugation, washed once with LB and resuspended in 0.2% LB–amoxicillin–arabinose medium and incubated for a further 3 h at 23°C with shaking for induction of ClbPss-GFP expression.

Cellular fractionation experiments were performed as previously described.^{26,27} Crude extracts were prepared by centrifugation of 1 ml culture at OD_{600nm}=0.5. The pellets were suspended in 0.1 ml Tris–HCl buffer (40 mM, pH 8.0) and 0.02 ml of 1% protamine sulfate. Periplasm and spheroplasts were prepared by lysozyme/ethylene-diaminetetraacetic acid/cold osmohock method and separated by centrifugation. Membranes and cytoplasmic soluble fractions were prepared from spheroplasts by breaking the cells by sonication and separated by ultracentrifugation at 120,000g for 90 min. Protease inhibitor Complete (Roche Diagnostics, Mannheim, Germany) was added at the maximum recommended concentration in all buffers used.

Western-blot analysis

Western blots were performed with polyclonal rabbit anti-GFP (Invitrogen) at a dilution of 1:10,000 and monoclonal anti-RGS6His according to the manufacturer's instructions (Qiagen). Western blots performed with monoclonal rabbit anti-HA (Sigma) at a dilution of 1:1000, mice anti-alkaline phosphatase (Millipore) at a dilution of 1:5000 and horseradish-peroxidase-tagged goat anti-rabbit or goat anti-mouse secondary antibodies (Cell Signaling) at a dilution of 1:10,000 according to the manufacturer's instructions.

Monitoring of alkaline phosphatase activity

ClbP fused to alkaline phosphatase were obtained by recombination using primers listed in Table S2. Cloning experiments were performed with In-Fusion HD cloning kit (Clontech) using the primers described in Table S2. *E. coli* cells were grown on LB plates with amoxicillin (32 µg/ml) and 50 µg/ml of 5-bromo-4-chloro-3-indolyl phosphate (the chromogenic substrate of alkaline phosphatase purchased from Sigma) at 30°C for 2 days. The blue-stained colony-forming units reveal the presence of the PhoA alkaline phosphatase domain in the periplasm as previously reported.¹⁵

Cloning and site-directed mutagenesis

fmtA-like gene was amplified from *S. aureus* strain mu50. After a double digestion with NdeI and NotI (New England Biolabs), PCR products were cloned in pET-M13 (Bioscience) expression vectors. ClbP mutants were obtained by site-directed mutagenesis using primers listed in Table S2. Site-directed mutagenesis experiments were performed with QuikChange™ Site-Directed Mutagenesis Kit II (Stratagene, La Jolla, CA) using the

mutagenic primers described in Table S2. The substitutions D334E, D336N, D334E-D336N, D334Q, D336L and D334Q-D336L were individually introduced in ClbP cloned into p33ClbP plasmid. The changes were confirmed by double-stranded DNA sequencing. The final p33clbP derivative vectors were used to transform *E. coli* DH10B pBAC*pk Δ clbP* strain, which is defective for ClbP production.

Eukaryotic cell culture, bacterial infections, comet assay and cell cycle analysis

Cell assays were performed as previously described, with minor modifications.^{1,5} For bacterial infections, overnight LB–amoxicillin cultures of bacteria were diluted in interaction medium [Dulbecco's modified Eagle's medium, 10% de complemented fetal calf serum (Lonza), 25 mM Hepes (Sigma) and 32 μ g/ml amoxicillin]. Tetracycline was added to overnight cultures and interacting medium to induce the expression of ClbP and derivatives. Approximately 50% confluent HeLa cell cultures (American Type Culture Collection CCL2) were then infected with a multiplicity of infection (number of bacteria per eukaryotic cell) of 500. Cells were washed twice in phosphate-buffered saline (pH 7.4) 3 h after inoculation and incubated in Dulbecco's modified Eagle's medium, 10% fetal calf serum and 200 μ g/ml gentamicin (Sigma) to stop the infection and were kept until analysis. Cell morphology was examined by Giemsa staining at 72 h postinfection. For cell cycle analysis, cellular suspensions were fixed and permeabilized with 70% ethanol at 48 h postinfection, then treated with 50 μ g/ml propidium iodide and 250 μ g/ml RNase A (Roche). DNA-content data were acquired with an SLR II or FACScalibur flow cytometer (Becton Dickinson) and analyzed with FlowJo software (Tree Star). DNA breaks were analyzed by comet assay using SYBR Green as fluorescent DNA staining agent as described previously.^{29,30} Comets were scored with the image analysis software CometScore™ (TriTek Corporation) using the mean comet tail moment (product of tail length and fraction of DNA in the tail).

FmtA-like production and purification

pFmtA-like plasmid encodes the peptidase domain of FmtA-like without the first 87 nucleotides and the last 384 nucleotides of the *fmtA*-like gene. These sequences encode the signal peptide and three putative hydrophobic transmembrane helices that were removed to improve cytoplasmic expression and protein purification. *E. coli* BL21(DE3) Δ *ampC* transformed with pFmtA-like plasmid was cultured in Terrific broth medium supplemented with 50 μ g/ml kanamycin, 0.4 M D-sorbitol (MP) and 2.5 mM β -betaine (Sigma). The cell culture was grown at 37 °C with shaking up to reach an OD_{620nm} = 0.7. Overexpression of the *fmtA*-like encoding gene was induced with 0.2 mM isopropyl β -D-thiogalactopyranoside overnight at 25 °C with shaking. Cells were harvested by centrifugation (10,000g for 10 min at 4 °C) and suspended in 20 mM Tris–HCl buffer (pH 7.5). After addition of 5 μ g lysozyme (2 h at 4 °C), the cells were disrupted by ultrasonic treatment (four times for 30 s, each time at 20 W). The extract was clarified by centrifugation at 14,000g for

60 min at 4 °C. After addition of 2 mg DNase I (Roche) for 15 min at room temperature, the supernatant (10,000g for 60 min at 4 °C) was dialyzed overnight against 20 mM Tris–HCl buffer (pH 7.5) and filtered (0.22- μ m filter) (Millipore). Purification was carried out by ion-exchange chromatography on a 10-ml HiTrap™ Q Sepharose™ high-performance column (GE Healthcare) equilibrated with 20 mM Tris–HCl (pH 7.5) and eluted with a linear NaCl gradient (0–1 M). The FmtA-like-containing elution peak (160 mM NaCl) was then dialyzed 4 h against 50 mM of 4-morpholineethanesulfonic acid buffer (pH 6.0) and purified by a second ion-exchange chromatography on a HiTrap™ Q Sepharose™ high-performance column equilibrated with 50 mM of 4-morpholineethanesulfonic acid buffer (pH 6.0) and eluted with a linear NaCl gradient (0–1 M). The FmtA-like-containing elution peak (50 mM NaCl) was extensively dialyzed overnight against 5 mM Tris–HCl and 50 mM NaCl buffer (pH 7) and concentrated by ultrafiltration to 10 mg/ml for crystallization. The concentration was carried out using YM-10000 centricon (Millipore). The enzyme was more than 95% homogeneous as determined by Coomassie blue staining after 10% SDS-PAGE analysis.

Enzymatic activity

Standard photometric assays using *p*-nitroaniline-conjugated amino acids X-pNA (X: Gly, Gly-Gly, L-Ala, D-Ala, D-Ala-D-Ala, Bz-Asn or L-Lys) as substrates were performed at 25 °C in 5 mM Tris–HCl and 50 mM NaCl, pH 7.0, containing 4 mM of the substrate. Data were collected on the UVICON930 system. Absorption coefficients at 405 nm of 9450 M⁻¹ cm⁻¹ for X-pNA were used to calculate the enzymatic activity. K_m and k_{cat} were determined from kinetic data using ENCORA software (Kluyver laboratory).

Crystallization and structure determination

fmtA-like crystals were grown in hanging drops over a well solution of 0.2 M ammonium acetate (Sigma), 0.1 M sodium acetate trihydrate (pH 4.6) (Sigma) and 30% (w/v) polyethylene glycol 4000 (Sigma). Crystals reached a size of 40 μ m \times 40 μ m \times 70 μ m within 4 months. Before data collection, crystals were immersed for about 30 s in a cryoprotectant solution [30% sucrose, 0.2 M ammonium acetate, 0.1 M sodium acetate trihydrate (pH 4.6) and 30% (w/v) polyethylene glycol 4000] and were flashcooled in liquid nitrogen. Data were collected using a Q315r ADSC-CCD detector on European Synchrotron Radiation Facility (ESRF) beamline ID-23-1 at the ESRF (Grenoble, France). Reflections were indexed, integrated and scaled using CCP4 package.^{31–33} The initial model was obtained by molecular replacement with the program Phaser,³⁴ and the ClbP X-ray structures (PDB ID: 3O3V) were used as search models. The structure was automatically and manually refined with the REFMAC5 and Coot programs, respectively.^{35,36} Cross-validation was used throughout, and 5% of the data were used for the R_{free} calculation. The stereochemical quality of the models was monitored with the PROCHECK program.³⁷ Ramachandran plots were calculated by RAMPAGE.¹⁶ Processing and crystallographic refinement statistics for FmtA-like crystal structure are listed in Table S3 (PDB ID: 4GDN).

Bioinformatic analysis and molecular modeling

ClbP primary sequence (accession number CAJ76284.1) was analyzed using InterProScan.³⁸ Signal peptide analysis was performed using SignalP 4.0 server³⁹ and lipoP 1.0 server.⁴⁰ Transmembrane helix location was predicted using TMHMM server† and TMBASE.⁴¹

Substitutions at position 334 and/or 336 were introduced in ClbP structure using Chimera version 1.5.3‡ visualization software and Dunbrack rotamer library.⁴² The models were then protonated with PROPKA using the PDB2PQR server,⁴³ and the protons were subjected to energy minimization (100 steps steepest descent) using the MMTK routine under Chimera⁴⁴ with AMBER parameters. The quality and stereochemistry of the models were evaluated using the program PROCHECK program.³⁷ The APBS program¹⁹ was used to calculate the electrostatic surface potential map of protein models, which were visualized with the Chimera program.

Acknowledgements

We thank Christelle Blavignac, Claire Szczepaniak and Lorraine Gameiro from the Centre Imagerie Cellulaire Santé (Université d'Auvergne) for their technical assistance in photonic microscopy and in flow cytometry. We thank Marlène Jan and Roland Perroux for their excellent technical assistance. This work was supported by the Ministère de la Recherche et de la Technologie, l'Institut National de la Santé et de la Recherche Médicale (INSERM U1071) and l'Institut National de la Recherche Agronomique (USC INRA 2018). These experiments were performed on the ID-23-1 beamline at the ESRF, Grenoble, France. We are grateful to Local Contact at ESRF for providing assistance in using beamline ID-23-1.

Supplementary Data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jmb.2012.09.017>

Received 17 August 2012;

Received in revised form 20 September 2012;

Accepted 21 September 2012

Keywords:

serine peptidases;
colibactin;
non-ribosomal peptide;
electrostatic surface charge;
Escherichia coli

† software@cbs.dtu.dk

‡ <http://plato.cgl.ucsf.edu/chimera/>

Abbreviations used:

ESRF, European Synchrotron Radiation Facility;
PK, polyketide;
PKS, polyketide synthetase;
NRP, non-ribosomal peptide;
NRPS, non-ribosomal peptide synthetase.

References

- Nougayrède, J.-P., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G. *et al.* (2006). *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science*, **313**, 848–851.
- Putze, J., Hennequin, C., Nougayrède, J.-P., Zhang, W., Homburg, S., Karch, H. *et al.* (2009). Genetic structure and distribution of the colibactin genomic island among members of the family *Enterobacteriaceae*. *Infect. Immun.* **77**, 4696–4703.
- Cuevas-Ramos, G., Petit, C. R., Marcq, I., Boury, M., Oswald, E. & Nougayrède, J.-P. (2010). *Escherichia coli* induces DNA damage *in vivo* and triggers genomic instability in mammalian cells. *Proc. Natl Acad. Sci. USA*, **107**, 11537–11542.
- Homburg, S., Oswald, E., Hacker, J. & Dobrindt, U. (2007). Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. *FEMS Microbiol. Lett.* **275**, 255–262.
- Dubois, D., Baron, O., Cougnoux, A., Delmas, J., Pradel, N., Boury, M. *et al.* (2011). ClbP is the prototype of a peptidase subgroup involved in the biosynthesis of nonribosomal peptides. *J. Biol. Chem.* <http://dx.doi.org/10.1074/jbc.M111.221960>.
- Kevany, B. M., Rasko, D. A. & Thomas, M. G. (2009). Characterization of the complete zwittermixin A biosynthesis gene cluster from *Bacillus cereus*. *Appl. Environ. Microbiol.* **75**, 1144–1155.
- Reimer, D., Pos, K. M., Thines, M., Grün, P. & Bode, H. B. (2011). A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat. Chem. Biol.* **7**, 888–890.
- Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. & Barrett, A. J. (2008). MEROPS: the peptidase database. *Nucleic Acids Res.* **36**, D320–D325.
- Komatsuzawa, H., Choi, G. H., Fujiwara, T., Huang, Y., Ohta, K., Sugai, M. & Suginaka, H. (2000). Identification of a *fmtA*-like gene that has similarity to other PBPs and β -lactamases in *Staphylococcus aureus*. *FEMS Microbiol. Lett.* **188**, 35–39.
- Usher, K. C., Blaszczyk, L. C., Weston, G. S., Shoichet, B. K. & Remington, S. J. (1998). Three-dimensional structure of AmpC β -lactamase from *Escherichia coli* bound to a transition-state analogue: possible implications for the oxyanion hypothesis and for inhibitor design. *Biochemistry*, **37**, 16082–16092.
- Jacoby, G. A. (2009). AmpC β -lactamases. *Clin. Microbiol. Rev.* **22**, 161–182; Table of Contents.
- Izard, J. W. & Kendall, D. A. (1994). Signal peptides: exquisitely designed transport promoters. *Mol. Microbiol.* **13**, 765–773.
- Shih, Y.-L., Kawagishi, I. & Rothfield, L. (2005). The MreB and Min cytoskeletal-like systems play independent roles in prokaryotic polar differentiation. *Mol. Microbiol.* **58**, 917–928.

14. Marrichi, M., Camacho, L., Russell, D. G. & DeLisa, M. P. (2008). Genetic toggling of alkaline phosphatase folding reveals signal peptides for all major modes of transport across the inner membrane of bacteria. *J. Biol. Chem.* **283**, 35223–35235.
15. Drew, D., Sjöstrand, D., Nilsson, J., Urbig, T., Chin, C., de Gier, J.-W. & von Heijne, G. (2002). Rapid topology mapping of *Escherichia coli* inner-membrane proteins by prediction and PhoA/GFP fusion analysis. *Proc. Natl Acad. Sci. USA*, **99**, 2690–2695.
16. Laskar, A., Chatterjee, S., Roy, A., Kumar Dey, S. & Mandal, C. (2011). Molecular modeling and structural analysis of five SE clan (S12 family) serine proteases. *Asian J. Biotechnol.* **3**, 435–448.
17. Sharp, K. A. & Honig, B. (1990). Electrostatic interactions in macromolecules: theory and applications. *Annu. Rev. Biophys. Chem.* **19**, 301–332.
18. Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
19. Demangel, C., Stinear, T. P. & Cole, S. T. (2009). Buruli ulcer: reductive evolution enhances pathogenicity of *Mycobacterium ulcerans*. *Nat. Rev., Microbiol.* **7**, 50–60.
20. Fischbach, M. A. & Walsh, C. T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev.* **106**, 3468–3496.
21. Delfosse, V., Girard, E., Birck, C., Delmarcelle, M., Delarue, M., Poch, O. *et al.* (2009). Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family. *PLoS One*, **4**, e4712.
22. Antalis, T. M., Buzza, M. S., Hodge, K. M., Hooper, J. D. & Netzel-Arnett, S. (2010). The cutting edge: membrane-anchored serine protease activities in the pericellular microenvironment. *Biochem. J.* **428**, 325–346.
23. Rausch, C., Hoof, I., Weber, T., Wohlleben, W. & Huson, D. H. (2007). Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution. *BMC Evol. Biol.* **7**, 78.
24. Konz, D., Doekel, S. & Marahiel, M. A. (1999). Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin. *J. Bacteriol.* **181**, 133–140.
25. Pradel, N., Delmas, J., Wu, L. F., Santini, C. L. & Bonnet, R. (2009). Sec- and Tat-dependent translocation of β -lactamases across the *Escherichia coli* inner membrane. *Antimicrob. Agents Chemother.* **53**, 242–248.
26. Santini, C. L., Ize, B., Chanal, A., Müller, M., Giordano, G. & Wu, L. F. (1998). A novel Sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J.* **17**, 101–112.
27. Chen, M., Xie, K., Jiang, F., Yi, L. & Dalbey, R. E. (2002). YidC, a newly defined evolutionarily conserved protein, mediates membrane protein assembly in bacteria. *Biol. Chem.* **383**, 1565–1572.
28. Olive, P. L. & Banáth, J. P. (2006). The comet assay: a method to measure DNA damage in individual cells. *Nat. Protoc.* **1**, 23–29.
29. Collins, A. R. (2004). The comet assay for DNA damage and repair: principles, applications, and limitations. *Mol. Biotechnol.* **26**, 249–261.
30. Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R. *et al.* (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 235–242.
31. Batty, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 271–281.
32. Evans, P. (2006). Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **62**, 72–82.
33. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674.
34. Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A. *et al.* (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **67**, 355–367.
35. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **66**, 486–501.
36. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.
37. Lovell, S. C., Davis, I. W., Arendall, W. B., 3rd, de Bakker, P. I. W., Word, J. M., Prisant, M. G. *et al.* (2003). Structure validation by C α geometry: phi, psi and C β deviation. *Proteins*, **50**, 437–450.
38. Syed, A. & Upton, C. (2006). Java GUI for InterProScan (JIPS): a tool to help process multiple InterProScans and perform ortholog analysis. *BMC Bioinformatics*, **7**, 462.
39. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods*, **8**, 785–786.
40. Rahman, O., Cummings, S. P., Harrington, D. J. & Sutcliffe, I. C. (2008). Methods for the bioinformatic identification of bacterial lipoproteins encoded in the genomes of Gram-positive bacteria. *World J. Microbiol. Biotechnol.* **24**, 2377–2382.
41. Hofmann, K. & Stoffel, W. (1993). TMBASE—a database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler*, **374**, 166.
42. Dunbrack, R. L., Jr. (2002). Rotamer libraries in the 21st century. *Curr. Opin. Struct. Biol.* **12**, 431–440.
43. Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G. & Baker, N. A. (2007). PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35**, W522–W525.
44. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612.

3. Inhibition de l'activité génotoxique et protumorale de bactéries productrices de colibactine.

Publication n°4 :

The NRP peptidase ClbP as a target for the inhibition of genotoxicity, cell proliferation and tumorigenesis mediated by *pks*-harbouring bacteria.

Antony Cougnoux, Julien Delmas, Lucie Gibold, Chiara Romagnoli, Gabriel Cuevas-Ramos, Eric Oswald, Arlette Darfeuille-Michaud, Fabio Prati, Guillaume Dalmaso and Richard Bonnet.

Soumission en attente car dépôt de brevet en cour.

Nous avons montré que ClbP est une peptidase à sérine active dont l'activité enzymatique est essentielle à l'activité génotoxique de l'îlot génomique *pks*. Le domaine enzymatique de ClbP est localisé dans le périplasme bactérien, un site facilement accessible à des xénobiotiques. Ces deux caractéristiques font de ClbP une cible de choix pour le développement de molécules « drug-like » bloquant la synthèse de colibactine. Nous avons donc entrepris la recherche de petites molécules susceptibles de fixer spécifiquement et donc d'inhiber le site actif de ClbP pour bloquer l'activité délétère de bactéries productrices de colibactine.

L'industrie pharmaceutique porte un intérêt croissant aux dérivés de l'acide boronique pour les utiliser dans l'inhibition d'enzymes à sérine active [385]. Un principe actif basé sur cette molécule, un inhibiteur du protéasome, est déjà utilisé dans le traitement du myélome multiple (bortezomib -Velcade[®]-, approuvé par la FDA en 2003) et plusieurs autres font l'objet d'essais cliniques dans différents domaines médicaux comme l'antibiothérapie (GSK'52, Glasgow Smith Kline) ou le traitement du psoriasis (AN2728, Anacor pharmaceutical) [386]. L'intérêt de ces molécules tient à l'atome de bore. Il présente une structure atomique marquée par la présence d'une orbitale *p* vide qui lui permet d'accepter un doublet d'électrons en provenance de nucléophiles biologiques. Ainsi, il peut former une liaison covalente réversible avec la fonction hydroxyle des sérines réactives. Les acides boroniques peuvent donc agir comme des inhibiteurs compétitifs des enzymes à sérine active. De plus, les acides boroniques portent un substituant, qui peut être modifié pour adapter la spécificité de la molécule à la cible considérée.

Le site actif de ClbP forme un sillon électronégatif se prolongeant vers l'extrémité N-

terminale du brin β 11. Par contre, le site actif se ferme précocement à son extrémité C-terminale. Afin d'identifier des acides boroniques susceptibles de fixer efficacement la sérine active de ClbP, nous avons donc sélectionné trois acides boroniques dépourvus de charge négative et de taille réduite (composés 1 à 3, **Table 1**). La première molécule dérive de l'acide phényle boronique et les 2 autres, d'un acide boronique sulfonamide [387]. Ces molécules candidates ont été insérées dans une chimiothèque contenant 450 acides boroniques et l'ensemble a été criblé *in silico* pour leur capacité à fixer le site actif de ClbP par la méthode du « docking » grâce au logiciel Autodock Vina [388]. Un composé chargé négativement et de grande taille, connu pour fixer le site actif de l'enzyme à sérine active AmpC (composé 4, **Table 1**), ne pouvait pas être positionné dans le site actif de ClbP par « docking ». A l'inverse, les composés 1 à 3 présentaient un positionnement compatible avec une acylation de la sérine active de ClbP et le score d'interaction (Efficacité de fixation : quantité d'énergie libre de fixation par atome lourd) classait ces composés dans les douze meilleurs ligands putatifs de ClbP. Ces résultats suggéraient que les trois molécules candidates étaient susceptibles d'inhiber l'activité enzymatique de ClbP.

Nous avons déterminé les constantes d'inhibition (K_i) des 4 composés précédents vis-à-vis de ClbP, de la ClbP-like ZmaM, et du plus proche paralogue, FmtA-like par spectrophométrie à l'aide de substrats chromogènes. Les composés 1 à 3 présentaient une activité inhibitrice contre ClbP et ZmaM avec des valeurs de K_i comprises entre 240-440 nM et 1,43-4,56 μ M respectivement (**Table 1**), alors que le composé 4 n'avait pas d'activité détectable contre ces deux enzymes (K_i , > 1 mM). Inversement, les composés 1 à 3 n'avaient pas ou peu d'activité contre FmtA-like (K_i , 0,966 mM à > 1 mM), contrairement au composé 4 (K_i , 530 nM). Ces données étaient cohérentes avec les prédictions obtenues *in silico*. Les trois molécules candidates présentaient une activité nanomolaire contre ClbP et leur efficacité de fixation ($\sim -0,9$ kcal/mol, **Table 1**) était compatible avec un développement pour des applications thérapeutiques [389]. De plus, on observait un lien entre la structure des composés et leur spécificité d'action ; l'absence de charge négative et une taille réduite étaient associées à l'activité inhibitrice sur ClbP et son homologue ZmaM alors que la présence d'une charge négative et une augmentation de la taille amélioraient l'activité sur le paralogue FmtA-like.

Les données obtenues *in silico* suggèrent que les trois molécules retenues interagissaient avec le site actif de ClbP et étaient capables de former une liaison covalente avec la sérine active. Afin de vérifier ces hypothèses et comprendre comment des molécules

différentes peuvent inhiber de façon similaire ClbP, nous avons déterminé la structure tridimensionnelle de ClbP en complexe avec les deux meilleurs inhibiteurs (le composé 1 dérivé d'acide phényle boronique et le composé 2 dérivé d'acide boronique sulfonamide) par diffraction aux rayons X. Le traitement des données de diffraction a permis de résoudre la structure des complexes à la résolution de 2,19 pour le composé 1 et de 2,24Å pour le composé 2. La densité des électrons permettait de voir l'ensemble du site actif, les composés et une liaison de l'atome O_γ de la sérine active (Ser95) avec l'atome de bore. Les complexes obtenus confirmaient donc la capacité des composés retenus à interagir avec le site actif de ClbP.

Malgré des structures différentes, les deux types de composé interagissaient avec les mêmes résidus du site actif (**Figure 1, Publication 4**). Le substituant sulfonamide ou phényle des composés est orienté vers l'extrémité N-terminale du brin β11, contrairement à ce qui est observé avec d'autres enzymes voisines comme AmpC qui fixent leurs substrats et inhibiteurs au contact de l'extrémité C-terminale du brin β11 [387,390]. Ces données renforcent l'hypothèse évoquée précédemment d'une accommodation du substrat de ClbP vers l'extrémité C-terminale du brin β11. Par ailleurs, les structures obtenues sont marquées par une densité élevée d'interactions électrostatiques non conventionnelles, qui permettent de différencier les modalités de fixation des deux composés (**Figure 2, Publication 4**). Ces données révélaient donc deux modes de fixation permettant d'accéder au même site et un large éventail d'interactions pouvant être mises à profit pour concevoir de nouvelles molécules ciblant ClbP et ses homologues.

Afin d'explorer la capacité des composés à atteindre leur cible *in bacteria*, des souches de *E. coli pks+* ont été pré-incubées ou non avec les composés 1 à 3 à différentes concentrations et l'activité génotoxique des bactéries a été évaluée sur des cellules épithéliales d'origine humaine en culture ainsi que dans un modèle murin de anse ligaturée. L'activité génotoxique, évaluée au travers du test des mégaloctes et de l'immunomarquage de la forme phosphorylée de l'histone γH2Ax, était bloquée par les composés 1 à 3 de façon dose dépendante *in vitro* (**Figure 3 et 4, Publication 4**) sans affecter la survie bactérienne et sans toxicité apparente sur les cellules eucaryotes. Leur efficacité a été confirmée dans le modèle murin (**Figure 5, Publication 4**). Les composés 1 et 2 étaient les plus actifs et étaient capables d'abroger l'activité génotoxique. De même, nous avons exploré la capacité des composés à bloquer l'activité pro-proliférative des *E. coli pks+* sur une lignée cellulaire épithéliale intestinale d'origine humaine et dans le modèle murin de xénogreffe utilisé dans la première étude. Dans les deux cas, les composés prévenaient la prolifération cellulaire et la croissance tumorale

induite par les cellules infectées par les *E. coli pks+* (**Figure 6, Publication 4**). Comme attendu au regard des travaux précédents, cette inhibition était associée à un blocage de la sénescence cellulaire induite par les bactéries et une diminution de l'expression du facteur de croissance HGF (**Figure 6, Publication 4**). L'augmentation de l'HGF est un facteur de mauvais pronostic d'évolution du cancer colorectal [391,392]. L'inhibition de son récepteur c-Met est une option thérapeutique dans le cadre de cette maladie et fait l'objet de nombreuses recherches [392,393]. Par ailleurs, plusieurs molécules inhibant c-Met ont fait ou font l'objet d'essais cliniques avec des résultats encourageants [393]. Nos inhibiteurs, qui bloquent l'apparition de cassures de l'ADN ainsi que l'expression de HGF induite par les bactéries *pks+*, pourraient être une alternative ou un complément aux traitements évoqués ci-dessus.

The NRP peptidase ClbP as a target for the inhibition of genotoxicity, cell proliferation and tumourigenesis mediated by *pks*-harbouring bacteria

Antony Cougnoux^{1*}, Julien Delmas^{1,2*}, Lucie Gibold^{1,2}, Chiara Romagnoli³, Gabriel Cuevas-Ramos⁴, Eric Oswald^{4,5,6}, Arlette Darfeuille-Michaud^{1,7}, Fabio Prati³, Guillaume Dalmasso¹ and Richard Bonnet^{1,2}

* These authors contributed equally to this work

1. Clermont Université, Université d'Auvergne, Inserm U1071, INRA USC2018, 63000 Clermont-Ferrand, France
2. Centre Hospitalier Universitaire, 63000 Clermont-Ferrand, France
3. Dipartimento di Chimica, Università degli Studi di Modena e Reggio Emilia, Via Campi 183, 41100 Modena, Italy
4. INRA; USC 1360; Toulouse, France; Inserm; UMR1043; Toulouse, France; CNRS; UMR5282; Toulouse, France; Université de Toulouse;
5. UPS, Centre de Physiopathologie de Toulouse Purpan (CPTP); Toulouse, France;
6. CHU Toulouse; Hôpital Purpan; Service de bactériologie-Hygiène; Toulouse, France
7. Institut Universitaire de Technologie, Université d'Auvergne, Aubière, France.

Corresponding author: Richard Bonnet

Address : CHU, laboratoire de Bacteriologie

58 rue Montalembert

63000 Clermont-Ferrand, France

e-mail : rbonnet@chu-clermontferrand.fr, guillaume.dalmasso@udamail.fr

Tel : +33 (0)4 73 754 920

Short title: inhibition of the genotoxic and tumourigenic activities of *pks*-harbouring bacteria

Abstract

Microbiota may promote colorectal cancer. *Escherichia coli* isolated from colorectal cancer harbour the *pks* genomic island (*pks*⁺ *E. coli*) responsible for the synthesis of colibactin, a genotoxic compound. A transient contact of malignant cells with colibactin-producing *E. coli* increases tumour growth in a mouse xenograft model. ClbP was identified as a crucial enzyme in colibactin maturation. In this work, we identified boron-based small molecules, which inhibit ClbP, with K_i values in the nanomolar range (240 to 440 nM) and determined the crystal structure of ClbP in complex with these compounds. These inhibitors suppressed *in vitro* and *in vivo* the genotoxic activity of *pks*-harbouring *E. coli*. In addition, they prevented *in vitro* and *in vivo* cellular proliferation, HGF production and the tumourigenesis mediated by *pks*⁺ *E. coli*, demonstrating that ClbP is a valuable target for the prevention of the toxicity and the carcinogenic effect mediated by *pks*⁺ bacteria.

INTRODUCTION

Bacteria constitute about 90% of all cells in the human body¹. The majority of these bacteria colonize the colon. This resident bacterial consortium, called “microbiota,” generally influences the host in a beneficial fashion exerting protection against pathogens and contributing to metabolic pathways. However, it is also increasingly recognized as playing a major role in human health and diseases, notably colorectal cancer (CRC). The high density of microorganisms observed in the colon, (10^{13} commensal bacteria colonizing the colon vs from 10^2 for duodenum to 10^8 for the distal ileum), is associated with a greater risk of cancer than in the small intestine². In addition, mutant mice genetically susceptible to colorectal cancer or mice with chemical-induced chronic inflammation develop significantly fewer tumours under germ-free conditions than when they are kept under specific pathogen-free conditions and harbour conventional microbiota^{3,4}.

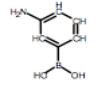
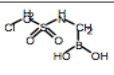
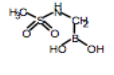
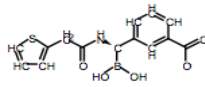
Escherichia coli is a consistent member of the human intestinal microbiota. As a commensal bacterium, *E. coli* coexists with its mammalian host in good harmony, promotes normal intestinal homeostasis and rarely causes diseases. However, some strains carry virulence genes enabling them to colonize certain ecological niches⁵. Mucosa-associated *Escherichia coli*, which are more frequently identified in colon tissue from patients with adenocarcinomas than in controls, can have pro-carcinogenic features^{6,7}. Certain *Escherichia coli* strains harbour the pathogenic island *pks*, which encodes enzymes responsible for the synthesis of a genotoxin called colibactin. Colibactin induces double-strand DNA breaks, cell cycle arrest and megalocytosis⁸. DNA damage was associated with mutations and chromosomal instability⁹; lesions frequently observed in cancers. Interestingly, *pks*-harbouring *E. coli* (*pks*+ *E. coli*) strains are observed in 50 to 60% of human CRC specimens *versus* 20% in controls^{10,11} and *pks*+ *E. coli* have a carcinogenic effect independent of inflammation¹¹. In addition, it was recently demonstrated in a xenograft mouse model that the underlying mechanism involves

pks+ *E. coli*-induced senescent cells, which promote tumour progression by the secretion of the growth factor, notably HGF.

Besides several attempts, colibactin has not been purified and its precise structure is yet unknown. However, the genetic and functional analyses of the *pks* island indicate that it encodes non-ribosomal peptide synthetases (NRPS) and polyketide synthetases (PKS) necessary to the production of one or more polyketide - non-ribosomal peptide (PK-NRP) hybrid compounds corresponding to colibactin¹². In addition to NRPS and PKS, the *pks* island encodes nine accessory, tailoring and editing enzymes. Among them, ClbP was identified as a crucial enzyme in *pks+* *E. coli*-induced genotoxic activity¹³. This active serine peptidase and its distant functional homologous, such as ZmaM, form a subgroup of enzymes, which activate NRP compounds^{13,14}.

In this work, we identified boron-based small molecules, which inhibit ClbP with K_i values in the nanomolar range and determined the crystal structure of ClbP in complex with these compounds. These inhibitors suppress the genotoxic activity of *pks*-harboring *E. coli*, both *in vitro* and *ex vivo*. In addition, they prevent the cellular proliferation and the tumourigenesis mediated by *pks+* *E. coli* *in vitro* and *in vivo*, demonstrating that ClbP is a valuable target for the prevention of the toxicity induced by *pks+* bacteria.

Table 1. Inhibition constant (K_i , μM), free binding energy (ΔG , Kcal/mol) and ligand efficiency (LE, Kcal/mol) of the compounds against the enzyme ClbP, the ortholog ZmaM and the paralog FmtA-like.

Compound		ClbP			ZmaM			FmtA-like		
ID	Structure	K_i	ΔG	LE	K_i	ΔG	LE	K_i	ΔG	LE
1		0.240	-9.02	-0.90	4.560	-7.28	-0.73	>1000	> -4.09	> -0.19
2		0.360	-8.78	-0.88	1.430	-7.96	-0.80	966.0	-4.11	-0.45
3		0.440	-8.66	-0.96	2.100	-7.74	-0.86	>1000	> -4.09	> -0.19
4		>1000	> -4.09	> -0.19	>1000	> -4.09	> -0.19	0.530	-8.55	-0.39

RESULTS

***In silico* screening of putative boron-based inhibitors of ClbP.**

Boronic acid-based compounds can form dative bond with biological nucleophiles such as reactive serine and therefore, might have a competitive inhibiting activity against ClbP. Because of the charge and the geometry of ClbP catalytic pockets^{13,15}, we selected small boronic acids devoid of negative-charged groups (compound 1 to 3, Table 1) as putative inhibitors of ClbP and a large and negative-charged compound as putative inactive compound (compound 4, Table 1). These four compounds were included in a set of 450 boronic acids, which were docked *in silico* into ClbP active site. Compound 4 was not accommodated in the vicinity of ClbP active serine, in contrast with compounds 1 to 3, which exhibited a positioning compatible with the acylation of the active serine residue. The binding modes of compounds were classified according to the ligand efficiency¹⁶. Interestingly, these bioinformatical approaches ranked the compounds 1 to 3 over the 12th position, suggesting there are ligands of ClbP active site.

Determination of the inhibition constants (K_i) of the selected compounds.

The inhibition constant (K_i) of compounds 1 to 4 was determined by spectrophotometry in a competition assay against ClbP, its functional homolog ZmaM and FmtA-like, which is the most closely related paralog of ClbP^{13,15}. Compounds 1 to 3 exhibited the best inhibitor activity against ClbP and ZmaM as expected, with K_i values in the ranges of 240-440 nM and 1.43-4.56 μ M, respectively (Table 1). In contrast, compound 4 had no detectable activity (K_i , > 1 mM). Inversely, compounds 1 to 3 exhibited no or only weak activity against FmtA-like (K_i , 0.966 mM to > 1 mM) and compound 4 exhibited better activity against FmtA-like (K_i , 530 nM) than against ClbP and its functional homologous ZmaM (K_i , > 1 mM). Overall, therefore we identified three nano-molar inhibitors of ClbP. The ligand efficiency of the inhibitors was similar despite differences in chemical structure with values around -0.9 kcal/mol (Table 1), suggesting that the selected molecules are interesting leads for medicinal chemistry compounds¹⁶.

Figure 1. Superimposition of ClbP active site in complex with inhibitors 1 and 2. The atoms of compounds overlap perfectly and are accommodated in the zone surrounded by the residues 95, 186 and 327-330. Carbon atoms are shown in gold for ClbP in complex with the compound 1, and in green for ClbP in complex the compound2. Oxygen atoms are shown in red, nitrogen atoms in blue, Boron atoms in magenta and sulfurs in yellow. The aromatic systems are indicated by dashed circles.

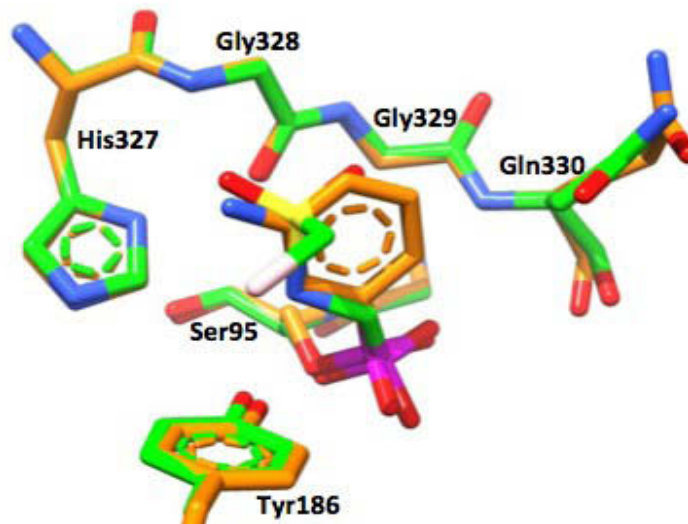
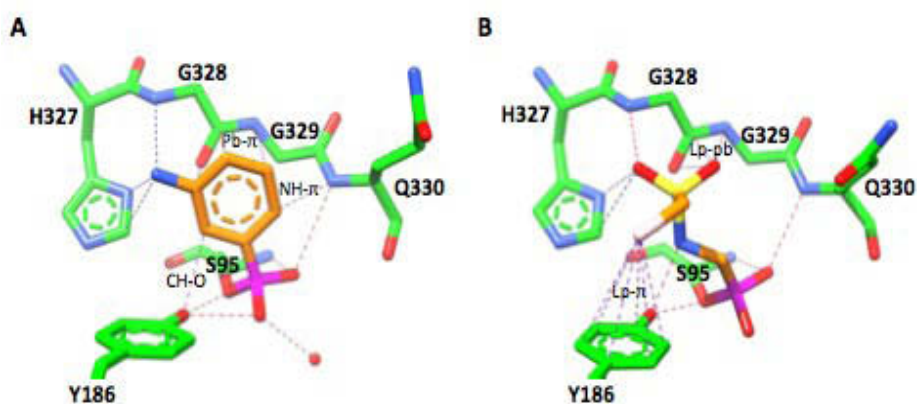


Figure 2. Structure of ClbP active site in complex with the inhibitors. The compounds (A, compound 1; B, compound 2) establish similar conventional interactions (van der Waals contact, NH...O and OH...O hydrogen bonds), but differ by the non-canonical interactions mediated by lone pair electrons (oxygen and chloride atoms, and π electrons in aromatic systems or the peptide bond G328-G329). Carbon atoms are shown in green for ClbP and in gold for the compounds. Oxygen atoms are shown in red, nitrogen atoms in blue, boron in magenta, sulfur in yellow and chloride in white. The aromatic systems are indicated by dashed circles. The hydrogen bonds are shown in red dashed lines. Van der Waals contacts are indicated by blue dashed lines. Non-canonical interactions are indicated by purple dashed lines (Pb- π , peptide bond- π stacking; Lp- π , lone pair... π interaction; Lp-pb, lone pair...peptide bond interaction; CH-O, CH...O hydrogen bond; NH- π , NH... π interaction).



Crystal structure of ClbP in complex with inhibitors

The structure of ClbP in complex with the best inhibitors (compounds 1 and 2) was determined by x-ray diffraction to confirm their ability to bind ClbP active site and understand how compounds having such differences in chemical structures can exhibit such closely related inhibiting activity. The structures were refined against diffraction data extending to 2.19 and 2.24 Å resolutions. The electron density map ([Supplementary figure 1](#)) showed well defined density throughout most of the structure. The three-dimensional structure of ClbP consists of two structural regions, all- α and α/β . The root-mean-square deviation (RMSD) of the C α positions is approximately 0.31 Å between the complex structures and the apoenzyme, showing no dramatic change in geometry resulting from the binding of inhibitors.

The positions of the inhibitors in the active site were unambiguously identified in the 2Fo-Fc omit density maps ([Supplementary figure 1](#)). The conformation of the inhibitors was similar in all monomers of the crystallographic asymmetric unit. Electron density connected the O γ of the catalytic residue Ser95 to the boron atoms of inhibitors, as expected. The other atoms of both compounds 1 and 2 were also accommodated near β -strand β 11 (residues 324-332) between residues 327-330 and Tyr150 ([Figure 1](#)). Intriguingly, most atoms of compounds perfectly overlapped despite major differences in chemical structure.

The moieties of compounds established similar conventional interactions (Van der Waals and hydrogen bond) with ClbP active site ([Figure 2](#)). Non-canonical interactions were also observed ([Figure 2](#)). For compound 1, two C atoms of the aromatic ring were at 2.8 Å and 3.0 Å of the Tyr188 hydroxyl function and the N atom of Gln330 respectively. The geometry of these close contacts is compatible with a NH \cdots π interaction involving Gln330 and a CH \cdots O interaction involving Tyr186^{17,18}. The plane of the aromatic ring was parallel with β -strand β 11 and established a π -stacking-like interaction¹⁹ with the peptide bond between residues Gly328 and Gly329. For compound 2, the chlorine atom established with the aromatic ring of Tyr188 a Cl \cdots π electrostatic interaction (Chlorine-aromatic centroid distance: 3.6 Å), which exhibited a face-on

Figure 3. Compounds 1 to 3 inhibited *pks*⁺ *E. coli*-mediated megalocytosis. HeLa cells were infected for 4 hours, then washed and fresh media containing gentamycin was added for 72h. Cells were stained with Giemsa. *pks*, cells infected by *pks*⁺ *E. coli*; Δ *pks*, cells infected by *pks*⁻ *E. coli*; Ninf, non-infected cells; *pks*/cpd1 to 3, cells infected by *pks*⁺ *E. coli* treated by compounds 1 to 3.

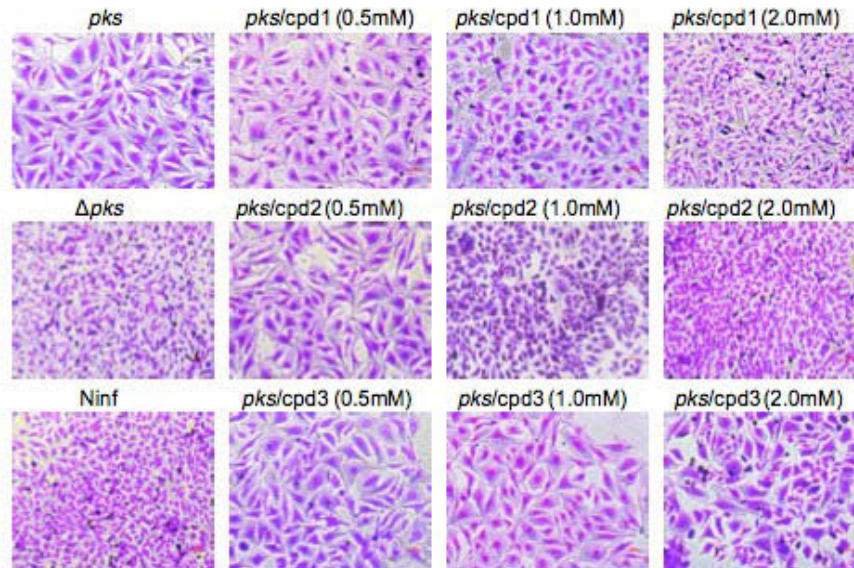
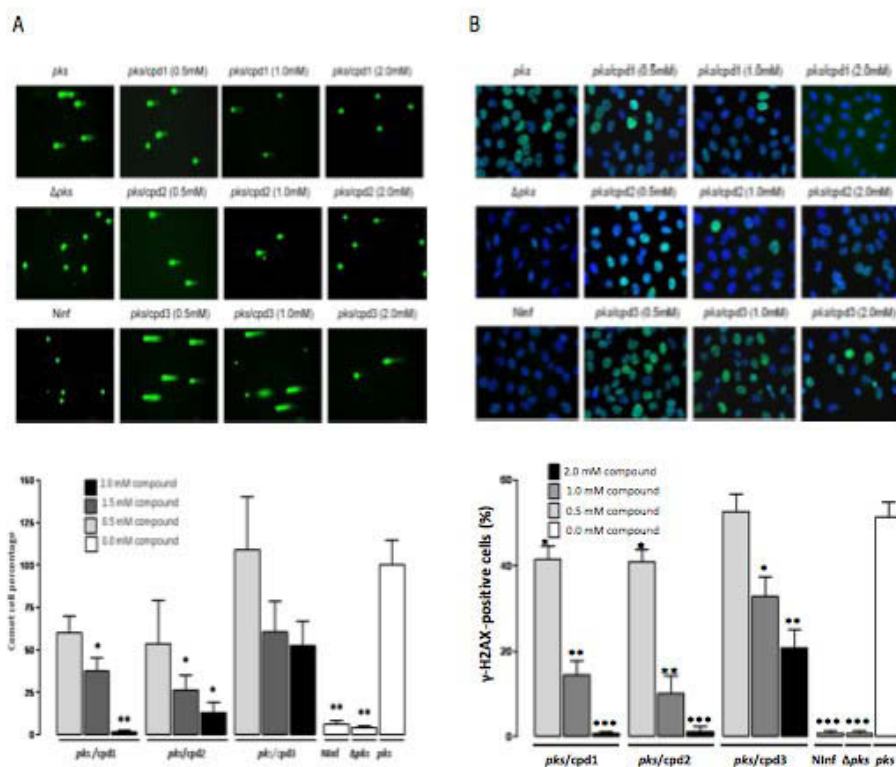


Figure 4. Compounds inhibited DNA damages induced by *pks*⁺ *E. coli*. The DNA damages were revealed by comet assays (A) and γ -H2Ax staining (B). *pks*, cells infected by *pks*⁺ *E. coli*; Δ *pks*, cells infected by *pks*⁻ *E. coli*; Ninf, non-infected cells; *pks*/cpd1 to 3, cells infected by *pks*⁺ *E. coli* treated by the compounds. Bar graphs on the bottom show the means \pm S.E.M of positive cells obtained from 3 independent experiments. * P <0.05; ** P <0.001; *** P <0.0001 vs *pks*⁺ *E. coli*.



geometry as previously reported²⁰. One oxygen atom of the sulfonamide was positioned in close contact with the C and N atoms of the Gly328-Gly329 peptide bond (3.1 Å and 3.6 Å respectively). Because of the π electron resonance cloud of peptide bonds, this close interaction resembled a Lone pair $\cdots\pi$ interaction²¹⁻²³.

In vitro* inhibition of the megalocytosis induced by *pks+* *E. coli

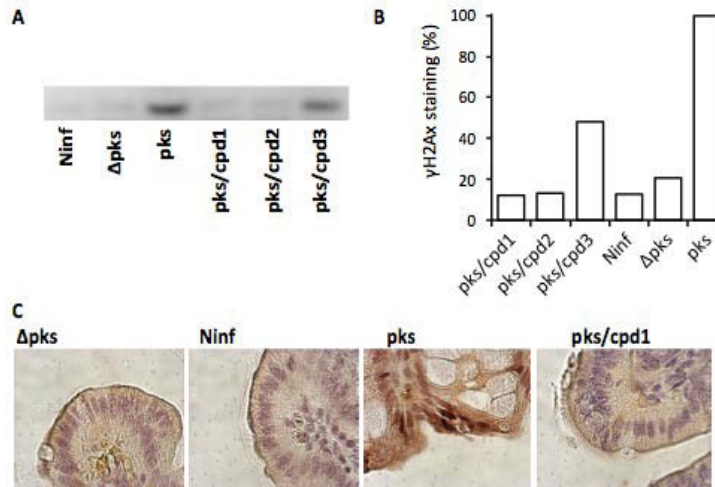
ClbP is required for megalocytosis mediated *pks*-harbouring *E. coli*^{12,13}. To investigate the efficiency of ClbP inhibition, epithelial cells were infected with *E. coli* DH10B hosting a bacterial artificial chromosome bearing the genomic island *pks* (*pks+* *E. coli*) pre-treated or not with increasing concentrations of compounds. *pks+* *E. coli* induced megalocytosis but not *E. coli* hosting a genomic *pks* island defective for ClbP (*pks-* *E. coli*) as expected (figure 3). Importantly, the compounds inhibited in a dose-dependent manner the megalocytosis induced *pks+* *E. coli*, which was abolished with the compounds 1 and 2 at concentration of 2 mM (figure 3). This inhibition was independent of an antibiotic effect since the compounds did not alter the bacterial growth at the concentration used (supplemental figure 2). In addition, the incubation of the compounds at the highest dose (2 mM) with epithelial cells did not affect cellular viability (supplemental figure 3). The compounds prevent therefore megalocytosis mediated by *pks+* *E. coli* through the inhibition of ClbP without evidence of toxicity against bacteria and eukaryotic cells.

In vitro* and *in vivo* inhibition of the DNA damage induced by *pks+* *E. coli

We next investigated the ability of compounds 1 to 3 to prevent the DNA damage induced by *pks+* *E. coli*. Epithelial cells were infected for 3 hours with *pks+* *E. coli* treated or not with increasing concentration of compounds (0.5 to 2mM). As negative controls, epithelial cells were infected or not with *pks-* *E. coli*. DNA damage was first assessed by comet assays (figure 4A). As expected, the percentage of comets was 20-fold higher in cells infected by *pks+* *E. coli* than in non-infected cells or cells infected by *pks-* *E. coli* ($P<0.001$). DNA damage was significantly decreased when *pks+* *E. coli* were treated with 2 mM of compounds 1 and 2 ($p<0.001$ and

Figure 5. The compounds inhibited DNA-damages induced by *pks*⁺ *E. coli* in a mouse colon loop model. A, Western blot analysis of γ H2AX in colonocytes 6 h after inoculation of colon loops with *pks*⁻ *E. coli*, *pks*⁺ *E. coli* treated or not with the compounds 1 to 3 (2mM). B, Relative band intensity of γ H2AX immunostaining *versus* protein yield. C, Paraffin tissue sections were stained for γ H2AX (brown) and counterstained with hematoxylin. *pks*⁻cpd 1 to 3, colon loops infected by *pks*⁺ *E. coli* treated with compounds 1 to 3 (2mM); Ninf, non-infected colon loops; Δ *pks*, colon loops infected by *pks*⁻ *E. coli*; *pks*, colon loops infected by *pks*⁺ *E. coli*.

(fold change)



$p=0.001$ respectively). Furthermore, as observed with the inhibition of megalocytosis, the compounds exhibited a dose response effect.

We next assessed DNA damage by immunostaining the S139-phosphorylation of histone H2Ax (γ -H2Ax) which is a canonical and sensitive marker of DNA double-strand breaks²⁴. As observed in figure 4B, γ -H2AX staining induced by *pks+* *E. coli* was significantly decreased in a dose dependent manner by all compounds. A concentration of 2 mM of compounds 1 and 2 restored a background level of γ -H2AX immunostaining, which was not significantly different to that of non-infected cells or cells infected with *pks-* *E. coli*. As observed with the comet assay, compound 3 was less effective since it did not abolish the DNA damage induced by *pks+* *E. coli*.

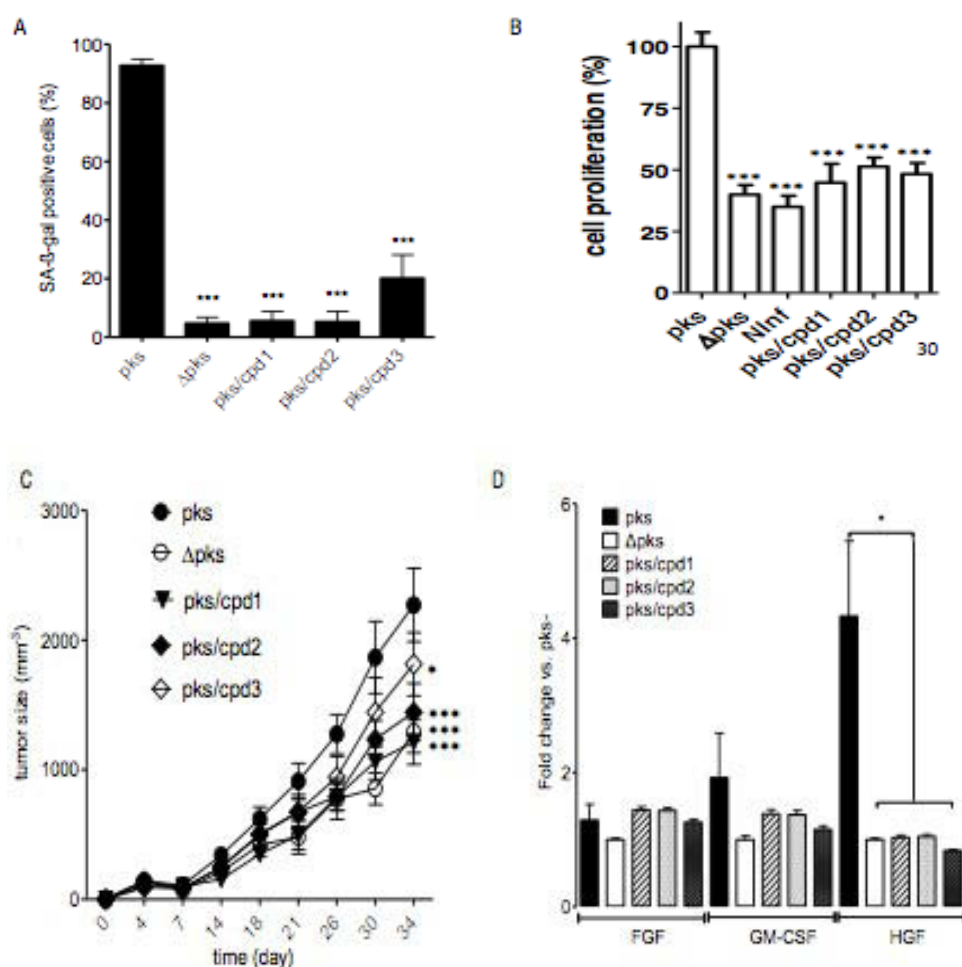
To substantiate these results, we investigated the ability of the compounds to prevent DNA damage induced by *pks+* *E. coli* in a mouse colon loop model. Colon loops were infected for 6 hours with *pks+* *E. coli* treated or not with the compounds 1 to 3 (2mM) and γ -H2Ax was assessed in colonocytes by Western blotting and immunohistology. Non-infected loops and loops inoculated by *pks-* *E. coli* were used as negative controls. Western blot and histologic analyses (Figure 5) revealed an increase in γ -H2Ax level in intestinal loops infected with *pks+* *E. coli*. Interestingly, the three compounds prevented this increase.

Taken together, our results indicate that the compounds were able to block the DNA damage induced by *pks+* *E. coli* *in vitro* and *in vivo*.

In vitro* and *in vivo* inhibition of cell proliferation and tumourigenesis induced by *pks+* *E. coli

Colibactin-producing *E. coli* induce the emergence of senescent cells, which promote tumour growth by the secretion of HGF and this growth factor is upregulated in human tumours colonized by *pks+* *E. coli* in contrast to tumour colonizing by *pks-* *E. coli*²⁵. We therefore investigated the ability of ClbP inhibitors to block cellular senescence mediated by *pks+* *E. coli*. Senescence-associated β -galactosidase (SA- β -gal) activity was assessed in intestine epithelial

Figure 6. ClbP inhibitors prevented cellular senescence, growth factor production and cell proliferation mediated by *pks+* *E. coli*. A, HCT116 cells were infected with *pks+* *E. coli* (*pks*), *pks-* *E. coli* (Δ *pks*) or *pks+* *E. coli* treated with compounds (*pks*/cpd) 1 to 3 (2mM) and percentage of senescence associated β -gal positive cells were determined at 5-days post-infection (Values represent means \pm S.E.M; n = 3 independent experiments in duplicate; *** P <0.001 vs *pks*). B, *In vitro* HCT116 cell proliferation induced by serum-free medium conditioned with HCT116 cells non infected (NInf), infected by *pks+* *E. coli* (*pks*), *pks-* *E. coli* (Δ *pks*) or *pks+* *E. coli* treated by compounds (*pks*/cpd) 1 to 3 (2 mM). The proliferation, assessed by MTT assay, is presented as percentage compared to medium conditioned with cells incubated with *pks+* *E. coli*. (Values represent means \pm S.E.M; n = 3 independent experiments in duplicate; *** P <0.001 vs. *pks*). C, *In vivo* tumour growth stimulated by HCT116 cells pre-incubated with *pks+* *E. coli* (*pks*), *pks-* *E. coli* (Δ *pks*) or *pks+* *E. coli* treated with compounds (*pks*/cpd) 1 to 3 (2mM). (Tumour size was measured as described in Materials and Methods; values represent means \pm S.E.M; n = 5 mice/group; * P <0.05 vs. Δ *pks*, *** P <0.001 vs *pks*)D, Growth factor production in tumours stimulated by HCT116 cells pre-incubated with *pks+* *E. coli* (*pks*), *pks-* *E. coli* (Δ *pks*) or *pks+* *E. coli* treated with compounds (*pks*/cpd) 1 to 3 (2mM). (Values represent means \pm S.E.M; n = 5 mice/group; * P <0.05 vs. *pks*).



cells (HCT116) not infected, infected by *pks+* *E. coli* or infected by *pks+* *E. coli* treated with the compounds 1 to 3. *pks+* *E. coli* increased SA- β -Gal activity with a marked increase in the number of SA- β -Gal positive cells. In not infected cells or cells infected by *pks-* *E. coli* or *pks+* *E. coli* treated with the compounds 1 to 3, a similar significant decrease in the number of senescent cells was observed compared to cells infected by *pks+* *E. coli* (Figure 6A).

The cell medium conditioned with cells infected by *pks+* *E. coli* increased cell proliferation compared to medium conditioned with non-infected cells or cells infected by *pks-* *E. coli* as previously reported²⁵. Interestingly, treatment of *pks+* *E. coli* with compounds 1 to 3 significantly inhibited the cellular proliferation mediated by these bacteria *in vitro* (Figure 6B). Intestinal epithelial cells infected by *pks+* *E. coli* stimulated the growth of co-transplanted cells in immunodeficient mice in contrast to cells infected by *pks-* *E. coli* (Figure 6D). Interestingly, a significant decrease in the tumour growth was observed from 26 and 34-day post-transplantation ($p < 0.05$) in presence of treatments by compounds 1/2 or 3 respectively (Figure 6D). The inhibition of the tumour growth was correlated with the K_i values (Table 1) and the significance of the growth decrease observed with compounds 1 and 2 rose between 26- and 30-day post-infection ($p < 0.001$). Accordingly, HGF RNA transcript production in the tumours was significantly decreased by the treatment of *pks+* *bacteria* with ClbP inhibitors 1 to 3. Overall, our results show that targeting ClbP is an interesting strategy in inhibiting *pks+*-induced tumourigenesis.

DISCUSSION

Non-ribosomal peptides (NRP) and NRP-polyketide hybrids (NRP-PK) are highly bioactive molecules, that are frequently toxic and interesting in drug development²⁶. Their biosynthesis occurs on multimodular enzymatic assembly lines called megasynthases²⁷ and the resulting compounds can be cleaved and matured by NRP-associated serine peptidases^{13,14}. In this work, the NRP-associated peptidase ClbP was targeted to inhibit the prodrug activation of colibactin, a NRP-PK-type genotoxin produced by *pks*-encoding enzymes in *E. coli* and other *Enterobacteria*. We identified three boronic acid-based compounds, which exhibit nanomolar inhibition constants against ClbP and ligand efficiency values of ~ -0.9 Kcal/mol, values compatible with medicinal drug development¹⁶. In addition, the compounds did not exhibit apparent toxicity against eukaryotic cells. The inhibitors had significant activity against the ortholog ZmaM, but not against the paralog FmtA-like, showing a relation between inhibitor activity and the function of these proteins. Our data demonstrate that these compounds efficiently inhibit ClbP and exhibit specificity against NRP-associated peptidases.

The inhibitors of ClbP were able to prevent the genotoxicity mediated by *pks*+ *E. coli* both *in vitro* and *in vivo*, thereby demonstrating that the NRP-associated peptidases are a valuable target to inhibit genomic islands synthesizing NRP and/or NRP-PK compounds. The activity of these inhibitors was dose dependent and varied according to the K_i values. The NRP-associated peptidases are proposed as a common mechanism of cleaving and prodrug activation of numerous NRP- and NRP-PK compounds. The inhibitors can therefore be useful effectors, allowing the synthesis of alternative NRP- and NRP-PK compounds or preventing the synthesis of highly bioactive molecules. Given the broad-spectrum activity of NRP- and NRP-PK (antibiotic, antifungal, genotoxic, etc)^{26,28,29}, this latter application might be useful in numerous domains such as the control of microbiote composition or screening for NRP and NRP-PK production.

The inhibition of the *pks* island might also have a direct application in the field of health.

E. coli is primarily a commensal inhabitant of the mammalian colon and the *pks* island is hosted by commensal *E. coli* strains belonging to the B2 phylogroup, which are commonly found and persistent in human microbiota^{12,30-34}. Recently studies show that *pks*-harbouring bacteria are overrepresented in human CRC specimens and favour tumour progression^{11,25}. Likewise, the *pks* island is also observed in 72% of *E. coli* strains causing prostatitis, which is associated with significantly increased risk for prostate cancer³⁵. The DNA damage induced by *pks*+ *E. coli* strains are estimated to be similar to that of a 0.5-Gy γ -ray whole-body irradiation and triggers mammalian cell transformation, gene mutations and chromosomal instability³⁶, which are central events in tumourigenesis^{37,38}. *pks*+ *E. coli* strains therefore represent a potential etiologic component of colorectal and prostate cancers, respectively the third most frequently diagnosed malignancy in both sexes and the most frequent cancer in males³⁹. In the present study, we report that inhibition of ClbP activity by boronic acid-based compounds prevents the genotoxic and tumourigenic activity induced by *pks* island both *in vitro* and *in vivo*; suggesting that they may be useful tools to prevent and/or control the deleterious actions of *pks*+ *E. coli*. Of note, the compounds have no apparent antibiotic activity. Their target is a virulence factor and not a component required for bacterial survival in contrast to antibiotics. This strategy, which rarely used in infectious disease, allows to avoid the selection of bacteria resistant to the treatment, as observed with antibiotics. This point is major issue in chronic diseases such as CRC.

Interestingly, the two best inhibitors (compounds 1 and 2) harbour a boronic acid function linked to a clearly different moiety, suggesting that two distinct families of inhibitors may be derived from these compounds for further drug improvement. Although exhibiting different chemical structures, these two inhibitors were accommodated in the same area of the ClbP binding site, towards the N-terminal extremity of the β 11 strand between the reactive serine S64 and residues Tyr188 and Gly328. The N-terminal extremity of the β 11 strand forms a broad groove accessible to solvent and mutagenesis experiments have shown that it contains key residues for the activity of ClbP in contrast to the C-terminal extremity^{13,15}. The N-terminal

extremity of the β 11 strand may therefore be the major site for colibactin accommodation.

The binding of the inhibitors in ClbP active site bring into play hydrogen bonds and van der Waals contacts with residues 95, 186, 327-328, 330. Non-canonical interactions are known to play an important role in protein stability and ligand binding. The binding of inhibitors in ClbP active site revealed an atypical high density of such interactions, which may be key determinants of their activity. Interestingly, residue Tyr186 established CH \cdots O and Cl \cdots π interactions with compounds 1 and 2 respectively, and the peptide bond Gly328-Gly329 established π -stacking-like and Lone pair \cdots π interaction with inhibitors 1 and 2 respectively. The type of interaction established by the inhibitors differs therefore from one molecule to the other according to their difference in chemical structure. However, the same structural elements of ClbP active site are involved in the interactions and the inhibition is therefore based on the same logic. Consequently, the structures show different ways to tiger the same goal and an interesting inventory of unusual interactions, which can be built into the designs of other inhibitors.

In conclusion, we have found boronic acid-based inhibitors of ClbP, that are able to suppress DNA damage and the proliferation/tumourigenesis induced by pks-harboured bacteria. This family of inhibitors might be useful in decreasing the production of genotoxic compounds synthesized by *pks* island and subsequently in reducing the carcinogenesis mediated by these bacteria.

MATERIAL AND METHODS

Strains, plasmids and compounds. For *in vitro* and *in vivo* infections, we used the *E. coli* strain DH10B hosting the pBeloBAC 11 bearing the *pks* island (pBACpks; pks+ *E. coli*), or hosting the pBACpks deleted of the ClbP enzyme (pBACpks Δ ClbP, pks- *E. coli*)¹² and 5 pks+ *E. coli* strains isolated from colorectal tumours. The *pks*-harbouring *E. coli* previously reported and its isogenic mutant was used for *ex vivo* experiments³⁶. The enzymatic domain of ClbP, was produced in the *E. coli* BL21(DE3) from the pET9a derivative plasmid pClbPpep¹³. ZmaM and FmtA-like enzymatic domains were produced likewise after cloning in the pETM13 plasmid (Stratagene, La Jolla, CA, USA). *E. coli* strains were routinely grown at 37°C in Luria Bertani broth (LB) or on LB agar plates. Compound 1 (3-aminophenyl boronic acid) was purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France). Compound 2 was obtained by a general procedure described elsewhere⁴⁰, through sulfonylation of pinacol aminomethaneboronate with commercially available chloromethanesulfonyl chloride (89% yield). Mp 59-61 °C. ¹H-NMR (400 MHz, CDCl₃): δ 1.28 (12H, s, pinacol protons), 2.87 (2H, d, *J* 4.6, BCH₂), 4.53 (2H, s, ClCH₂), 4.72 (1H, br, NH). ¹³C-NMR (100 MHz, CDCl₃): δ 24.8, 29.8 (br, CB), 53.9, 85.0. EI-MS, *m/z* (%): 256 (1) [M⁺], 254 (4), 213 (5), 211 (13), 191 (6), 170 (8), 156 (20), 127 (6), 106 (6), 104 (100), 103 (25), 85 (42), 74 (21), 60 (21), 59 (62), 55 (14), 49 (13), 43 (16), 41 (21). Compounds 3 and 4 were synthesized as previously described^{40,41}.

Protein production and purification. Proteins were expressed in *E. coli* BL21(DE3) cultured in terrific-broth II medium (Euromedex, Souffelweyersheim, France) supplemented with sorbitol 400mM (Sigma-Aldrich), betaine 2.5 mM (Sigma-Aldrich), and kanamycin 50 μ g/ml (Sigma-Aldrich). The expression of proteins was induced with 0.2 mM isopropyl- β -D-thiogalactopyranoside (Euromedex). The purification of ClbP was performed as previously described¹³. The purification of FmtA-like and ZmaM was carried out by ion-exchange chromatography into a HiTrapTM Q SepharoseTM High Performance column (Healthcare

Europe, Velizy-villacoublay, France) equilibrated with 20 mM Tris-HCl pH 7.5 and eluted with a linear NaCl gradient (0 to 500 mM). The elution peak was then purified by ion-exchange chromatography into a HiTrap™ Q Sepharose™ High Performance column (Healthcare Europe) equilibrated with 50 mM MES pH 6.0 and eluted with a linear NaCl gradient (0 to 500 mM). The elution peaks were extensively dialyzed against NaCl 50 mM Tris-HCl 5 mM pH 7.0 and concentrated by ultrafiltration to a final concentration of 10 mg/ml. The purified enzymes were more than 95% homogeneous as determined by Coomassie blue staining after SDS-PAGE (Laemmli 1970).

Enzyme assays. The enzymatic activity of ClbP and its homologous ZmaM was detected by the hydrolysis of 4 mM p-nitrophenylbutyrate (Sigma-Aldrich) at 405 nm in a 50 mM NaCl 5 mM Tris-HCl buffer pH 7.0 at 25°C. The enzymatic activity of FmtA-like was detected by monitoring the formation of p-nitroaniline from 5 mM L-Ala-p-nitroanilide (L-Ala-pNA, Bachem) at 405 nm in the same buffer. The enzymatic reactions were performed using 13 μM enzyme. K_m , k_{cat} and K_i were determined in at least a triplicate repeat analysis (s.d. ≤15%) with ENCORA software (Slats et al. 1999). The kinetic parameters of ClbP, ZmaM and Fmta-like were 0.01 min⁻¹, 0.18 min⁻¹ and 12.70 min⁻¹ for k_{cat} , and 1.96 μM, 34.0 μM and 1003 μM for K_m , respectively.

Crystallization and structure determination. ClbP crystals were grown in hanging drops over a solution of 0.8 M monosodium dipotassium phosphate buffer (pH 7.0) as previously reported¹³. The crystals were soaked in the crystallization buffer supplemented with 50 mM boronic acid compounds and cryo-protected with 30% sucrose before flashcooling in liquid nitrogen. Data were collected using a Q315r ADSC-CCD detector on ESRF beamline 23-1 at the European Synchrotron Radiation Facility (Grenoble, France). Reflexions were indexed, integrated, and scaled using CCP4 package⁴². The structure was solved by molecular replacement with the

program PHASER⁴³ and ClbP x-ray structure (pdb ID, 3O3V) as a search model. The structure was automatically and manually refined with REFMAC⁴⁴ and COOT⁴⁵ programs, respectively. Cross-validation was used throughout and 5 % of the data were used for the Rfree calculation. The stereochemical quality of the models was monitored with the PROCHECK program⁴⁶. Ramachandran plots were calculated by RAMPAGE⁴⁷. Processing and crystallographic refinement statistics are listed in supplemental table 2. The coordinates and structural factors of ClbP have been deposited in the RCSB Protein Data Bank (PDB IDs: 4E6W and 4E6X).

Molecular docking. A set of 450 boronic acid structures collected from supplier catalogs (Boron Molecular, Tianjin Boron PharmaTech Co. and Sigma-Aldrich) was screened by computational simulation for their ability to bind ClbP active site. The screening scheme comprised flexible molecular docking with Autodock Vina 1.1⁴⁸ and molecular dynamics simulation with Amber 10, followed by MM-PBSA binding free energy estimates⁴⁹. Three-dimensional coordinates of the boronic acid structures were generated with Marvin 5.3.8, 2010, ChemAxon (<http://www.chemaxon.com>), and Autodock Tools⁵⁰ was used to prepare the input files for AutoDock Vina. All bound waters, ligands and cofactors were removed from ClbP crystal structure (PDB ID, 3O3V). ClbP active site was treated as rigid during the docking simulations with Autodock Vina. A carbon atom with a correction of the bond lengths based on boron crystallographic data was used to mimic the boron atom, because no parameters were implemented for boron. This approximation was considered acceptable since the differences between the compounds resulted from changes remote from boron. The best binding mode of each ligand was collected for further investigations. The corresponding coordinates were refined and rescored by molecular dynamics simulations followed by MM-PBSA binding free energy estimates, as previously described with minor modifications⁴⁹. The amber force field atom types and ClbP atomic charges were assigned with the antechamber and parmcheck modules of Amber. The RESP charges of compounds were obtained from *ab initio* calculations using

GAMESS (HF/6-31G* theory level) with RED-III software as recommended^{51,52}. Because of the mechanism of action of boron-based inhibitors, the predicting binding modes exhibiting the boron atom located in the vicinity of the O_v atom of the active serine were collected for MM-PBSA binding free energy calculation and ranked according to ligand efficiency as previously described¹⁶.

In vitro infection assay. *In vitro* assays were performed using HeLa and HTC116 cells as previously described¹². Briefly, bacteria were grown overnight in LB medium in the presence or absence of inhibitors. Cells were then infected for 3 hours at a multiplicity of infection of 500. After 3 hours of infection, cells were washed three times with PBS and incubated in culture medium supplemented with 200 µg/ml gentamicin until analysis.

Comet assay. The comet assay (a single-cell gel electrophoresis assay) was performed at 24 hours post-infection. Infected cells were trypsinized, 5.10⁴ cells/ml were mixed with 0.5% low melting point agarose and spread on microscope slides pre-coated with 1.5% normal melting point agarose. After 45 min gelling at 4°C, the slides were immersed in neutral lysing solution (10 mM Tris, 2.5 M NaCl, 100 mM EDTA, DMSO 10%, Triton 1%) for 2 hours at 4°C. Electrophoresis was carried out at 4 °C for 40 min at 220 mA and 30 V in a migration buffer containing 0.3 M NaOH, 1 mM EDTA. The slides were then submerged in a neutralization buffer (0.4 M Tris-HCl, pH 7.5), dried at room temperature, fixed in 100% ethyl alcohol and stained with SYBR Green solution. DNA integrity was assessed at 400× magnification, using a fluorescence microscope (Zeiss LSM 510). The extent and distribution of DNA damage were assessed by examining at least 50 randomly selected cells in a triplicate repeat analysis. The cells were scored with the CometScore™ (Tritek corp.) image analysis software using the mean comet tail moment (product of tail length and fraction of DNA in the tail).

Cells immunostaining. Cells cultured on cover slips were fixed at 24 hours post-infection in 4% paraformaldehyde, then blocked 1 h at room temperature and incubated over-night at 4°C with anti-phospho(Ser139)-H2Ax antibodies (Ozyme, Saint Quentin Yvelines, France). After several washings with PBS, cells were incubated with FITC-conjugated secondary antibodies (Ozyme). DNA was stained with Hoechst (Invitrogen) and images were acquired with a Zeiss LSM 510 Meta laser scanning confocal microscope.

Cell proliferation and toxicity assays. Conditioned media were obtained by collecting the culture medium without fetal bovine serum from 5 days post-infected cells. Overnight serum-starved cells were stimulated with conditioned media and cellular proliferation was measured with the Cell Proliferation kit I (MTT) (Roche Applied Science, Mannheim, Germany). Cytolysis was investigated by measuring lactate dehydrogenase (LDH) release (LDH kit; Boehringer, Mannheim, Germany).

Senescence-associated β -galactosidase staining: Senescent cells were visualized using the Senescence Cells Histochemical Staining Kit (Sigma-Aldrich) according to the manufacturer's instructions. *In vivo* β -galactosidase staining was performed as previously described ⁷. Briefly, tumours were embedded into OCT (Thermo Fisher), freeze dried and immediately processed. 5 μ m sections were fixed for 1 min with 1% paraformaldehyde solution and stained overnight with X-gal staining solution (Sigma-Aldrich).

Murine colon loop model and tumourigenesis assays. The colon loop assay was performed as previously described ³⁶. Briefly, BALB/cJ mice (4 weeks old, Charles Rivers) were anesthetized, a midline abdominal incision was made, a ligature was placed at the proximal and distal ends of the colon, and 3×10^9 bacteria (300 μ L) were injected into the colonic lumen. The incision was sutured, and animals were allowed to recover. The mice were killed 6 hours later,

and the colons were collected immediately.

Human intestinal cancer xenografts were established in Nude (nu/nu) mice (Charles Rivers) as previously described⁵³ with minor modifications. Nude (nu/nu) mice (6 weeks old) were injected (100µl) s.c. into the dorsal flap with 1.6×10^6 HCT116 cells, alone or in combination with 3.0×10^5 5 day post-infected HCT116 cells. We measured three maximum diameters at x and y axes to derive tumour volume (size) using the equation, length \times (width)² \times 0.52. Animal experimentations were carried out in accordance with European Guidelines for the Care and Use of Animals for Research Purposes.

Immunohistochemistry: 5µm paraffin-embedded tissue sections were deparaffinized in xylene, incubated in 3% hydrogen peroxide in methanol for 30 minutes, rehydrated in ethanol gradient, and then treated with 10 mM sodium citrate buffer (pH 6.0) containing 0.05% Tween 20 at 120°C for 5 minutes in a microwave oven. Sections were blocked with 2% non-fat dry milk and 0.01% Tween 20 in PBS for 1 hour at 37°C, and then incubated with anti-phospho(Ser139)-H2Ax antibody (Ozyme) overnight at 4°C. After washes with PBS containing 0.01% Tween 20, sections were treated with appropriate biotinylated secondary antibodies for 30 minutes at 37°C, and colour development was performed using the Vectastain ABC kit (Vector Laboratories). Sections were then counterstained with hematoxylin, dehydrated, and coverslipped. Images were acquired using a Zeiss Axioskop LSM510 microscope (Carl Zeiss MicroImaging).

Western blot. Total proteins extracted from colonocytes prepared as previously described³⁶ were resolved on polyacrylamide gels and transferred to PVDF membranes. Membranes were blocked and then probed overnight at 4°C with anti-phospho(Ser139)-H2Ax antibody (Ozyme). After several washings, membranes were incubated with the appropriate HRP-conjugated secondary antibodies (Ozyme). Immunoreactive bands were detected using the enhanced chemiluminescence detection system (Amersham Biosciences).

Quantitative real-time PCR (qRT-PCR): Total RNAs isolated using the TRIzol reagent (Invitrogen) from tumours were reversely transcribed using the NCodeTM miRNA first-strand cDNA synthesis kit (Invitrogen) to quantify mature miRNA expression, or using the first-strand cDNA synthesis kit (Euromedex) to quantify mRNA expression according to the manufacturer's instructions. qRT-PCR was performed using SYBR Green qPCR Master Mix (Roche) on a Mastercycler Realplex⁴ (Eppendorf) using specific primers (FGF: 5'-GCC CCA GAA AAC CCG AGC GAG-3' and 5'-GGC ACC GCG TCC GCT AAT CT-3'; GM-CSF: 5'-TGT GGC TGC AGA GCC TGC TG-3' and 5'-CAG GTT CAG GAG ACG CCG GG-3'; HGF: 5'-GGC ATC AAA TGT CAG CCC TGG AGT-3' and 5'-ACC AGG GTC CCC CTT CTT CCC -3'). Fold-induction was calculated using the Ct method: $\Delta\Delta Ct = (Ct_{\text{Target gene}} - Ct_{\text{Housekeeping gene}})_{\text{group1}} - (Ct_{\text{Target gene}} - Ct_{\text{Housekeeping gene}})_{\text{group2}}$, and the final data were derived from $2^{-\Delta\Delta Ct}$.

Acknowledgements

The X-ray experiments were performed on the ID23-1 beamline at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Local Contact at ESRF for providing assistance in using beamline ID23-1. We also thank Marlène Jan and Roland Perroux for her excellent technical assistance. This work was supported by the Ministère de la Recherche et de la Technologie and Institut national de la santé et de la recherche médicale (UMR Inserm U1071), l'Institut National de la Recherche Agronomique (USC-2018), the Ligue contre le cancer and the Centre Hospitalier Régional Universitaire de Clermont-Ferrand, France.

References

1. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
2. Tjalsma, H., Boleij, A., Marchesi, J. R. & Dutilh, B. E. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nature reviews. Microbiology* (2012).doi:10.1038/nrmicro2819
3. Dove, W. F. *et al.* Intestinal neoplasia in the ApcMin mouse: independence from the microbial and natural killer (beige locus) status. *Cancer Res.* **57**, 812–814 (1997).
4. Uronis, J. M. *et al.* Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE* **4**, e6026 (2009).
5. Kaper, J. B., Nataro, J. P. & Mobley, H. L. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* **2**, 123–140 (2004).
6. Swidsinski, A. *et al.* Association between intraepithelial *Escherichia coli* and colorectal cancer. *Gastroenterology* **115**, 281–286 (1998).
7. Martin, H. M. *et al.* Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology* **127**, 80–93 (2004).
8. Nougayrède, J.-P. *et al.* *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848–851 (2006).
9. Cuevas-Ramos, G. *et al.* *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11537–11542 (2010).
10. Dubois, D. *et al.* High prevalence of cyclomodulin- and genotoxin-producing *E. coli* associated with colon tumors. *In submission* (2012).
11. Arthur, J. C. *et al.* Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science* (2012).doi:10.1126/science.1224820
12. Nougayrède, J.-P. *et al.* *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science* **313**, 848–51 (2006).
13. Dubois, D. *et al.* ClbP is a prototype of a peptidase subgroup involved in biosynthesis of nonribosomal peptides. *J Biol Chem* **286**, 35562–35570 (2010).
14. Reimer, D., Pos, K. M., Thines, M., Grun, P. & Bode, H. B. A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat Chem Biol* **7**, 888–890 (2011).
15. Cougnoux, A. *et al.* Analysis of structure-function relationships in the colibactin-maturing enzyme ClbP. *Journal of molecular biology* in press (2012).
16. Hopkins, A. L., Groom, C. R. & Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discov Today* **9**, 430–431 (2004).
17. Cotesta, S. & Stahl, M. The environment of amide groups in protein-ligand complexes: H-bonds and beyond. *J Mol Model* **12**, 436–444 (2006).
18. Malone, J. F., Murray, C. M., Charlton, M. H., Docherty, R. & Lavery, A. J. X-H... π (phenyl) interactions Theoretical and crystallographic observations. *J Chem Soc, Faraday Trans* **93**, 3429–3436 (1997).
19. Bendova, L., Jurecka, P., Hobza, P. & Vondrasek, J. Model of peptide bond-aromatic ring interaction: correlated ab initio quantum chemical study. *J Phys Chem B* **111**, 9975–9979 (2007).
20. Imai, Y. N., Inoue, Y., Nakanishi, I. & Kitaura, K. Cl- π interactions in protein-ligand complexes. *Protein Sci* **17**, 1129–1137 (2008).
21. Egli, M. & Sarkhel, S. Lone pair-aromatic interactions: to stabilize or not to stabilize. *Accounts of chemical research* **40**, 197–205 (2007).
22. Jain, A., Ramanathan, V. & Sankararamkrishnan, R. Lone pair... π interactions between water oxygens and aromatic residues: Quantum chemical studies based on high-resolution protein structures and model compounds. *Protein Sci* **18**, 595–605 (2009).
23. Mooibroek, T. J., Gamez, P. & Reedijk, J. Lone pair- π interactions: a new supramolecular bond? *Cryst Eng Comm* **10**, 1501–1515 (2008).
24. Olive, P. L. & Banath, J. P. The comet assay: a method to measure DNA damage in individual cells. *Nat Protoc* **1**, 23–29 (2006).
25. Cougnoux, A. *et al.* Bacterial genotoxin colibactin promotes tumourigenesis by impairing SUMOylation. *In submission*
26. Li, J. W.-H. & Vederas, J. C. Drug Discovery and Natural Products: End of an Era or an Endless frontier? *Science* **325**, 161–5 (2009).
27. Sattely, E. S., Fischbach, M. a & Walsh, C. T. Total biosynthesis: in vitro reconstitution of polyketide and nonribosomal peptide pathways. *Natural product reports* **25**, 757–93 (2008).
28. Walsh, C. T. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science (New York, N.Y.)* **303**, 1805–10 (2004).
29. Walsh, C. T. Natural insights for chemical biologists. *Nature Chemical Biology* **1**, 122–124 (2005).
30. Escobar-Paramo, P. *et al.* Large-scale population structure of human commensal *Escherichia coli* isolates.

- Appl Environ Microbiol* **70**, 5698–5700 (2004).
31. Johnson, J. R., Johnston, B., Kuskowski, M. A., Nougayrede, J. P. & Oswald, E. Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* pks genomic island. *J Clin Microbiol* **46**, 3906–3911 (2008).
 32. Kurokawa, K. *et al.* Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA research* **14**, 169–81 (2007).
 33. Putze, J. *et al.* Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infect Immun* **77**, 4696–4703 (2009).
 34. Zhang, L., Foxman, B. & Marrs, C. Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol* **40**, 3951–3955 (2002).
 35. De Marzo, A. M. *et al.* Inflammation in prostate carcinogenesis. *Nature reviews. Cancer* **7**, 256–69 (2007).
 36. Cuevas-Ramos, G. *et al.* *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc Natl Acad Sci U S A* **107**, 11537–11542 (2010).
 37. Grady, W. M. & Carethers, J. M. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* **135**, 1079–1099 (2008).
 38. Lengauer, C., Kinzler, K. W. & Vogelstein, B. DNA methylation and genetic instability in colorectal cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 2545–50 (1997).
 39. Siegel, R., Ward, E., Brawley, O. & Jemal, A. Cancer statistics, 2011: the impact of eliminating socioeconomic and racial disparities on premature cancer deaths. *CA Cancer J Clin* **61**, 212–236 (2011).
 40. Eidam, O. *et al.* Design, synthesis, crystal structures, and antimicrobial activity of sulfonamide boronic acids as beta-lactamase inhibitors. *J Med Chem* **53**, 7852–7863 (2010).
 41. Morandi, F. *et al.* Nanomolar inhibitors of AmpC beta-lactamase. *J Am Chem Soc* **125**, 685–695 (2003).
 42. Potterton, L. *et al.* Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2288–2294 (2004).
 43. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40**, 658–674 (2007).
 44. Steiner, R. A., Lebedev, A. A. & Murshudov, G. N. Fisher's information in maximum-likelihood macromolecular crystallographic refinement. *Acta Crystallogr. D Biol. Crystallogr.* **59**, 2114–2124 (2003).
 45. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
 46. Laskowski, R. A., Moss, D. S. & Thornton, J. M. Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**, 1049–1067 (1993).
 47. Lovell, S. C. *et al.* Structure validation by C α geometry: phi,psi and C β deviation. *Proteins* **50**, 437–450 (2003).
 48. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461 (2010).
 49. Rastelli, G., Del Rio, A., Degliesposti, G. & Sgobba, M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. *J Comput Chem* **31**, 797–810 (2010).
 50. Morris, G. M. *et al.* AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **30**, 2785–2791 (2009).
 51. Cornell, W. D., Cieplak, P., Bayly, C. I. & Kollmann, P. A. Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation. *J Am Chem Soc* **115**, 9620–9631 (1993).
 52. Dupradeau, F. Y. *et al.* R.E.D.D.B.: a database for RESP and ESP atomic charges, and force field libraries. *Nucleic Acids Res* **36**, D360–7 (2008).
 53. Krtolica, a, Parrinello, S., Lockett, S., Desprez, P. Y. & Campisi, J. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 12072–7 (2001).

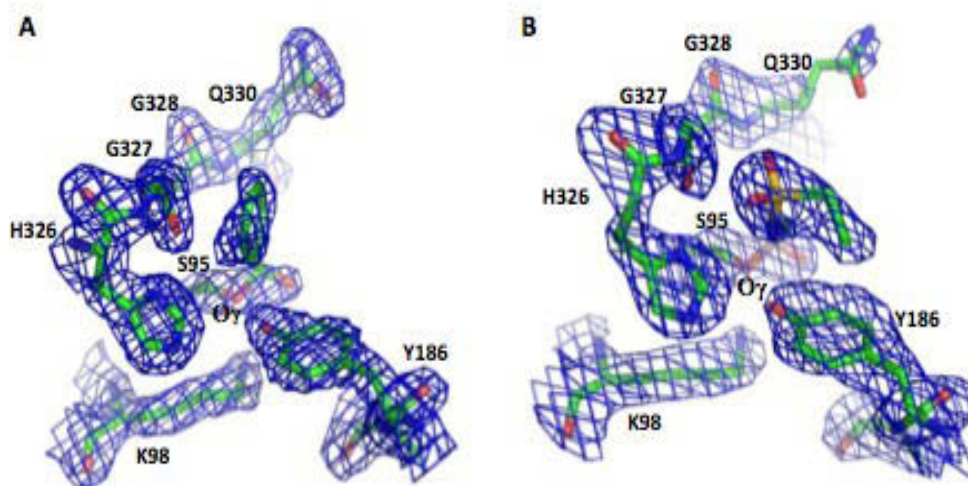
Supplementary data

Supplemental table 1. Data collection and refinement statistics.

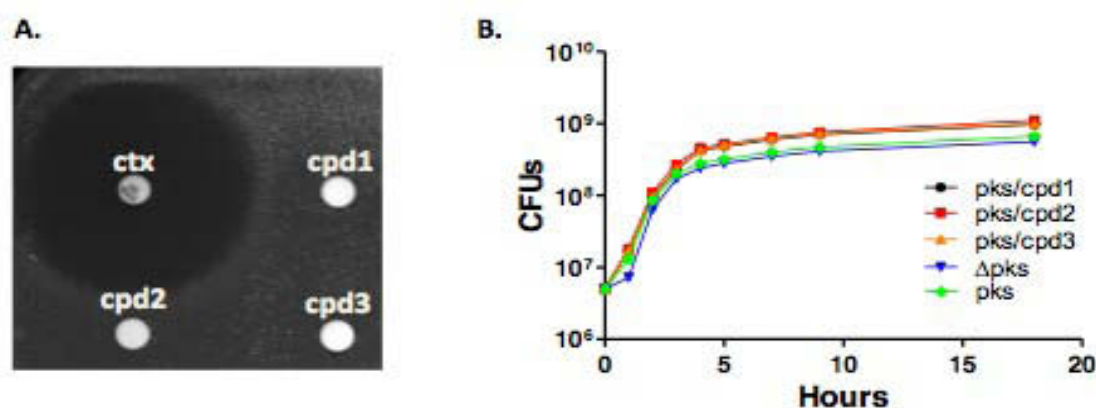
	C1bP / Compound 1	C1bP / Compound 2
Data collection		
Space group	<i>C121</i>	<i>C121</i>
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	104.5, 149.4, 88.2	105.1, 152.5, 86.9
α , β , γ (°)	90.0, 124.2, 90.0	90.0, 123.4, 90.00
Resolution (Å)	52.2-2.19 (2.31-2.19)*	52.6-2.24 (2.36-2.24)*
<i>R</i> _{merge}	0.087 (0.311)*	0.069 (0.290)*
<i>I</i> / σ <i>I</i>	10.4 (3.7)*	10.7 (3.6)*
Completeness (%)	99.7 (99.8)*	79.0 (80.9)*
Redundancy	3.3 (3.2)*	3.0 (2.9)*
Refinement		
Resolution (Å)	52.2 – 2.19	52.6 – 2.24
No. unique reflections	54118	40985
<i>R</i> _{work} / <i>R</i> _{free}	0.194 / 0.238	0.2004 / 0.2492
No. atoms		
Protein	7626	7054
Water	400	172
Ligand	30	20
<i>B</i> -factors		
Protein	27.45	38.66
Water	26.99	32.03
Ligand	23.11	48.19
R.m.s. deviations		
Bond lengths (Å)	0.013	0.014
Bond angles (°)	1.510	1.713
Ramachandran plot		
% residues in favored region	97.6	98.1
% residues in allowed region	2.4	1.9

* , values for the highest range of resolution

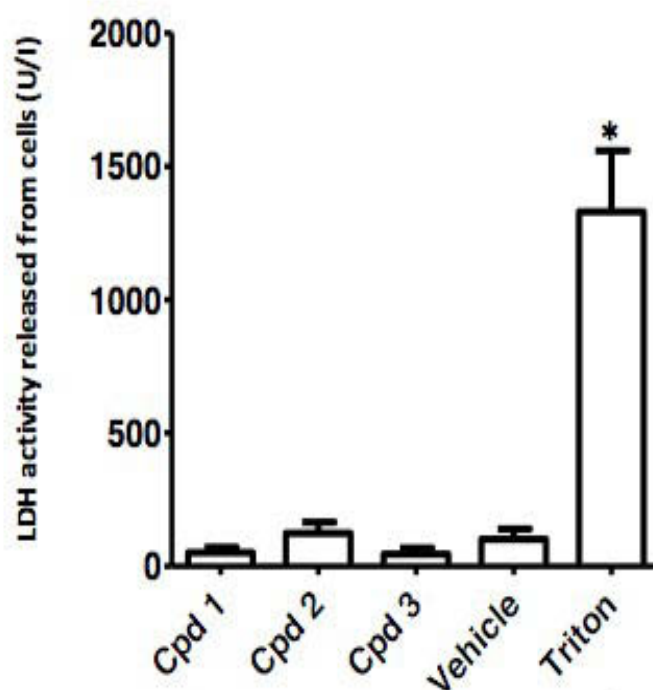
Supplemental Figure 1. Electron density at the vicinity of the serine active residue (S95) for C1bP in complex with compounds 1 and 2. The position of the inhibitors (A, compound 1; B, compound 2) in the active site is unambiguously identified and the electron density connected the O γ of the catalytic Ser95 to the boron atoms of the inhibitors as expected.



Supplemental figure 2. Inhibitors did not inhibit the growth of *pks*⁺ *E. coli*. A. Antibiogram of *pks*⁺ *E. coli* for the antibiotic cefotaxime (ctx, 30 μ g) and the compounds 1 to 3 (cpd1 to 3, 60 μ g). In contrast to cefotaxime, the compounds 1 to 3 did not inhibit the bacterial culture. B. Growth curves of *pks*⁻ *E. coli* (Δ *pks*) and *pks*⁺ *E. coli* (*pks*) in absence and in presence of 2 mM compounds (*pks*/cpd1 to 3) at 37°C in Müller-Hinton broth. The compounds did not inhibit the bacterial growth.



Supplemental figure 3. Inhibitors did not exhibit cellular toxicity. LDH release by HeLa cells incubated 24 hours with 2 mM of compounds (cpd) 1 to 3, the vehicle or 0.01% triton (positive control). Values represent means \pm S.E.M (n = 3 independent experiments in duplicate; **P*<0.05 vs vehicle).



III- Conclusions et perspectives

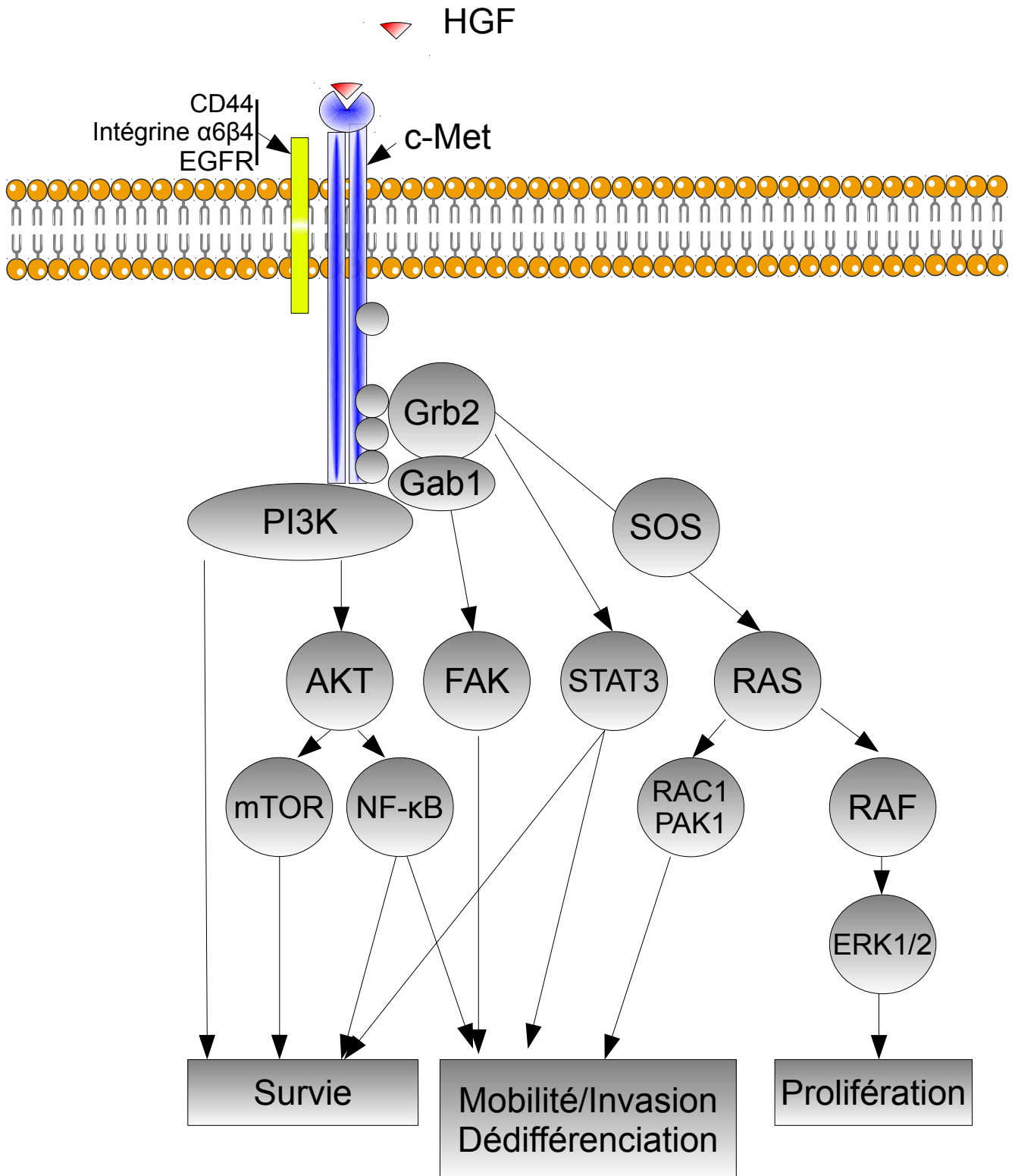


Figure 33 : la voie de signalisation HGF/c-Met. La fixation de HGF à son récepteur entraîne son autophosphorylation et l'activation de protéines adaptatrices. Ces dernières activent de nombreuses voies de signalisation impliquées dans la survie, la prolifération, la mobilité, l'invasion et la dédifférenciation des cellules. AKT : Protein Kinase B ; EGFR : epithelial growth factor receptor ; FAK : focal adhesion kinase ; Gab1 : Grb2-associated adaptor protein 1 ; Grb2 : growth factor receptor-bound protein 2 ; mTOR : mammalian Target Of Rapamycin ; NF-κB : nuclear factor kapa B ; PAK : p21-activated kinase ; PI3K : phosphoinositide-3-kinase ; PTEN : Phosphatase and tensin homolog ; RAS : rat sarcoma oncogene homolog ; RAC1 : Ras-related C3 botulinum toxin substrate 1 ; SOS : son of sevenless ; STAT3 : signal transducer and activator of transcription 3.

1. *E. coli* producteur de colibactine, sénescence prématurée et carcinogénèse colorectale

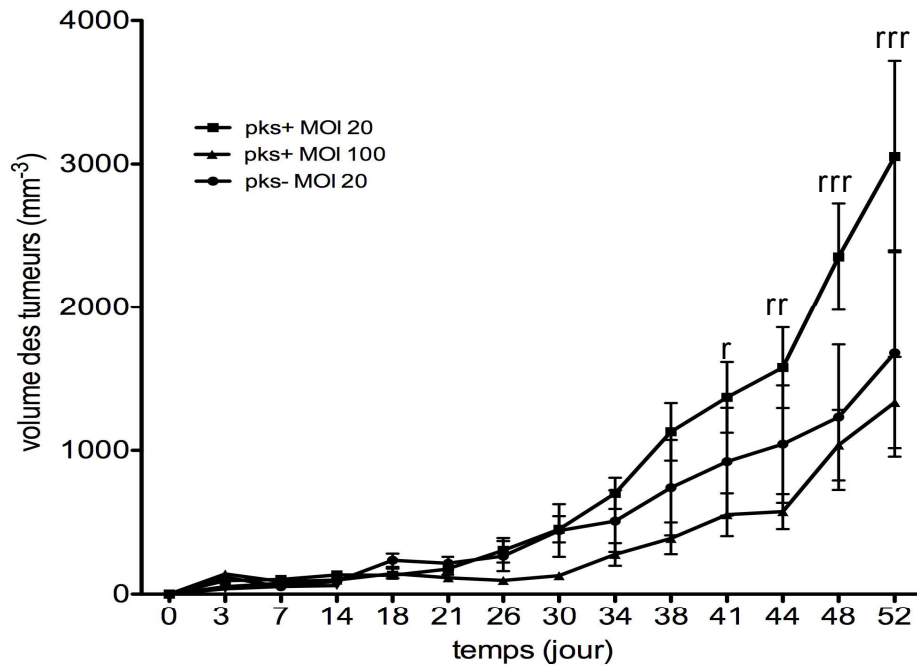
Le groupe du Professeur E. Oswald a montré que *E. coli* pks+ induit aussi bien *in vitro* qu'*in vivo* des cassures double brin de l'ADN dans les cellules eucaryotes infectées [14,26]. Ces lésions sont partiellement réparées *in vitro* et il en résulte une augmentation du taux de mutations ainsi qu'une instabilité chromosomique [14]. Récemment, des travaux menés par l'équipe du Professeur C. Jobin, ont montré qu'une souche murine porteuse de l'îlot *pks* induisait des cassures double brin de l'ADN des cellules hôte et présentait une forte activité carcinogène indépendante de l'inflammation dans un modèle murin IL-10^{-/-} [13]. L'activité carcinogène de *E. coli* pks+ semble donc liée à l'activité génotoxique et mutagène de la colibactine qui pourraient affecter les gènes suppresseurs de tumeurs et/ou les oncogènes.

Nos résultats obtenus *in vitro*, *in vivo* et sur des pièces opératoires issues de patients atteints de cancer colorectal montrent que l'activité pro-tumorale de *E. coli* pks+ peut aussi être indirecte. Les principaux résultats expérimentaux obtenus dans ce sens peuvent être résumés en deux points :

- Les *Escherichia coli* producteurs de colibactine induisent la sénescence prématurée des lignées cellulaires épithéliales intestinales humaines,
- La sénescence cellulaire induite par les *E. coli* producteurs de colibactine s'accompagne d'un phénotype sécrétoire marqué par la sécrétion de facteurs de croissance, principalement le HGF.

Le HGF entraîne la prolifération de cellules épithéliales intestinales non infectées *in vitro* et dans un modèle de xénogreffes murin. Cette activité repose très probablement sur l'action autocrine et paracrine du HGF sur les cellules exprimant le récepteur tyrosine kinase c-Met. La fixation du HGF à son récepteur conduit à une cascade de signalisations (**Figure 33**), qui active les voies PI3K/AKT et RAS/MAPK. La voie HGF/c-Met favorise ainsi la survie, la prolifération, et la migration cellulaire, ainsi que la dédifférenciation cellulaire [394,395].

La sénescence cellulaire est un processus connu pour avoir des fonctions antagonistes pro- et anti-tumorales [396]. Le mode d'action de composés anti-cancéreux comme la bléomycine repose pour partie sur l'induction massive de la sénescence cellulaire [397,398].

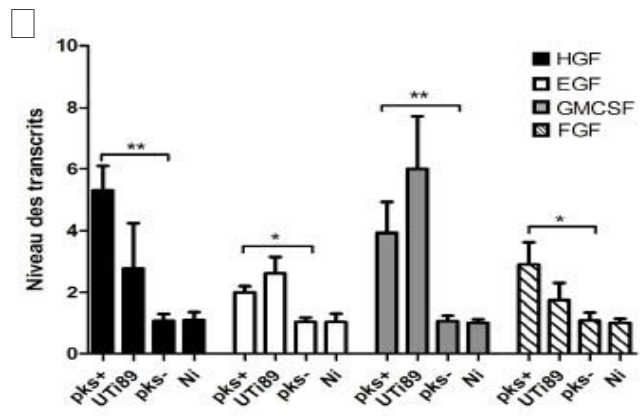
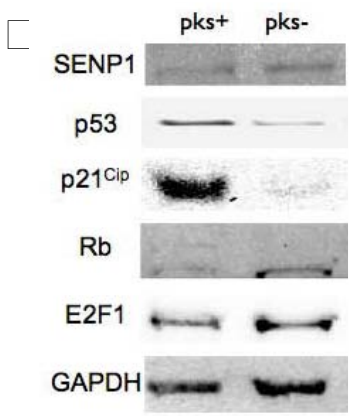
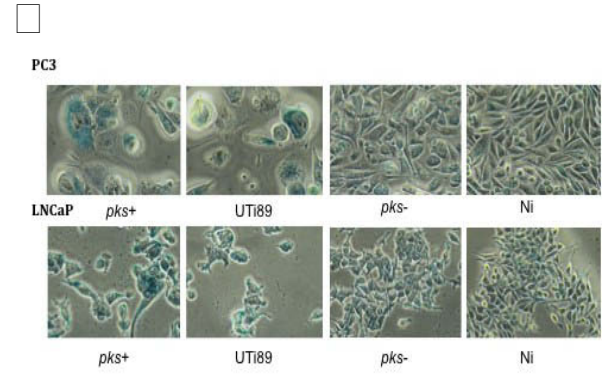
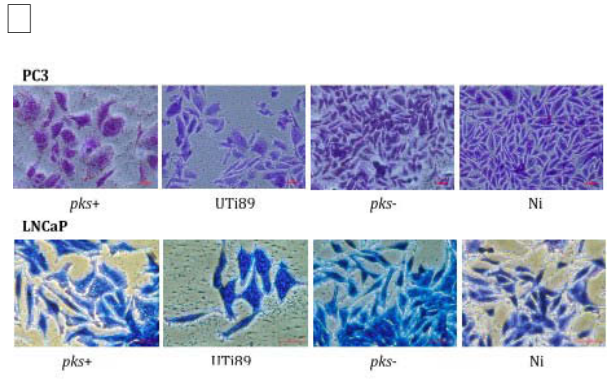
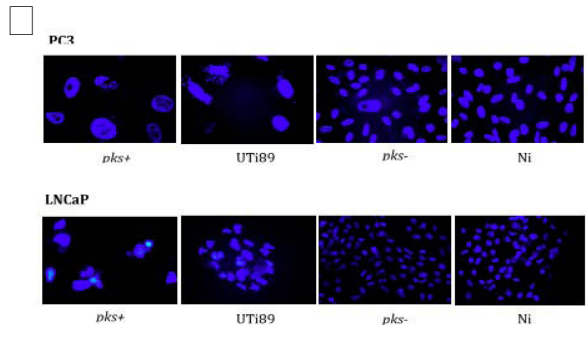
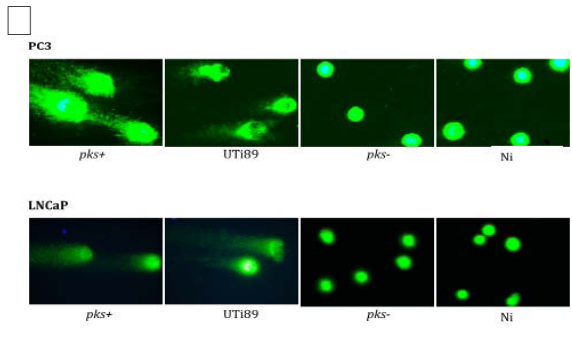


i Téj u . sTj c
 n²
 t uTé d n i é 1 Cl i é Téé 1 Cl Téé
 o c
 i éé 2 c
 6. a 6 r àééaerr àééTerr àéééT
 HC

Ces molécules ralentissent la croissance tumorale en induisant un blocage du cycle cellulaire définitif des cellules tumorales [172,396,399]. Il est récemment apparu que des fibroblastes sénescents peuvent au contraire favoriser la croissance tumorale [172,400]. Nos données suggèrent que ce phénomène peut être étendu aux cellules malignes infectées par des bactéries productrices de colibactine.

Nous proposons que l'activité pro-tumorale ou l'activité anti-tumorale de la sénescence pourrait prendre le pas sur l'autre activité selon les circonstances. Le ratio entre le nombre de bactéries *pks+* et le nombre de cellules cancéreuses ainsi que l'accessibilité des cellules au sein de la tumeur pourraient être des éléments modulant l'effet de la sénescence induite par les bactéries *pks+*. Si ce ratio est élevé et les cellules cancéreuses facilement accessibles, les bactéries *pks+* pourraient avoir une activité anti-tumorale en induisant la sénescence dans la majorité des cellules grâce au blocage du cycle cellulaire. A l'inverse, si les bactéries n'induisent la sénescence que dans une fraction des cellules cancéreuses, ces dernières pourraient avoir une activité trophique sur les cellules non infectées et ainsi avoir une activité pro-tumorale grâce au phénotype sécrétoire associé à la sénescence. Des résultats préliminaires obtenus dans le modèle murin de xénogreffe avec une dose infectieuse forte et faible confortent ces hypothèses. En effet, l'infection des xénogreffes à l'aide d'une forte dose bactérienne (~100 bactéries/cellule) tend à diminuer la vitesse de croissance des tumeurs comparativement à des xénogreffes non infectées (**Figure 34**). En revanche, comme présenté plus haut, une faible dose infectieuse (~20 bactéries/cellule) accélère la croissance tumorale. Pour confirmer ces résultats, nous souhaitons explorer la relation dose-réponse de *E. coli pks+* sur la croissance tumorale.

Par ailleurs, Krieger et collaborateurs ont rapporté que 72% des *E. coli* isolés lors de prostatites infectieuses étaient porteurs de l'îlot *pks* [374]. Une stimulation de la prolifération cellulaire associée à une dédifférenciation cellulaire ont été observées dans un modèle de prostatite aiguë à *E. coli* chez le rat [401]. Ces données, associées à la large distribution de l'îlot *pks* chez les *E. coli* uropathogènes (36,8-47%) [349,370], suggèrent une implication possible de ces bactéries dans la progression du cancer de la prostate. En parallèle des travaux réalisés sur les cellules d'origine intestinale, nous avons exploré l'action des *E. coli* producteurs de colibactine sur les lignées cellulaires d'origine prostatique LNCaP et PC3. Ces deux populations cellulaires ont été infectées par des *E. coli pks+* (DH10B *pks+* et la souche clinique UT189) et *pks-* (DH10b *pks-*). Les cellules infectées par les bactéries productrices de colibactine



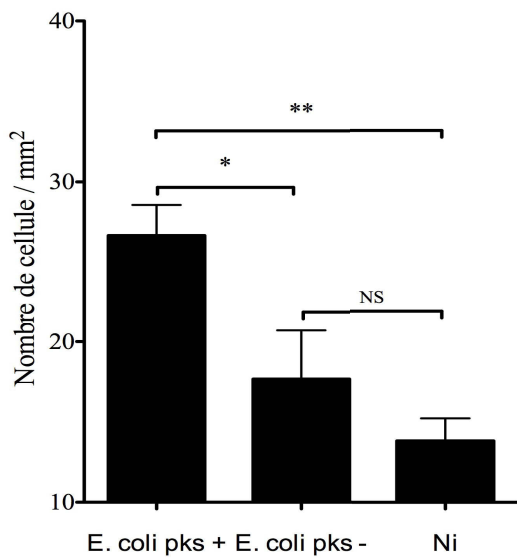
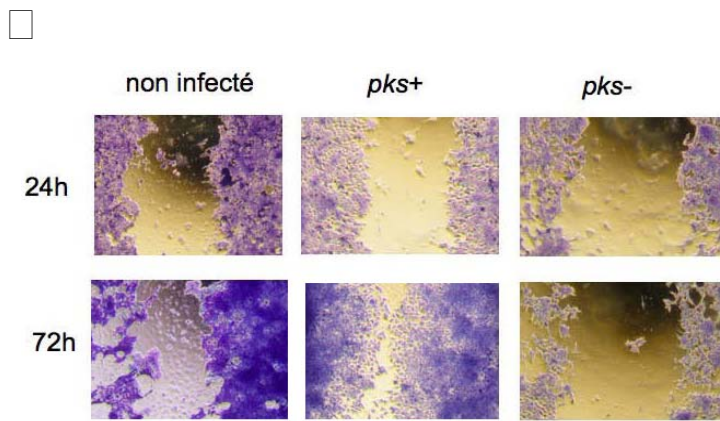
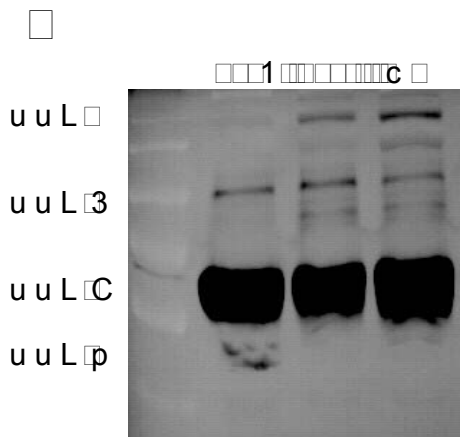
b' Df mè
 9a 2 x r 7b p 2
 8 r 9 . - A v
 o 2 2 l 8 i)
 2 8 vT bU
 (eb, U 8 Nx e 8
 p A(2
 i c, c, cu An, 8 v o AT
 x Df mè gyd dv tgg yd dU a+ 9 . -4) b t, c, 4
 , é t c . 9i 4 T t c u An, 4
 c t(4) tinea. U4nDu + m U

deviennent géantes et présentent une activité β -galactosidase lysosomiale à pH6. Une accumulation des protéines p53 et p21^{Cip} ainsi qu'une diminution de la peptidase SENP1 ont été détectées par Western-blot. Ces cinq éléments (**Figure 35**) concordent avec la mise en place d'une sénescence cellulaire. De plus, un phénotype sécrétoire capable de stimuler la prolifération de cellules non infectées a été observé (**Figure 35**). La sénescence cellulaire consécutive à une prostatite à *E. coli pks+* pourrait donc favoriser la progression du cancer de la prostate comme évoqué dans le cancer colorectal. Par ailleurs, Redon et collaborateurs rapportent que des tumeurs implantées dans la région dorsale de souris induisent une inflammation chronique responsable de lésions tissulaires dans le tractus digestif et au niveau cutané [402]. Une action à distance des cellules sénescents induites par *E. coli pks+* sur le cancer de la prostate pourrait également être proposée. Nous pouvons ainsi imaginer que l'apparition de cellules sénescents au niveau intestinal sous l'action de bactéries productrices de colibactine, pourrait stimuler la prolifération de cellules malignes prostatiques grâce à la diffusion des facteurs de croissance sécrétés. Des expérimentations animales permettant de soutenir cette hypothèse sont nécessaires.

2. *E. coli* producteur de colibactine, sénescence prématurée et métastases

Le développement de métastases correspond au stade le plus avancé du cancer. Ce dernier va envahir les tissus adjacents puis ceux situés à distance. Ce processus repose notamment sur la migration des cellules suite à un remodelage de la matrice extra-cellulaire sous l'action de protéases telles que les métalloprotéases (MMP). Une étude histologique récente met en exergue l'importance de l'expression de la MMP-1 dans l'invasion et l'apparition de métastases [403]. Aucun épithélium normal ni adénome exprime MMP-1, alors que 85,2% des carcinomes invasifs l'expriment. L'augmentation de l'expression de MMP-1 serait donc associée à la progression tumorale. Il est à noter que la surexpression de MMP-1 peut être liée à la diminution d'expression de SENP1 [404]. De plus, des fibroblastes, rendues sénescents sous l'action de la bléomycine, participent à la croissance tumorale par la sécrétion de métalloprotéases et notamment MMP-1 [400].

Des résultats préliminaires nous ont permis d'observer une augmentation de la migration cellulaire sous l'action du surnageant de cellules sénescents induites par les bactéries *pks+*. Cette migration cellulaire a été objectivée à l'aide d'un test de cicatrisation des tapis cellulaires et un test d'invasion en matrigel (**Figure 36**). Par ailleurs, un zymogramme en gel de collagène



i g v C4 y
 4zf h y p i d v épp2 U n
 i y.n. i kC
 a v b o i mi
 d v épp2 i
 kC i i
 o o
 v o 2 q D. U q D. p cr u u L s
 u u L s o c s c e s

a révélé la production de MMP-1 dans les cellules infectées par les bactéries productrices de colibactine à 24 heures post-infection. Parallèlement, la quantité de transcrits codant pour les protéines inhibitrices de MMP, TIMP1 et TIMP2, est significativement diminuée dans ces cellules (20 et 5 fois respectivement par rapport à des cellules infectées par *E. coli pks-*). Ces résultats préliminaires suggèrent un rôle potentiel des bactéries *pks+* dans la dissémination des métastases. Les bactéries productrices de colibactine en stimulant la production de MMP-1 pourraient désorganiser la matrice extra-cellulaire. Nous proposons d'explorer plus avant ce processus dans un modèle murin de métastases.

3. *E. coli* producteur de colibactine, sénescence prématurée et angiogenèse

L'angiogenèse ou néovascularisation correspond à la formation de nouveaux vaisseaux sanguins. Ce processus permet d'irriguer les tumeurs et contribue à la dissémination des cellules malignes dans l'organisme. En conditions normales, il est régulé par l'établissement d'un équilibre entre la production de molécules pro- et anti-angiogéniques [405]. Au niveau des tumeurs, la croissance rapide des cellules entraîne une hypoxie, qui induit un blocage de la dégradation du facteur de transcription HIF1- α (hypoxia-inducible factor-1 α). Ce facteur de transcription cible des effecteurs pro-angiogéniques tels que le VEGF (vascular endothelial growth factor), le FGF, le HGF, l'uPAR (urokinase receptor), les MMPs, etc... D'autres facteurs décrits pour stimuler l'angiogenèse sont les médiateurs de l'inflammation IL-8 et les prostaglandines. L'épigénétique jouerait également un rôle central dans l'angiogenèse. Il a récemment été démontré que l'histone méthyl-tranfèrase EZH2 (Enhancer of Zest homologue 2) régulerait les niveaux de transcription du VEGF et de son récepteur lors de la carcinogenèse [406].

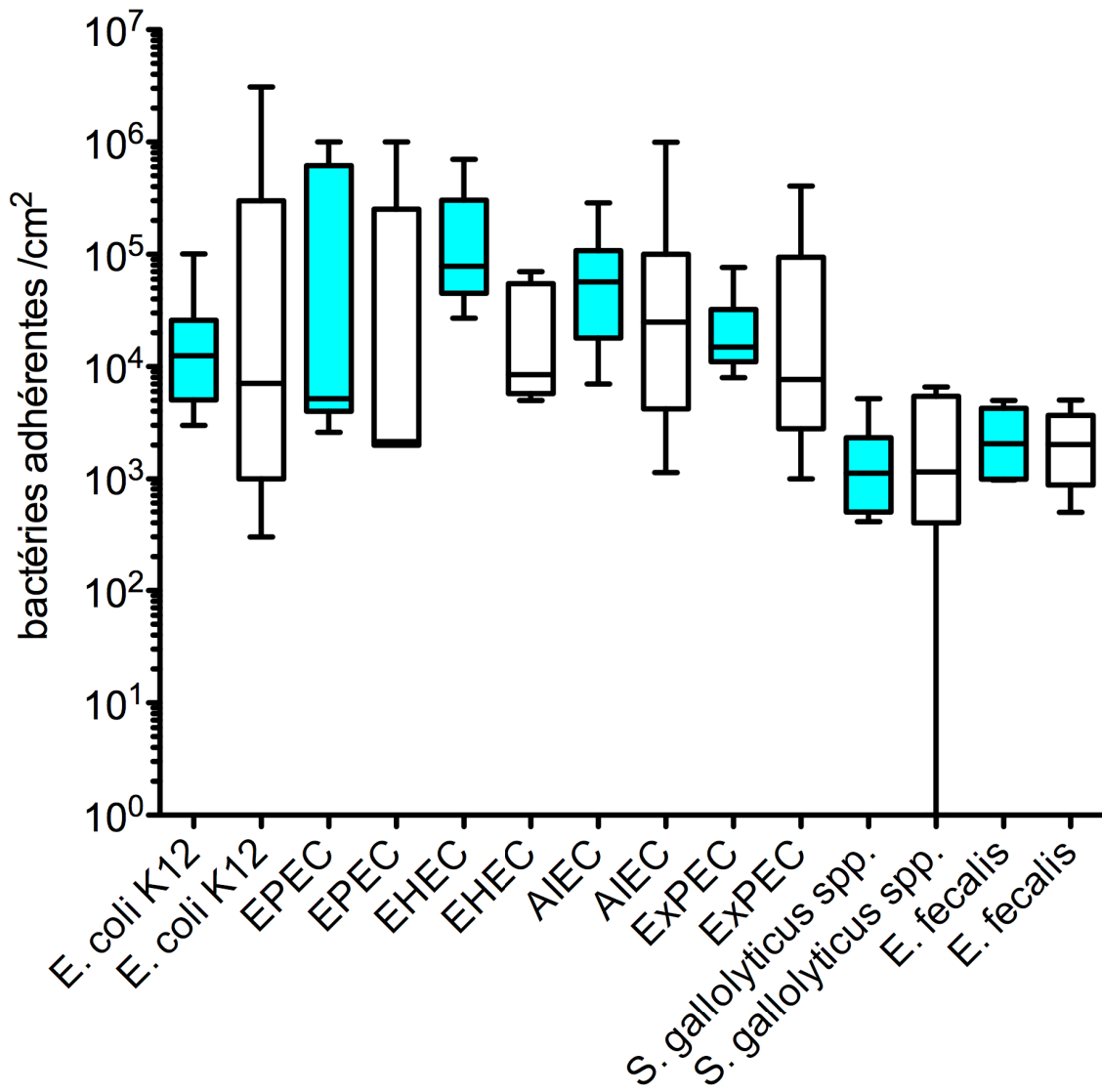
La caractérisation des cellules rendues sénescents sous l'action des bactéries productrices de colibactine a révélé une production accrue de HGF et FGF. Nous avons également observé une augmentation des transcrits codant l'IL-8, le VEGF, uPAR ainsi qu'une production accrue de MMP-1. L'ensemble de ces facteurs est décrit pour stimuler l'angiogenèse [250,403,405,407–409]. Les cellules sénescents pourraient donc également contribuer au développement tumoral en favorisant la vascularisation des tumeurs. Il serait donc intéressant d'étudier l'impact des produits du phénotype sécrétoire des cellules sénescents sur des cellules endothéliales. Une autre approche serait une évaluation histologique du niveau de vascularisation des tumeurs en présence et absence de cellules sénescents induites par les

bactéries *pks+*. Nous proposons également de rechercher une accumulation des protéines EZH2 et HIF1- α en réponse à l'infection par des bactéries *pks+*.

4. *E. coli* producteur de colibactine, sénescence prématurée et signalisation cellulaire

Nous avons identifié une nouvelle voie de signalisation cellulaire responsable de la sénescence cellulaire. Elle repose sur c-Myc, le microARN-20a-5p et SENP1, qui affecte la SUMOylation des protéines. Son originalité tient notamment dans le lien qu'elle établit entre des effecteurs globaux de la régulation que sont les microARN et la modification post-transcriptionnelle des protéines par SUMOylation. Parmi les protéines affectées par la SUMOylation, nous avons retrouvé p53, un acteur majeur de la sénescence cellulaire dont certaines cibles étaient modulées (p21^{Cip} et Rb). Cependant, il émerge actuellement de la littérature une signalisation de la sénescence plus complexe faisant intervenir des MAPK et de la voie PTEN/p27^{Kip1} [151,410]. La voie de signalisation dépendante des MAPK fait intervenir p16^{Ink4A}. L'expression de cette dernière ne semble pas modifiée dans les cellules rendues sénescents par les *E. coli pks+* car l'utilisation d'un ARN interférant dirigé contre p16^{INK4A} n'affectait pas l'entrée des cellules en sénescence. Reste à explorer une éventuelle modification de la voie PTEN/p27^{Kip1} en réponse à l'infection par *E. coli pks+*.

La sénescence cellulaire prématurée induite par des bactéries productrices de colibactine est associée à une dérégulation de la machinerie de SUMOylation des protéines. Des effecteurs bactériens et viraux ont été décrits pour détourner cette machinerie en mimant ou en affectant la stabilité ou l'expression de certains effecteurs majeurs de la SUMOylation comme la SUMO peptidase SENP1, l'enzyme d'activation E2 Ubc9 ou des E3-SUMO ligase [411–414]. A côté de la machinerie cellulaire de SUMOylation, il existe d'autres processus de régulation basés sur la modification post-transcriptionnelle des protéines comme l'Ubiquitination et NEDDylation qui contribuent à l'orientation des protéines vers le protéasome, ou l'ISGylation qui modulent la stabilité des protéines cibles [153]. La cyclomoduline Cif des *E. coli* entéropathogènes bloquent le cycle cellulaire en modifiant les protéines NEDD8 et Ubiquitine [357,415]. Il serait donc intéressant d'explorer l'effet des bactéries productrices de colibactine sur ces processus et d'explorer les conséquences éventuelles sur la progression du cancer colorectal.



LL s t c
 L e l LL o
 o



Nous avons observé une surexpression du microARN 20a-5p en réponse à l'infection par les *E. coli* pks+. Une exploration de l'expression des principaux oncomirs en réponse à l'infection montre que ces bactéries induisent une modification plus large du profil d'expression des microARN (données non publiées). L'analyse globale du microARNome de cellules en sénescence répliquative a été déjà réalisée [416], mais aucune donnée n'est disponible sur des cellules en sénescence prématurées. Il serait donc intéressant d'explorer les spécificités du microARNome et du transcriptome des cellules sénescents induites par *E. coli* pks+, afin de mieux comprendre l'implication des microARNs dans le blocage du cycle cellulaire et le phénotype sécrétoire de ces cellules.

Ces perspectives de travaux sur la signalisation cellulaire doivent être conduits en conservant à l'esprit qu'un des objectifs majeurs pourrait être de définir des cibles permettant de découpler les effets néfastes du phénotype sécrétoire associé à la sénescence des effets bénéfiques de la sénescence (blocage du cycle cellulaire). La sénescence induite par les bactéries productrices de colibactine pourrait alors être mise à profit dans le traitement du cancer colorectal en bloquant spécifiquement les voies de signalisation responsables du phénotype sécrétoire tout en préservant le blocage de la prolifération cellulaire induit par ces bactéries.

5. *E. coli* producteurs de colibactine, sénescence et dysbiose intestinale

Les protéines de surface « platelet-activating factor receptor » (PAFr), « Keratin 10 » (K10) et « Laminin Receptor » (LR) sont sur-exprimées dans les cellules sénescents [417]. Ces protéines sont utilisées comme récepteurs par des bactéries telles que *Streptococcus pneumoniae*, *Haemophilus influenzae* ou *Neisseria meningitidis* [280,417–420]. Les cellules sénescents pourraient donc porter un panel de protéines de surface original, pouvant influencer sur l'adhésion des bactéries et ainsi induire une dysbiose associée à la sénescence. Des tests d'adhésion sur cellules HCT116 normales ou sénescents ont été réalisés au laboratoire. Les données obtenues ont révélé que les cellules sénescents ne favorisent pas l'adhésion de *E. coli* quel que soit son pathovar (ExPEC, EPEC, EHEC, AIEC ou non pathogènes) ou de bactéries associées au cancer colorectal comme *Enterococcus faecalis* ou *Streptococcus gallolyticus* spp. *gallolyticus* (**Figure 37**) [13,110,294,319].

Cependant, il a été également montré que la présence de cellules sénescents

déstabilise les épithéliums en affectant les jonctions entre cellules [421]. Cette atteinte de l'intégrité de la barrière épithéliale pourrait favoriser l'infiltration d'agents pathogènes bactériens ou viraux et contribuer ainsi à l'apparition de septicémies ou d'endocardites lors de cancers colorectaux [309,318,422]. De même, la désorganisation de l'épithélium induite par la sénescence pourrait favoriser la diffusion des micro-organismes pathogènes, bactéries ou virus. L'état des jonctions serrées pourrait être exploré en réalisant l'immunomarquage des protéines ZO1 et des claudines, qui sont impliquées dans l'établissement de ces jonctions. En parallèle, l'activité déstabilisatrice des bactéries *pks*⁺ pourrait également être étudiée en explorant la translocation de bactéries et de macromolécules marquées (Dextran couplé à la fluoresceine isothiocyanate).

6. ClbP et la synthèse de la colibactine

Nos travaux associés à ceux de Reimer et collaborateurs [384] montrent que la protéine ClbP codée par l'îlot *pks* appartient à un sous-groupe de peptidases à sérine active assurant l'activation de composés de type NRP. Pour ce faire, ClbP possède une triade catalytique et un site actif chargé négativement, qui sont indispensables à l'activité biologique de l'îlot *pks*.

La bioactivité de ClbP nécessite sa translocation au périplasma grâce à son extrémité N-terminale *via* le système SRP/Sec/YidC. L'intérêt d'une localisation périplasmique de ClbP peut s'expliquer par l'activité génotoxique de la colibactine mature. En effet, localisée dans le périplasma, ClbP assure la bioactivation de la toxine à distance de l'ADN bactérien. Elle peut également s'expliquer par une impossibilité de la colibactine mature à être sécrétée hors du cytoplasme. Le peptide NRP clivé par ClbP pourrait servir de séquence d'adressage vers le périplasma à l'image d'un peptide signal classique.

L'activité génotoxique de la colibactine requiert également un ancrage à la membrane plasmique *via* 3 hélices transmembranaires localisées à l'extrémité C-terminale de la protéine. L'intérêt de cet ancrage à la membrane peut être justifié par une interaction entre ClbP et un système d'exportation de la pré-colibactine du cytoplasme vers le périplasma. Cette hypothèse est soutenue par la fusion de nombreuses ClbP-like avec un domaine codant une pompe d'efflux. Une multimérisation des hélices transmembranaires de ClbP pourrait constituer un système d'exportation ou ces hélices pourraient interagir avec des pompes d'efflux comme la pompe de type MATE codée par l'îlot *pks* (gène *clbM*). Cependant, des travaux en double hybride adaptés aux protéines membranaires et de « GST-pull-down » n'ont pas permis

d'observer une multimérisation de ClbP ou d'éventuelles interactions avec la pompe MATE. Une alternative à ces hypothèses serait une interaction avec la chaîne d'assemblage NRPS/PKS cytoplasmique ou directement avec la pré-colibactine. Cette dernière pourrait être associée à la membrane cytoplasmique car l'analyse des enzymes de la chaîne d'assemblage montre la présence d'un module enzymatique pouvant assurer l'intégration d'acides gras. Cette association de la pré-colibactine avec la membrane cytoplasmique grâce à un acide gras justifierait une localisation de ClbP à la membrane pour assurer sa maturation. Aussi, des expériences de « GST-pull down » ou d'immunoprécipitation pourraient être réalisées et suivies de l'analyse des ligands. Ces approches permettraient de déterminer la fonction de l'extrémité C-terminale de ClbP.

7. Inhibition de ClbP et prévention des effets délétères de la colibactine

La détermination de la structure, la connaissance du processus catalytique, la localisation extra-cytoplasmique et le rôle bioactivateur de ClbP font de cette dernière une cible de choix pour le développement de molécules bloquant la synthèse de colibactine. Nous avons montré qu'elle pouvait être inhibée par des acides boroniques « drug-like » et qu'elle constituait une cible permettant de bloquer l'activité génotoxique et pro-tumorale des *E. coli* producteurs de colibactine sans affecter la viabilité des bactéries. Le ciblage de ClbP permet donc de contrôler la virulence de la souche sans affecter sa survie. Il permet de bloquer l'activité délétère des bactéries sans exercer de pression de sélection susceptible de faire émerger des bactéries résistantes au traitement. Cette stratégie, qui est rarement utilisée dans le traitement des maladies infectieuses d'origine bactérienne, permettrait aussi d'éviter les effets collatéraux provoqués par une antibiothérapie. En effet, cette dernière est associée à des risques de modification majeure de la flore digestive pouvant conduire à des pathologies graves comme la colite pseudomembraneuse à *Clostridium difficile* [423]. Une stratégie visant à contrôler les effets de l'îlot *pks* sur la progression tumorale sans rechercher l'élimination des bactéries porteuses permet de contourner ces problématiques, qui peuvent être majeures dans une affection chronique telle que le cancer colorectal.

En conclusion, nos travaux ont permis de confirmer le rôle des bactéries productrices de colibactine dans la progression du cancer colorectal. Elles permettraient l'émergence d'un microenvironnement sénescence susceptible de favoriser la prolifération des cellules adjacentes.

Étant donné la fréquence de ces bactéries au contact des tumeurs colorectales, il paraissait intéressant de développer des stratégies visant à bloquer l'expression de la colibactine. L'étude de la structure et de la fonction de ClbP a permis d'identifier des molécules « drug-like » prévenant l'activité génotoxique et pro-tumorale des bactéries productrices. Cette stratégie pourrait compléter l'arsenal thérapeutique utilisé pour prévenir ou traiter le cancer colorectal.

IV Références bibliographiques

1. Cancer of the Colon and Rectum - SEER Stat Fact Sheets. <http://seer.cancer.gov/statfacts/html>
2. Facteurs de risque du cancer colorectale. <http://info.cancer.ca>
3. Delattre O, Olschwang S, Law DJ, Melot T, Remvikos Y, Salmon RJ, et al. Multiple genetic alterations in distal and proximal colorectal cancer. *Lancet*. 1989 Aug 12;2(8659):353–6.
4. Cunningham D, Atkin W, Lenz H-J, Lynch HT, Minsky B, Nordlinger B, et al. Colorectal cancer. *Lancet*. 2010 Mar 20;375(9719):1030–47.
5. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990 Jun 1;61(5):759–67.
6. Weitz J, Koch M, Debus J, Höhler T, Galle PR, Büchler MW. Colorectal cancer. *Lancet*. 2005 Jan 8;365(9454):153–65.
7. Nougayrède J-P, Oswald E. [Microbiota and colorectal cancer: genotoxic bacteria in the intestinal tract]. *Bull. Acad. Natl. Med.* 2011 Jun;195(6):1295–1304; discussion 1304–1305.
8. Hope ME, Hold GL, Kain R, El-Omar EM. Sporadic colorectal cancer--role of the commensal microbiota. *FEMS Microbiol. Lett.* 2005 Mar 1;244(1):1–7.
9. van Ooij C. Bacterial toxins: *Escherichia coli* damages host DNA. *Nat. Rev. Microbiol.* 2010 Aug;8(8):534.
10. Tjalsma H, Boleij A, Marchesi JR, Dutilh BE. A bacterial driver-passenger model for colorectal cancer: beyond the usual suspects. *Nature reviews. Microbiology.* 2012 Jun 25.
11. Wu S, Rhee K-J, Zhang M, Franco A, Sears CL. *Bacteroides fragilis* toxin stimulates intestinal epithelial cell shedding and gamma-secretase-dependent E-cadherin cleavage. *J. Cell. Sci.* 2007 Jun 1;120(Pt 11):1944–52.
12. Wu S, Rhee K-J, Albesiano E, Rabizadeh S, Wu X, Yen H-R, et al. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* 2009 Sep;15(9):1016–22.
13. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T-J, et al. Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. *Science*. 2012 Aug 16.
14. Cuevas-Ramos G, Petit CR, Marcq I, Boury M, Oswald E, Nougayrède J-P. *Escherichia coli* induces DNA damage in vivo and triggers genomic instability in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* 2010 Jun 22;107(25):11537–42.
15. Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat. Rev. Microbiol.* 2004 Feb;2(2):123–40.
16. Johnson JR, Russo TA. Extraintestinal pathogenic *Escherichia coli*: 'the other bad *E. coli*'. *J. Lab. Clin. Med.* 2002 Mar;139(3):155–62.
17. Hacker J, Carniel E. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep.* 2001 May;2(5):376–81.
18. Clements A, Young J, Constantinou N, Frankel G. Infection strategies of enteric pathogenic *E. coli*. *Gut microbes.* 2012 Mar 1.
19. Mukhopadhyay I, Hansen R, El-Omar EM, Hold GL. IBD-what role do Proteobacteria play? *Nat Rev Gastroenterol Hepatol.* 2012 Apr;9(4):219–30.
20. Köhler C-D, Dobrindt U. What defines extraintestinal pathogenic *Escherichia coli*? *Int. J. Med. Microbiol.* 2011 Dec;301(8):642–7.
21. Pitout JDD. Extraintestinal Pathogenic *Escherichia coli*: A Combination of Virulence with Antibiotic Resistance. *Front Microbiol.* 2012;3:9.

22. Wiles TJ, Kulesus RR, Mulvey MA. Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp. Mol. Pathol.* 2008 Aug;85(1):11–9.
23. Mokady D, Gophna U, Ron EZ. Virulence factors of septicemic *Escherichia coli* strains. *Int. J. Med. Microbiol.* 2005 Oct;295(6-7):455–62.
24. Escobar-Páramo P, Clermont O, Blanc-Potard A-B, Bui H, Le Bouguéne C, Denamur E. A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol. Biol. Evol.* 2004 Jun;21(6):1085–94.
25. Nougayrède J-P, Taieb F, De Rycke J, Oswald E. Cyclomodulins: bacterial effectors that modulate the eukaryotic cell cycle. *Trends Microbiol.* 2005 Mar;13(3):103–10.
26. Nougayrède J-P, Homburg S, Taieb F, Boury M, Brzuszkiewicz E, Gottschalk G, et al. *Escherichia coli* induces DNA double-strand breaks in eukaryotic cells. *Science.* 2006 Aug 11;313(5788):848–51.
27. Walsh CT. Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science.* 2004 Mar 19;303(5665):1805–10.
28. McCracken VJ, Lorenz RG. The gastrointestinal ecosystem: a precarious alliance among epithelium, immunity and microbiota. *Cell. Microbiol.* 2001 Jan;3(1):1–11.
29. Shirota K, LeDuy L, Yuan SY, Jothy S. Interleukin-6 and its receptor are expressed in human intestinal epithelial cells. *Virchows Arch., B, Cell Pathol.* 1990;58(4):303–8.
30. Eckmann L, Kagnoff MF, Fierer J. Epithelial cells secrete the chemokine interleukin-8 in response to bacterial entry. *Infect. Immun.* 1993 Nov;61(11):4569–74.
31. Jung HC, Eckmann L, Yang SK, Panja A, Fierer J, Morzycka-Wroblewska E, et al. A distinct array of proinflammatory cytokines is expressed in human colon epithelial cells in response to bacterial invasion. *J. Clin. Invest.* 1995 Jan;95(1):55–65.
32. Schiffrin EJ, Blum S. Interactions between the microbiota and the intestinal mucosa. *Eur J Clin Nutr.* 2002 Aug;56 Suppl 3:S60–64.
33. Eckmann L, Kagnoff MF, Fierer J. Intestinal epithelial cells as watchdogs for the natural immune system. *Trends Microbiol.* 1995 Mar;3(3):118–20.
34. Huttner KM, Bevins CL. Antimicrobial peptides as mediators of epithelial host defense. *Pediatr. Res.* 1999 Jun;45(6):785–94.
35. Müller CA, Autenrieth IB, Peschel A. Innate defenses of the intestinal epithelial barrier. *Cell. Mol. Life Sci.* 2005 Jun;62(12):1297–307.
36. Masuda K, Nakamura K, Yoshioka S, Fukaya R, Sakai N, Ayabe T. Regulation of microbiota by antimicrobial peptides in the gut. *Adv. Otorhinolaryngol.* 2011;72:97–9.
37. Furness JB, Kunze WA, Clerc N. Nutrient tasting and signaling mechanisms in the gut. II. The intestine as a sensory organ: neural, endocrine, and immune responses. *Am. J. Physiol.* 1999 Nov;277(5 Pt 1):G922–928.
38. Bourlioux P, Koletzko B, Guarner F, Braesco V. The intestine and its microflora are partners for the protection of the host: report on the Danone Symposium 'The Intelligent Intestine,' held in Paris, June 14, 2002. *Am. J. Clin. Nutr.* 2003 Oct;78(4):675–83.
39. Corfield AP, Myerscough N, Longman R, Sylvester P, Arul S, Pignatelli M. Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. *Gut.* 2000 Oct;47(4):589–94.
40. Plaut AG. Trefoil peptides in the defense of the gastrointestinal tract. *N. Engl. J. Med.* 1997 Feb 13;336(7):506–7.

41. Ouellette AJ. Paneth cells and innate immunity in the crypt microenvironment. *Gastroenterology*. 1997 Nov;113(5):1779–84.
42. Stappenbeck TS, Hooper LV, Gordon JI. Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. *Proc. Natl. Acad. Sci. U.S.A.* 2002 Nov 26;99(24):15451–5.
43. Fujihashi K, Kweon MN, Kiyono H, VanCott JL, van Ginkel FW, Yamamoto M, et al. A T cell/B cell/epithelial cell internet for mucosal inflammation and immunity. *Springer Semin. Immunopathol.* 1997;18(4):477–94.
44. Wells JM, Rossi O, Meijerink M, van Baarlen P. Colloquium Paper: Epithelial crosstalk at the microbiota-mucosal interface. *Proceedings of the National Academy of Sciences*. 2010 Sep 8;108(Supplement_1):4607–14.
45. Gregorieff A, Clevers H. Wnt signaling in the intestinal epithelium: from endoderm to cancer. *Genes Dev.* 2005 Apr 15;19(8):877–90.
46. Crosnier C, Vargesson N, Gschmeissner S, Ariza-McNaughton L, Morrison A, Lewis J. Delta-Notch signalling controls commitment to a secretory fate in the zebrafish intestine. *Development*. 2005 Jan 3;132(5):1093–104.
47. Casali A, Battle E. Intestinal stem cells in mammals and *Drosophila*. *Cell Stem Cell*. 2009 Feb 6;4(2):124–7.
48. Crosnier C, Stamatakis D, Lewis J. Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. *Nat. Rev. Genet.* 2006 May;7(5):349–59.
49. Heldin CH, Miyazono K, ten Dijke P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*. 1997 Dec 4;390(6659):465–71.
50. Hardwick JCH, Van Den Brink GR, Bleuming SA, Ballester I, Van Den Brande JMH, Keller JJ, et al. Bone morphogenetic protein 2 is expressed by, and acts upon, mature epithelial cells in the colon. *Gastroenterology*. 2004 Jan;126(1):111–21.
51. Ingham PW, McMahon AP. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev.* 2001 Dec 1;15(23):3059–87.
52. van den Brink GR, Hardwick JCH, Nielsen C, Xu C, ten Kate FJ, Glickman J, et al. Sonic hedgehog expression correlates with fundic gland differentiation in the adult gastrointestinal tract. *Gut*. 2002 Nov;51(5):628–33.
53. van den Brink GR, Bleuming SA, Hardwick JCH, Schepman BL, Offerhaus GJ, Keller JJ, et al. Indian Hedgehog is an antagonist of Wnt signaling in colonic epithelial cell differentiation. *Nat. Genet.* 2004 Mar;36(3):277–82.
54. Jiménez E, Fernández L, Marín ML, Martín R, Odriozola JM, Nueno-Palop C, et al. Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. *Curr. Microbiol.* 2005 Oct;51(4):270–4.
55. Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* 1999 May;69(5):1035S–1045S.
56. Savage DC. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* 1977;31:107–33.
57. Holdeman LV, Good IJ, Moore WE. Human fecal flora: variation in bacterial composition within individuals and a possible effect of emotional stress. *Appl. Environ. Microbiol.* 1976 Mar;31(3):359–75.
58. Wei C, Brent MR. Using ESTs to improve the accuracy of de novo gene prediction. *BMC Bioinformatics*. 2006;7:327.
59. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010 Mar 4;464(7285):59–65.

60. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. *Science*. 2008 Jun 20;320(5883):1647–51.
61. Bäckhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. Host-bacterial mutualism in the human intestine. *Science*. 2005 Mar 25;307(5717):1915–20.
62. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell*. 2005 Jul 15;122(1):107–18.
63. Hooper LV, Gordon JI. Commensal host-bacterial relationships in the gut. *Science*. 2001 May 11;292(5519):1115–8.
64. Dethlefsen L, McFall-Ngai M, Relman DA. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*. 2007 Oct 18;449(7164):811–8.
65. Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD, et al. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl. Environ. Microbiol.* 1999 Nov;65(11):4799–807.
66. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, et al. Diversity of the human intestinal microbial flora. *Science*. 2005 Jun 10;308(5728):1635–8.
67. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*. 2006 Feb;55(2):205–11.
68. Rigottier-Gois L, Rochet V, Garrec N, Suau A, Doré J. Enumeration of *Bacteroides* species in human faeces by fluorescent in situ hybridisation combined with flow cytometry using 16S rRNA probes. *Syst. Appl. Microbiol.* 2003 Mar;26(1):110–8.
69. Rigottier-Gois L, Bourhis A-G, Gramet G, Rochet V, Doré J. Fluorescent hybridisation combined with flow cytometry and hybridisation of total RNA to analyse the composition of microbial communities in human faeces using 16S rRNA probes. *FEMS Microbiol. Ecol.* 2003 Mar 1;43(2):237–45.
70. Favier CF, Vaughan EE, De Vos WM, Akkermans ADL. Molecular monitoring of succession of bacterial communities in human neonates. *Appl. Environ. Microbiol.* 2002 Jan;68(1):219–26.
71. Zoetendal EG, Akkermans AD, De Vos WM. Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Appl. Environ. Microbiol.* 1998 Oct;64(10):3854–9.
72. Callister SJ, McCue LA, Turse JE, Monroe ME, Auberry KJ, Smith RD, et al. Comparative Bacterial Proteomics: Analysis of the Core Genome Concept. *PLoS ONE*. 2008 Feb 6;3(2):e1542.
73. Turnbaugh PJ, Gordon JI. The core gut microbiome, energy balance and obesity. *J. Physiol. (Lond.)*. 2009 Sep 1;587(Pt 17):4153–8.
74. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, et al. Human gut microbiome viewed across age and geography. *Nature*. 2012 Jun 14;486(7402):222–7.
75. Turnbaugh PJ, Quince C, Faith JJ, McHardy AC, Yatsunenko T, Niazi F, et al. Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proc. Natl. Acad. Sci. U.S.A.* 2010 Apr 20;107(16):7503–8.
76. Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, et al. A core gut microbiome in obese and lean twins. *Nature*. 2009 Jan 22;457(7228):480–4.
77. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. The human microbiome project. *Nature*. 2007 Oct 18;449(7164):804–10.
78. van Tongeren SP, Slaets JPJ, Harmsen HJM, Welling GW. Fecal microbiota composition and frailty. *Appl.*

- Environ. Microbiol. 2005 Oct;71(10):6438–42.
79. Woodmansey EJ. Intestinal bacteria and ageing. *J. Appl. Microbiol.* 2007 May;102(5):1178–86.
 80. Mariat D, Firmesse O, Levenez F, Guimaraes V, Sokol H, Doré J, et al. The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiol.* 2009;9:123.
 81. Guigoz Y, Doré J, Schiffrin EJ. The inflammatory status of old age can be nurtured from the intestinal environment. *Curr Opin Clin Nutr Metab Care.* 2008 Jan;11(1):13–20.
 82. Dobson A, Cotter PD, Ross RP, Hill C. Bacteriocin production: a probiotic trait? *Appl. Environ. Microbiol.* 2012 Jan;78(1):1–6.
 83. Sonnenburg JL, Xu J, Leip DD, Chen C-H, Westover BP, Weatherford J, et al. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. *Science.* 2005 Mar 25;307(5717):1955–9.
 84. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, et al. The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. U.S.A.* 2004 Nov 2;101(44):15718–23.
 85. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature.* 2006 Dec 21;444(7122):1027–31.
 86. Bäckhed F, Manchester JK, Semenkovich CF, Gordon JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc. Natl. Acad. Sci. U.S.A.* 2007 Jan 16;104(3):979–84.
 87. Gershwin ME, Schiffrin E. Probiotics and immunity. Introduction. *Clin Rev Allergy Immunol.* 2002 Jun;22(3):205–6.
 88. Alam M, Midtvedt T, Uribe A. Differential cell kinetics in the ileum and colon of germfree rats. *Scand. J. Gastroenterol.* 1994 May;29(5):445–51.
 89. Bondada S, Wu H, Robertson DA, Chelvarajan RL. Accessory cell defect in unresponsiveness of neonates and aged to polysaccharide vaccines. *Vaccine.* 2000 Oct 15;19(4-5):557–65.
 90. Sudo N, Yu X-N, Aiba Y, Oyama N, Sonoda J, Koga Y, et al. An oral introduction of intestinal bacteria prevents the development of a long-term Th2-skewed immunological memory induced by neonatal antibiotic treatment in mice. *Clin. Exp. Allergy.* 2002 Jul;32(7):1112–6.
 91. Wostmann BS, Pleasants JR. The germ-free animal fed chemically defined diet: a unique tool. *Proc. Soc. Exp. Biol. Med.* 1991 Oct;198(1):539–46.
 92. Kiyono H, Michalek SM, Mosteller LM, Torii M, Hamada S, McGhee JR. Enhancement of murine immune responses to orally administered haptened Streptococcus mutans. *Scand. J. Immunol.* 1982 Dec;16(6):455–63.
 93. Sudo N, Sawamura S, Tanaka K, Aiba Y, Kubo C, Koga Y. The requirement of intestinal bacterial flora for the development of an IgE production system fully susceptible to oral tolerance induction. *J. Immunol.* 1997 Aug 15;159(4):1739–45.
 94. Strobel S, Mowat AM. Immune responses to dietary antigens: oral tolerance. *Immunol. Today.* 1998 Apr;19(4):173–81.
 95. Macpherson AJ, Harris NL. Interactions between commensal intestinal bacteria and the immune system. *Nat. Rev. Immunol.* 2004 Jun;4(6):478–85.
 96. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: human gut microbes associated with obesity. *Nature.* 2006 Dec 21;444(7122):1022–3.
 97. Baker PI, Love DR, Ferguson LR. Role of gut microbiota in Crohn's disease. *Expert Rev Gastroenterol Hepatol.* 2009 Oct;3(5):535–46.

98. Hammer HF. Gut microbiota and inflammatory bowel disease. *Dig Dis.* 2011;29(6):550–3.
99. Rajilic-Stojanovic Mirjana. The core microbiota within us [Internet]. Wageningen University; 2007. Available from: <http://www.cost853.ch/pdf-files/Wageningen2006/Rajilic-Stojanovic-presentationCOST853-june06.pdf>
100. Macdonald TT, Monteleone G. Immunity, inflammation, and allergy in the gut. *Science.* 2005 Mar 25;307(5717):1920–5.
101. Penders J, Stobberingh EE, van den Brandt PA, Thijs C. The role of the intestinal microbiota in the development of atopic disorders. *Allergy.* 2007 Nov;62(11):1223–36.
102. Penders J, Stobberingh EE, Thijs C, Adams H, Vink C, van Ree R, et al. Molecular fingerprinting of the intestinal microbiota of infants in whom atopic eczema was or was not developing. *Clin. Exp. Allergy.* 2006 Dec;36(12):1602–8.
103. Compare D, Nardone G. Contribution of gut microbiota to colonic and extracolonic cancer development. *Dig Dis.* 2011;29(6):554–61.
104. Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* 2009 May;9(5):313–23.
105. Sansonetti PJ. Host-bacteria homeostasis in the healthy and inflamed gut. *Curr. Opin. Gastroenterol.* 2008 Jul;24(4):435–9.
106. Sansonetti PJ. War and peace at the intestinal epithelial surface: an integrated view of bacterial commensalism versus bacterial pathogenicity. *J. Pediatr. Gastroenterol. Nutr.* 2008 Apr;46 Suppl 1:E6–7.
107. Sansonetti PJ, Di Santo JP. Debugging how bacteria manipulate the immune response. *Immunity.* 2007 Feb;26(2):149–61.
108. McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, et al. The Macaque Gut Microbiome in Health, Lentiviral Infection, and Chronic Enterocolitis. *PLoS Pathogens.* 2008;4(2):e20.
109. Rawls JF. Enteric infection and inflammation alter gut microbial ecology. *Cell Host Microbe.* 2007 Aug 16;2(2):73–4.
110. Martin HM, Campbell BJ, Hart CA, Mpofu C, Nayar M, Singh R, et al. Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. *Gastroenterology.* 2004 Jul;127(1):80–93.
111. Strober W. Adherent-invasive *E. coli* in Crohn disease: bacterial 'agent provocateur'. *J. Clin. Invest.* 2011 Mar;121(3):841–4.
112. Jensen SR, Fink LN, Struve C, Sternberg C, Andersen JB, Brynskov J, et al. Quantification of specific *E. coli* in gut mucosa from Crohn's disease patients. *J. Microbiol. Methods.* 2011 Jul;86(1):111–4.
113. Keighley MR, Arabi Y, Dimock F, Burdon DW, Allan RN, Alexander-Williams J. Influence of inflammatory bowel disease on intestinal microflora. *Gut.* 1978 Dec;19(12):1099–104.
114. Darfeuille-Michaud A, Boudeau J, Bulois P, Neut C, Glasser A-L, Barnich N, et al. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. *Gastroenterology.* 2004 Aug;127(2):412–21.
115. Stecher B, Hardt W-D. The role of microbiota in infectious disease. *Trends Microbiol.* 2008 Mar;16(3):107–14.
116. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, et al. *Salmonella enterica* serovar typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol.* 2007 Oct;5(10):2177–89.
117. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, et al. Host-mediated

- inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. *Cell Host Microbe*. 2007 Sep 13;2(3):204.
118. Salminen S, Bouley C, Boutron-Ruault MC, Cummings JH, Franck A, Gibson GR, et al. Functional food science and gastrointestinal physiology and function. *Br. J. Nutr.* 1998 Aug;80 Suppl 1:S147–171.
 119. Thomas J. Kindt, Richard A. Goldsby, Barbara A. Osborne. *Immunologie : le cour de Janis Kuby*. 6ème édition.
 120. Neutra MR, Kraehenbuhl JP. Transepithelial transport and mucosal defence I: the role of M cells. *Trends Cell Biol.* 1992 May;2(5):134–8.
 121. Kraehenbuhl JP, Neutra MR. Epithelial M cells: differentiation and function. *Annu. Rev. Cell Dev. Biol.* 2000;16:301–32.
 122. Kraehenbuhl JP, Neutra MR. Molecular and cellular basis of immune protection of mucosal surfaces. *Physiol. Rev.* 1992 Oct;72(4):853–79.
 123. Kraehenbuhl JP, Neutra MR. Transepithelial transport and mucosal defence II: secretion of IgA. *Trends Cell Biol.* 1992 Jun;2(6):170–4.
 124. Mowat AM, Donachie AM, Parker LA, Robson NC, Beacock-Sharp H, McIntyre LJ, et al. The role of dendritic cells in regulating mucosal immunity and tolerance. *Novartis Found. Symp.* 2003;252:291–302; discussion 302–305.
 125. Sansonetti PJ, Medzhitov R. Learning tolerance while fighting ignorance. *Cell.* 2009 Aug 7;138(3):416–20.
 126. Maynard CL, Weaver CT. Intestinal effector T cells in health and disease. *Immunity.* 2009 Sep 18;31(3):389–400.
 127. Hochegger H, Takeda S, Hunt T. Cyclin-dependent kinases and cell-cycle transitions: does one fit all? *Nat. Rev. Mol. Cell Biol.* 2008 Nov;9(11):910–6.
 128. Bennett RA, Wilson DM 3rd, Wong D, Demple B. Interaction of human apurinic endonuclease and DNA polymerase beta in the base excision repair pathway. *Proc. Natl. Acad. Sci. U.S.A.* 1997 Jul 8;94(14):7166–9.
 129. Prasad R, Dianov GL, Bohr VA, Wilson SH. FEN1 stimulation of DNA polymerase beta mediates an excision step in mammalian long patch base excision repair. *J. Biol. Chem.* 2000 Feb 11;275(6):4460–6.
 130. Robins P, Harris AL, Goldsmith I, Lindahl T. Cross-linking of DNA induced by chloroethylnitrosourea is presented by O6-methylguanine-DNA methyltransferase. *Nucleic Acids Res.* 1983 Nov 25;11(22):7743–58.
 131. de Jager M, Dronkert ML, Modesti M, Beerens CE, Kanaar R, van Gent DC. DNA-binding and strand-annealing activities of human Mre11: implications for its roles in DNA double-strand break repair pathways. *Nucleic Acids Res.* 2001 Mar 15;29(6):1317–25.
 132. Hefferin ML, Tomkinson AE. Mechanism of DNA double-strand break repair by non-homologous end joining. *DNA Repair (Amst.)*. 2005 Jun 8;4(6):639–48.
 133. Ameisen JC, Groux H, Capron A. [Programmed death (apoptosis) of T CD4 lymphocytes and AIDS pathogenesis]. *C. R. Acad. Sci. III, Sci. Vie.* 1992;314(9 Suppl):47–50.
 134. LeBrun DP, Warnke RA, Cleary ML. Expression of bcl-2 in fetal tissues suggests a role in morphogenesis. *Am. J. Pathol.* 1993 Mar;142(3):743–53.
 135. Wyllie AH, Kerr JF, Currie AR. Cell death: the significance of apoptosis. *Int. Rev. Cytol.* 1980;68:251–306.
 136. Alison MR, Sarraf CE. Apoptosis: a gene-directed programme of cell death. *J R Coll Physicians Lond.* 1992 Jan;26(1):25–35.

137. Fink SL, Cookson BT. Apoptosis, Pyroptosis, and Necrosis: Mechanistic Description of Dead and Dying Eukaryotic Cells. *Infect Immun*. 2005 Apr;73(4):1907–16.
138. d'Adda di Fagagna F. Living on a break: cellular senescence as a DNA-damage response. *Nature Reviews Cancer*. 2008 Jul;8(7):512–22.
139. Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*. 1997 Mar 7;88(5):593–602.
140. Yates KE, Korbel GA, Shtutman M, Roninson IB, DiMaio D. Repression of the SUMO-specific protease Snp1 induces p53-dependent premature senescence in normal human fibroblasts. *Aging Cell*. 2008 Oct;7(5):609–21.
141. Cristofalo VJ, Lorenzini A, Allen RG, Torres C, Tresini M. Replicative senescence: a critical review. *Mech. Ageing Dev*. 2004 Nov;125(10-11):827–48.
142. HAYFLICK L. THE LIMITED IN VITRO LIFETIME OF HUMAN DIPLOID CELL STRAINS. *Exp. Cell Res*. 1965 Mar;37:614–36.
143. Shrivastav M, De Haro LP, Nickoloff JA. Regulation of DNA double-strand break repair pathway choice. *Cell Res*. 2008 Jan;18(1):134–47.
144. Larsson ola. cellular senescence-an integrated perspective. 2005;Cancer Therapy Vol 3:495–510.
145. Ben-Porath I, Weinberg RA. The signals and pathways activating cellular senescence. *Int. J. Biochem. Cell Biol*. 2005 May;37(5):961–76.
146. Banin S, Moyal L, Shieh S, Taya Y, Anderson CW, Chessa L, et al. Enhanced phosphorylation of p53 by ATM in response to DNA damage. *Science*. 1998 Sep 11;281(5383):1674–7.
147. Ogawara Y, Kishishita S, Obata T, Isazawa Y, Suzuki T, Tanaka K, et al. Akt enhances Mdm2-mediated ubiquitination and degradation of p53. *J. Biol. Chem*. 2002 Jun 14;277(24):21843–50.
148. Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, Brown JP, et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science*. 1998 Nov 20;282(5393):1497–501.
149. Han Z, Wei W, Dunaway S, Darnowski JW, Calabresi P, Sedivy J, et al. Role of p21 in apoptosis and senescence of human colon cancer cells treated with camptothecin. *J. Biol. Chem*. 2002 May 10;277(19):17154–60.
150. McConnell BB, Starborg M, Brookes S, Peters G. Inhibitors of cyclin-dependent kinases induce features of replicative senescence in early passage human diploid fibroblasts. *Curr. Biol*. 1998 Mar 12;8(6):351–4.
151. Alimonti A, Nardella C, Chen Z, Clohessy JG, Carracedo A, Trotman LC, et al. A novel type of cellular senescence that can be enhanced in mouse models and human tumor xenografts to suppress prostate tumorigenesis. *J. Clin. Invest*. 2010 Mar;120(3):681–93.
152. Andreou AM, Tavernarakis N. SUMOylation and cell signalling. *Biotechnol J*. 2009 Dec;4(12):1740–52.
153. van der Veen AG, Ploegh HL. Ubiquitin-like proteins. *Annu. Rev. Biochem*. 2012;81:323–57.
154. Bayer P, Arndt A, Metzger S, Mahajan R, Melchior F, Jaenicke R, et al. Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol*. 1998 Jul 10;280(2):275–86.
155. Grant MM. Identification of SUMOylated proteins in neuroblastoma cells after treatment with hydrogen peroxide or ascorbate. *BMB Rep*. 2010 Nov;43(11):720–5.
156. Wilson VG. Sumoylation at the Host-Pathogen Interface. *Biomolecules*. 2012 Apr 5;2(2):203–27.
157. Andreou AM, Tavernarakis N. Roles for SUMO modification during senescence. *Adv. Exp. Med. Biol*.

2010;694:160–71.

158. Liu B, Shuai K. Targeting the PIAS1 SUMO ligase pathway to control inflammation. *Trends Pharmacol. Sci.* 2008 Oct;29(10):505–9.
159. Vertegaal ACO. Uncovering ubiquitin and ubiquitin-like signaling networks. *Chem. Rev.* 2011 Dec 14;111(12):7923–40.
160. Lima CD, Reverter D. Structure of the human SENP7 catalytic domain and poly-SUMO deconjugation activities for SENP6 and SENP7. *J. Biol. Chem.* 2008 Nov 14;283(46):32045–55.
161. Xu Z, Chau SF, Lam KH, Chan HY, Ng TB, Au SWN. Crystal structure of the SENP1 mutant C603S-SUMO complex reveals the hydrolytic mechanism of SUMO-specific protease. *Biochem. J.* 2006 Sep 15;398(3):345–52.
162. Bischof O, Dejean A. SUMO is growing senescent. *Cell Cycle.* 2007 Mar 15;6(6):677–81.
163. Bischof O, Schwamborn K, Martin N, Werner A, Sustmann C, Grosschedl R, et al. The E3 SUMO ligase PIASy is a regulator of cellular senescence and apoptosis. *Mol. Cell.* 2006 Jun 23;22(6):783–94.
164. Sugrue MM, Shin DY, Lee SW, Aaronson SA. Wild-type p53 triggers a rapid senescence program in human tumor cells lacking functional p53. *Proc. Natl. Acad. Sci. U.S.A.* 1997 Sep 2;94(18):9648–53.
165. Xu Y, Li J, Zuo Y, Deng J, Wang L-S, Chen G-Q. SUMO-specific protease 1 regulates the in vitro and in vivo growth of colon cancer cells with the upregulated expression of CDK inhibitors. *Cancer Lett.* 2011 Oct 1;309(1):78–84.
166. Dou H, Huang C, Van Nguyen T, Lu L-S, Yeh ETH. SUMOylation and de-SUMOylation in response to DNA damage. *FEBS Lett.* 2011 Sep 16;585(18):2891–6.
167. David-Watine B. Silencing Nuclear Pore Protein Tpr Elicits a Senescent-Like Phenotype in Cancer Cells. *PLoS ONE.* 2011 Jul 19;6(7):e22423.
168. Dimri GP, Lee X, Basile G, Acosta M, Scott G, Roskelley C, et al. A biomarker that identifies senescent human cells in culture and in aging skin in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 1995 Sep 26;92(20):9363–7.
169. te Poele RH, Okorokov AL, Jardine L, Cummings J, Joel SP. DNA damage is able to induce senescence in tumor cells in vitro and in vivo. *Cancer Res.* 2002 Mar 15;62(6):1876–83.
170. Bischof O, Dejean A, Pineau P. [A review of cellular senescence: friend or foe of tumorigenesis?]. *Med Sci (Paris).* 2009 Feb;25(2):153–60.
171. Lee BY, Han JA, Im JS, Morrone A, Johung K, Goodwin EC, et al. Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. *Aging Cell.* 2006 Apr;5(2):187–95.
172. Coppé J-P, Desprez P-Y, Krtolica A, Campisi J. The Senescence-Associated Secretory Phenotype: The Dark Side of Tumor Suppression. *Annual Review of Pathology: Mechanisms of Disease.* 2010;5(1):99–118.
173. Malumbres M, Barbacid M. RAS oncogenes: the first 30 years. *Nat. Rev. Cancer.* 2003 Jun;3(6):459–65.
174. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell.* 2000 Jan 7;100(1):57–70.
175. InVS | BEH n°02-03 (13 janvier 2009). Numéro thématique - Dépistage organisé du cancer colorectal en France [Internet]. [cited 2012 Sep 4]. Available from: http://www.invs.sante.fr/beh/2009/02_03/index.htm#2
176. Belot A, Grosclaude P, Bossard N, Jouglu E, Benhamou E, Delafosse P, et al. Cancer incidence and mortality in France over the period 1980-2005. *Rev Epidemiol Sante Publique.* 2008 Jun;56(3):159–75.
177. Lynch HT, Lynch JF. Genetics of colonic cancer. *Digestion.* 1998 Aug;59(5):481–92.

178. Li G-M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* 2008 Jan;18(1):85–98.
179. Kantelinen J, Kansikas M, Candelin S, Hampel H, Smith B, Holm L, et al. Mismatch repair analysis of inherited MSH2 and/or MSH6 variation pairs found in cancer patients. *Human mutation* [Internet]. 2012 May 11 [cited 2012 Jun 3]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22581703>
180. Obul J, Itoga S, Abliz M, Sato K, Ishige T, Utsuno E, et al. High-Resolution Melting Analyses for Gene Scanning of APC, MLH1, MSH2, and MSH6 Associated with Hereditary Colorectal Cancer. *Genet Test Mol Biomarkers.* 2012 May;16(5):406–11.
181. Bonis PA, Trikalinos TA, Chung M, Chew P, Ip S, DeVine DA, et al. Hereditary nonpolyposis colorectal cancer: diagnostic strategies and their implications. *Evid Rep Technol Assess (Full Rep).* 2007 May;(150):1–180.
182. Desai DC, Murday V, Phillips RK, Neale KF, Milla P, Hodgson SV. A survey of phenotypic features in juvenile polyposis. *J. Med. Genet.* 1998 Jun;35(6):476–81.
183. Howe JR, Roth S, Ringold JC, Summers RW, Järvinen HJ, Sistonen P, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science.* 1998 May 15;280(5366):1086–8.
184. Freeman TJ, Smith JJ, Chen X, Washington MK, Roland JT, Means AL, et al. Smad4-mediated signaling inhibits intestinal neoplasia by inhibiting expression of β -catenin. *Gastroenterology.* 2012 Mar;142(3):562–571.e2.
185. Tenière P, Songne K, Frebourg T, Moguelet P, Foulatier O, Michot F, et al. [Juvenile polyposis coli. The usefulness of a genetic study and the role of surgical treatment]. *Gastroenterol. Clin. Biol.* 2002 Nov;26(11):1047–50.
186. Itzkowitz SH, Yio X. Inflammation and cancer IV. Colorectal cancer in inflammatory bowel disease: the role of inflammation. *Am. J. Physiol. Gastrointest. Liver Physiol.* 2004 Jul;287(1):G7–17.
187. Shacter E, Weitzman SA. Chronic inflammation and cancer. *Oncology (Williston Park, N.Y.).* 2002 Feb;16(2):217–226, 229; discussion 230–232.
188. Lowe DB, Storkus WJ. Chronic inflammation and immunologic-based constraints in malignant disease. *Immunotherapy.* 2011 Oct;3(10):1265–74.
189. Sethi G, Shanmugam MK, Ramachandran L, Kumar AP, Tergaonkar V. Multifaceted link between cancer and inflammation. *Biosci. Rep.* 2012 Feb;32(1):1–15.
190. Bernstein CN, Blanchard JF, Kliewer E, Wajda A. Cancer risk in patients with inflammatory bowel disease: a population-based study. *Cancer.* 2001 Feb 15;91(4):854–62.
191. Fabre E, Spano JP, Atlan D, Braud AC, Mitry E, Panis Y, et al. [Cancer of the colon: an update]. *Bull Cancer.* 2000 Sep;Suppl 4:5–20.
192. Norat T, Bingham S, Ferrari P, Slimani N, Jenab M, Mazuir M, et al. Meat, fish, and colorectal cancer risk: the European Prospective Investigation into cancer and nutrition. *J. Natl. Cancer Inst.* 2005 Jun 15;97(12):906–16.
193. Huxley RR, Ansary-Moghaddam A, Clifton P, Czernichow S, Parr CL, Woodward M. The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. *Int. J. Cancer.* 2009 Jul 1;125(1):171–80.
194. Pretlow TP, Barrow BJ, Ashton WS, O’Riordan MA, Pretlow TG, Jurgisek JA, et al. Aberrant crypts: putative preneoplastic foci in human colonic mucosa. *Cancer Res.* 1991 Mar 1;51(5):1564–7.
195. Park DH, Kim HS, Kim WH, Kim TI, Kim YH, Park DI, et al. Clinicopathologic characteristics and malignant potential of colorectal flat neoplasia compared with that of polypoid neoplasia. *Dis. Colon Rectum.* 2008 Jan;51(1):43–49; discussion 49.

196. Whibley C, Pharoah PDP, Hollstein M. p53 polymorphisms: cancer implications. *Nat. Rev. Cancer.* 2009 Feb;9(2):95–107.
197. Schmitt CA, Fridman JS, Yang M, Lee S, Baranov E, Hoffman RM, et al. A senescence program controlled by p53 and p16INK4a contributes to the outcome of cancer therapy. *Cell.* 2002 May 3;109(3):335–46.
198. Fridman JS, Lowe SW. Control of apoptosis by p53. *Oncogene.* 2003 Dec 8;22(56):9030–40.
199. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J.* 1997 Jul 1;16(13):3797–804.
200. Behrens J, von Kries JP, Kühl M, Bruhn L, Wedlich D, Grosschedl R, et al. Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature.* 1996 Aug 15;382(6592):638–42.
201. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. *Nature.* 2005 Jun 9;435(7043):839–43.
202. Vafa O, Wade M, Kern S, Beeche M, Pandita TK, Hampton GM, et al. c-Myc can induce DNA damage, increase reactive oxygen species, and mitigate p53 function: a mechanism for oncogene-induced genetic instability. *Mol. Cell.* 2002 May;9(5):1031–44.
203. Wu S, Morin PJ, Maouyo D, Sears CL. *Bacteroides fragilis* enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology.* 2003 Feb;124(2):392–400.
204. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, et al. Identification of c-MYC as a target of the APC pathway. *Science.* 1998 Sep 4;281(5382):1509–12.
205. He TC, Chan TA, Vogelstein B, Kinzler KW. PPARdelta is an APC-regulated target of nonsteroidal anti-inflammatory drugs. *Cell.* 1999 Oct 29;99(3):335–45.
206. Kaplan KB, Burds AA, Swedlow JR, Bekir SS, Sorger PK, Näthke IS. A role for the Adenomatous Polyposis Coli protein in chromosome segregation. *Nat. Cell Biol.* 2001 Apr;3(4):429–32.
207. Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. *Nat. Rev. Cancer.* 2001 Oct;1(1):55–67.
208. Bienz M, Clevers H. Linking colorectal cancer to Wnt signaling. *Cell.* 2000 Oct 13;103(2):311–20.
209. Berthon A, Martinez A, Bertherat J, Val P. Wnt/ β -catenin signalling in adrenal physiology and tumour development. *Mol. Cell. Endocrinol.* 2012 Mar 31;351(1):87–95.
210. Sastre-Perona A, Santisteban P. Role of the wnt pathway in thyroid cancer. *Front Endocrinol (Lausanne).* 2012;3:31.
211. Samowitz WS, Powers MD, Spirio LN, Nollet F, van Roy F, Slattery ML. Beta-catenin mutations are more frequent in small colorectal adenomas than in larger adenomas and invasive carcinomas. *Cancer Res.* 1999 Apr 1;59(7):1442–4.
212. Shimizu Y, Ikeda S, Fujimori M, Kodama S, Nakahara M, Okajima M, et al. Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. *Genes Chromosomes Cancer.* 2002 Jan;33(1):73–81.
213. Schmierer B, Hill CS. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* 2007 Dec;8(12):970–82.
214. Grady WM, Myeroff LL, Swinler SE, Rajput A, Thiagalingam S, Lutterbaugh JD, et al. Mutational inactivation of transforming growth factor beta receptor type II in microsatellite stable colon cancers. *Cancer Res.* 1999 Jan 15;59(2):320–4.
215. McCubrey JA, Steelman LS, Chappell WH, Abrams SL, Wong EWT, Chang F, et al. Roles of the

- Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. *Biochim. Biophys. Acta.* 2007 Aug;1773(8):1263–84.
216. Roberts PJ, Der CJ. Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene.* 2007 May 14;26(22):3291–310.
 217. Lowe SW, Cepero E, Evan G. Intrinsic tumour suppression. *Nature.* 2004 Nov 17;432(7015):307–15.
 218. Samuels Y, Ericson K. Oncogenic PI3K and its role in cancer. *Curr Opin Oncol.* 2006 Jan;18(1):77–82.
 219. Nomura M, Shimizu S, Sugiyama T, Narita M, Ito T, Matsuda H, et al. 14-3-3 Interacts directly with and negatively regulates pro-apoptotic Bax. *J. Biol. Chem.* 2003 Jan 17;278(3):2058–65.
 220. Loeb LA. Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* 1991 Jun 15;51(12):3075–9.
 221. Hoeijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature.* 2001 May 17;411(6835):366–74.
 222. Choi S-W, Lee KJ, Bae Y-A, Min K-O, Kwon M-S, Kim K-M, et al. Genetic Classification of Colorectal Cancer Based on Chromosomal Loss and Microsatellite Instability Predicts Survival. *Clin Cancer Res.* 2002 Jan 7;8(7):2311–22.
 223. Barber TD, McManus K, Yuen KWY, Reis M, Parmigiani G, Shen D, et al. Chromatid cohesion defects may underlie chromosome instability in human colorectal cancers. *PNAS.* 2008 Apr 3;105(9):3443–8.
 224. Pérez-Carbonell L, Alenda C, Payá A, Castillejo A, Barberá VM, Guillén C, et al. Methylation Analysis of MLH1 Improves the Selection of Patients for Genetic Testing in Lynch Syndrome. *The Journal of Molecular Diagnostics.* 2010 Jul;12(4):498–504.
 225. Vilar E, Gruber SB. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol.* 2010 Mar;7(3):153–62.
 226. de Fraipont F, Richard M-J. L'hyperméthylation des gènes suppresseurs de tumeur comme marqueur en cancérologie. *Immuno-analyse & Biologie Spécialisée.* 2009 Feb;24(1):9–15.
 227. Deltour S, Chopin V, Leprince D. [Epigenetics and cancer]. *Med Sci (Paris).* 2005 Apr;21(4):405–11.
 228. Robertson KD, A.Jones P. DNA methylation: past, present and future directions. *Carcinogenesis.* 2000 Jan 3;21(3):461–7.
 229. Draht MXG, Riedl RR, Niessen H, Carvalho B, Meijer GA, Herman JG, et al. Promoter CpG island methylation markers in colorectal cancer: the road ahead. *Epigenomics.* 2012 Apr;4(2):179–94.
 230. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.* 1997 Mar 1;57(5):808–11.
 231. Toyota M, Ahuja N, Suzuki H, Itoh F, Ohe-Toyota M, Imai K, et al. Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. *Cancer Res.* 1999 Nov 1;59(21):5438–42.
 232. Nosho K, Shima K, Irahara N, Kure S, Baba Y, Kirkner GJ, et al. DNMT3B expression might contribute to CpG island methylator phenotype in colorectal cancer. *Clin. Cancer Res.* 2009 Jun 1;15(11):3663–71.
 233. Wong JJJ, Hawkins NJ, Ward RL. Colorectal cancer: a model for epigenetic tumorigenesis. *Gut.* 2007 Jan;56(1):140–8.
 234. Amato RJ, Stephenson J, Hotte S, Nemunaitis J, Bélanger K, Reid G, et al. MG98, a second-generation DNMT1 inhibitor, in the treatment of advanced renal cell carcinoma. *Cancer Invest.* 2012 Jun;30(5):415–21.

235. Kim HJ, Kim JH, Chie EK, Da Young P, Kim IA, Kim IH. DNMT (DNA methyltransferase) inhibitors radiosensitize human cancer cells by suppressing DNA repair activity. *Radiat Oncol.* 2012;7:39.
236. Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell.* 1993 Dec 3;75(5):843–54.
237. Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. *Nat. Rev. Mol. Cell Biol.* 2005 May;6(5):376–85.
238. Taganov KD, Boldin MP, Chang K-J, Baltimore D. NF-kappaB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. *Proc. Natl. Acad. Sci. U.S.A.* 2006 Aug 15;103(33):12481–6.
239. Krutovskikh VA, Herceg Z. Oncogenic microRNAs (OncomiRs) as a new class of cancer biomarkers. *Bioessays.* 2010 Oct;32(10):894–904.
240. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012 Jul 19;487(7407):330–7.
241. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA.* 2008 Jan 30;299(4):425–36.
242. Sarver AL, French AJ, Borralho PM, Thayanithy V, Oberg AL, Silverstein KAT, et al. Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer.* 2009;9:401.
243. Wu WKK, Law PTY, Lee CW, Cho CH, Fan D, Wu K, et al. MicroRNA in colorectal cancer: from benchtop to bedside. *Carcinogenesis.* 2011 Mar;32(3):247–53.
244. Philip M, Rowley DA, Schreiber H. Inflammation as a tumor promoter in cancer induction. *Semin. Cancer Biol.* 2004 Dec;14(6):433–9.
245. Luo J-L, Maeda S, Hsu L-C, Yagita H, Karin M. Inhibition of NF-kappaB in cancer cells converts inflammation- induced tumor growth mediated by TNFalpha to TRAIL-mediated tumor regression. *Cancer Cell.* 2004 Sep;6(3):297–305.
246. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature.* 2008 Jul 24;454(7203):436–44.
247. Karin M, Greten FR. NF-kappaB: linking inflammation and immunity to cancer development and progression. *Nat. Rev. Immunol.* 2005 Oct;5(10):749–59.
248. Balkwill F, Mantovani A. Inflammation and cancer: back to Virchow? *Lancet.* 2001 Feb 17;357(9255):539–45.
249. Coussens LM, Werb Z. Inflammation and cancer. *Nature.* 2002 Dec 19;420(6917):860–7.
250. Sparmann A, Bar-Sagi D. Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis. *Cancer Cell.* 2004 Nov;6(5):447–58.
251. Sumimoto H, Imabayashi F, Iwata T, Kawakami Y. The BRAF-MAPK signaling pathway is essential for cancer-immune evasion in human melanoma cells. *J. Exp. Med.* 2006 Jul 10;203(7):1651–6.
252. Shchors K, Shchors E, Rostker F, Lawlor ER, Brown-Swigart L, Evan GI. The Myc-dependent angiogenic switch in tumors is mediated by interleukin 1beta. *Genes Dev.* 2006 Sep 15;20(18):2527–38.
253. Brier B, Moses HL. TGF-beta and cancer. *Cytokine Growth Factor Rev.* 2006 Apr;17(1-2):29–40.
254. Greten FR, Eckmann L, Greten TF, Park JM, Li Z-W, Egan LJ, et al. IKKbeta links inflammation and

- tumorigenesis in a mouse model of colitis-associated cancer. *Cell*. 2004 Aug 6;118(3):285–96.
255. Pikarsky E, Porat RM, Stein I, Abramovitch R, Amit S, Kasem S, et al. NF-kappaB functions as a tumour promoter in inflammation-associated cancer. *Nature*. 2004 Sep 23;431(7007):461–6.
 256. Karin M, Cao Y, Greten FR, Li Z-W. NF-kappaB in cancer: from innocent bystander to major culprit. *Nat. Rev. Cancer*. 2002 Apr;2(4):301–10.
 257. Kishimoto T. Interleukin-6: from basic science to medicine--40 years in immunology. *Annu. Rev. Immunol.* 2005;23:1–21.
 258. Schneider MR, Hoeflich A, Fischer JR, Wolf E, Sordat B, Lahm H. Interleukin-6 stimulates clonogenic growth of primary and metastatic human colon carcinoma cells. *Cancer Lett*. 2000 Apr 3;151(1):31–8.
 259. Hsu C-P, Chung Y-C. Influence of interleukin-6 on the invasiveness of human colorectal carcinoma. *Anticancer Res*. 2006 Dec;26(6B):4607–14.
 260. Brozek W, Bises G, Girsch T, Cross HS, Kaiser HE, Peterlik M. Differentiation-dependent expression and mitogenic action of interleukin-6 in human colon carcinoma cells: relevance for tumour progression. *Eur. J. Cancer*. 2005 Oct;41(15):2347–54.
 261. Grivennikov S, Karin M. Autocrine IL-6 signaling: a key event in tumorigenesis? *Cancer Cell*. 2008 Jan;13(1):7–9.
 262. Grivennikov S, Karin E, Terzic J, Mucida D, Yu G-Y, Vallabhapurapu S, et al. IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer. *Cancer Cell*. 2009 Feb 3;15(2):103–13.
 263. Popivanova BK, Kostadinova FI, Furuichi K, Shamekh MM, Kondo T, Wada T, et al. Blockade of a chemokine, CCL2, reduces chronic colitis-associated carcinogenesis in mice. *Cancer Res*. 2009 Oct 1;69(19):7884–92.
 264. Schimanski CC, Schwald S, Simiantonaki N, Jayasinghe C, Gönner U, Wilsberg V, et al. Effect of chemokine receptors CXCR4 and CCR7 on the metastatic behavior of human colorectal cancer. *Clin. Cancer Res*. 2005 Mar 1;11(5):1743–50.
 265. Kim J, Takeuchi H, Lam ST, Turner RR, Wang H-J, Kuo C, et al. Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. *J. Clin. Oncol*. 2005 Apr 20;23(12):2744–53.
 266. Iglesias D, Nejda N, Azcoita MM, Schwartz S Jr, González-Aguilera JJ, Fernández-Peralta AM. Effect of COX2 -765G>C and c.3618A>G polymorphisms on the risk and survival of sporadic colorectal cancer. *Cancer Causes Control*. 2009 Oct;20(8):1421–9.
 267. Arber N, Eagle CJ, Spicak J, Rácz I, Dite P, Hajer J, et al. Celecoxib for the prevention of colorectal adenomatous polyps. *N. Engl. J. Med*. 2006 Aug 31;355(9):885–95.
 268. Benelli R. Aspirin, COX-2, and the risk of colorectal cancer. *N. Engl. J. Med*. 2007 Aug 23;357(8):824–825; author reply 824–825.
 269. Pereira C, Medeiros RM, Dinis-Ribeiro MJ. Cyclooxygenase polymorphisms in gastric and colorectal carcinogenesis: are conclusive results available? *Eur J Gastroenterol Hepatol*. 2009 Jan;21(1):76–91.
 270. Slattery ML, Herrick JS, Bondurant KL, Wolff RK. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. *Int. J. Cancer*. 2012 Jun 15;130(12):2974–80.
 271. Rakoff-Nahoum S, Medzhitov R. Regulation of spontaneous intestinal tumorigenesis through the adaptor protein MyD88. *Science*. 2007 Jul 6;317(5834):124–7.
 272. Huang B, Zhao J, Li H, He K-L, Chen Y, Chen S-H, et al. Toll-like receptors on tumor cells facilitate evasion

- of immune surveillance. *Cancer Res.* 2005 Jun 15;65(12):5009–14.
273. Balkwill F. Tumour necrosis factor and cancer. *Nat. Rev. Cancer.* 2009 May;9(5):361–71.
274. Popivanova BK, Kitamura K, Wu Y, Kondo T, Kagaya T, Kaneko S, et al. Blocking TNF- α in mice reduces colorectal carcinogenesis associated with chronic colitis. *J. Clin. Invest.* 2008 Feb;118(2):560–70.
275. Tudek B, Speina E. Oxidatively damaged DNA and its repair in colon carcinogenesis. *Mutation research* [Internet]. 2012 Apr 25 [cited 2012 May 30]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22561673>
276. NODA Noriko, WAKASUGI Hiro. *Cancer and Oxidative Stress.* 2001;JMAJ 44(12):535–9.
277. Noguchi M, Urushibara A, Yokoya A, O'Neill P, Shikazono N. The mutagenic potential of 8-oxoG/single strand break-containing clusters depends on their relative positions. *Mutat. Res.* 2012 Apr 1;732(1-2):34–42.
278. Seril DN, Liao J, Yang G-Y, Yang CS. Oxidative stress and ulcerative colitis-associated carcinogenesis: studies in humans and animal models. *Carcinogenesis.* 2003 Mar;24(3):353–62.
279. Huycke MM, Gaskins HR. Commensal bacteria, redox stress, and colorectal cancer: mechanisms and models. *Exp. Biol. Med. (Maywood).* 2004 Jul;229(7):586–97.
280. Marchesi JR, Dutilh BE, Hall N, Peters WHM, Roelofs R, Boleij A, et al. Towards the human colorectal cancer microbiome. *PLoS ONE.* 2011;6(5):e20447.
281. Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res.* 2012 Feb;22(2):292–8.
282. Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res.* 2012 Feb;22(2):299–306.
283. Bates RC, Bellovin DI, Brown C, Maynard E, Wu B, Kawakatsu H, et al. Transcriptional activation of integrin β 6 during the epithelial-mesenchymal transition defines a novel prognostic indicator of aggressive colon carcinoma. *J. Clin. Invest.* 2005 Feb;115(2):339–47.
284. Heine M, Nollau P, Masslo C, Nielsen P, Freund B, Bruns OT, et al. Investigations on the Usefulness of CEACAMs as Potential Imaging Targets for Molecular Imaging Purposes. *PLoS ONE.* 2011 Dec 5;6(12):e28030.
285. Niu J, Dorahy DJ, Gu X, Scott RJ, Draganic B, Ahmed N, et al. Integrin expression in colon cancer cells is regulated by the cytoplasmic domain of the β 6 integrin subunit. *Int. J. Cancer.* 2002 Jun 1;99(4):529–37.
286. Pizarro-Cerdá J, Cossart P. Bacterial adhesion and entry into host cells. *Cell.* 2006 Feb 24;124(4):715–27.
287. Roda G, Dahan S, Mezzanotte L, Caponi A, Roth-Walter F, Pinn D, et al. Defect in CEACAM family member expression in Crohn's disease IECs is regulated by the transcription factor SOX9. *Inflammatory Bowel Diseases.* 2009 Dec;15(12):1775–83.
288. Voges M, Bachmann V, Kammerer R, Gophna U, Hauck CR. CEACAM1 recognition by bacterial pathogens is species-specific. *BMC Microbiology.* 2010 Apr 20;10(1):117.
289. Berger CN, Billker O, Meyer TF, Servin AL, Kansau I. Differential recognition of members of the carcinoembryonic antigen family by Afa/Dr adhesins of diffusely adhering *Escherichia coli* (Afa/Dr DAEC). *Mol. Microbiol.* 2004 May;52(4):963–83.
290. Mimouna S, Gonçalves D, Barnich N, Darfeuille-Michaud A, Hofman P, Vouret-Craviari V. Crohn disease-associated *Escherichia coli* promote gastrointestinal inflammatory disorders by activation of HIF-dependent responses. *Gut microbes.* 2011 Nov 1.
291. Hauck CR. Cell adhesion receptors - signaling capacity and exploitation by bacterial pathogens. *Med. Microbiol. Immunol.* 2002 Oct;191(2):55–62.

292. Huycke MM, Abrams V, Moore DR. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. *Carcinogenesis*. 2002 Mar;23(3):529–36.
293. zur Hausen H. Streptococcus bovis: causal or incidental involvement in cancer of the colon? *Int. J. Cancer*. 2006 Nov 1;119(9):xi–xii.
294. Abdulmir AS, Hafidh RR, Abu Bakar F. The association of Streptococcus bovis/galloyticus with colorectal tumors: the nature and the underlying mechanisms of its etiological role. *J. Exp. Clin. Cancer Res*. 2011;30:11.
295. Wu S, Lim KC, Huang J, Saidi RF, Sears CL. Bacteroides fragilis enterotoxin cleaves the zonula adherens protein, E-cadherin. *Proc. Natl. Acad. Sci. U.S.A.* 1998 Dec 8;95(25):14979–84.
296. Scanlan PD, Shanahan F, Clune Y, Collins JK, O'Sullivan GC, O'Riordan M, et al. Culture-independent analysis of the gut microbiota in colorectal cancer and polyposis. *Environ. Microbiol.* 2008 Mar;10(3):789–98.
297. O'Keefe SJD. Nutrition and colonic health: the critical role of the microbiota. *Curr. Opin. Gastroenterol.* 2008 Jan;24(1):51–8.
298. Obe G, Anderson D. International Commission for Protection against Environmental Mutagens and Carcinogens. ICPEMC Working Paper No. 15/1. Genetic effects of ethanol. *Mutat. Res.* 1987 Nov;186(3):177–200.
299. Levitt MD, Furne J, Springfield J, Suarez F, DeMaster E. Detoxification of hydrogen sulfide and methanethiol in the cecal mucosa. *J. Clin. Invest.* 1999 Oct;104(8):1107–14.
300. Ridlon JM, Kang D-J, Hylemon PB. Bile salt biotransformations by human intestinal bacteria. *J. Lipid Res.* 2006 Feb;47(2):241–59.
301. Costarelli V, Key TJ, Appleby PN, Allen DS, Fentiman IS, Sanders TAB. A prospective study of serum bile acid concentrations and colorectal cancer risk in post-menopausal women on the island of Guernsey. *Br. J. Cancer*. 2002 Jun 5;86(11):1741–4.
302. Bernstein H, Bernstein C, Payne CM, Dvorakova K, Garewal H. Bile acids as carcinogens in human gastrointestinal cancers. *Mutat. Res.* 2005 Jan;589(1):47–65.
303. Uronis JM, Mühlbauer M, Herfarth HH, Rubinas TC, Jones GS, Jobin C. Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. *PLoS ONE*. 2009;4(6):e6026.
304. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U.S.A.* 2008 Oct 28;105(43):16731–6.
305. Engle SJ, Ormsby I, Pawlowski S, Boivin GP, Croft J, Balish E, et al. Elimination of Colon Cancer in Germ-free Transforming Growth Factor Beta 1-deficient Mice. *Cancer Res*. 2002 Nov 15;62(22):6362–6.
306. Souza DG, Vieira AT, Soares AC, Pinho V, Nicoli JR, Vieira LQ, et al. The essential role of the intestinal microbiota in facilitating acute inflammatory responses. *J. Immunol.* 2004 Sep 15;173(6):4137–46.
307. Balish E, Warner T. Enterococcus faecalis induces inflammatory bowel disease in interleukin-10 knockout mice. *Am. J. Pathol.* 2002 Jun;160(6):2253–7.
308. Elson CO, Cong Y, McCracken VJ, Dimmitt RA, Lorenz RG, Weaver CT. Experimental models of inflammatory bowel disease reveal innate, adaptive, and regulatory mechanisms of host dialogue with the microbiota. *Immunol. Rev.* 2005 Aug;206:260–76.
309. Steinberg D, Naggar CZ. Streptococcus bovis endocarditis with carcinoma of the colon. *N. Engl. J. Med.* 1977 Dec 15;297(24):1354–5.

310. Abdulmir AS, Hafidh RR, Mahdi LK, Al-jeboori T, Abubaker F. Investigation into the controversial association of *Streptococcus gallolyticus* with colorectal cancer and adenoma. *BMC Cancer*. 2009;9:403.
311. Boleij A, Schaeps RMJ, Tjalsma H. Association between *Streptococcus bovis* and colon cancer. *J. Clin. Microbiol.* 2009 Feb;47(2):516.
312. Waisberg J, Matheus C de O, Pimenta J. Infectious endocarditis from *Streptococcus bovis* associated with colonic carcinoma: case report and literature review. *Arq Gastroenterol.* 2002 Sep;39(3):177–80.
313. Abdulmir AS, Hafidh RR, Bakar FA. Molecular detection, quantification, and isolation of *Streptococcus gallolyticus* bacteria colonizing colorectal tumors: inflammation-driven potential of carcinogenesis via IL-1, COX-2, and IL-8. *Mol. Cancer*. 2010;9:249.
314. Corredoira J, Alonso MP, Coira A, Casariego E, Arias C, Alonso D, et al. Characteristics of *Streptococcus bovis* endocarditis and its differences with *Streptococcus viridans* endocarditis. *Eur. J. Clin. Microbiol. Infect. Dis.* 2008 Apr;27(4):285–91.
315. Gold JS, Bayar S, Salem RR. Association of *Streptococcus bovis* bacteremia with colonic neoplasia and extracolonic malignancy. *Arch Surg.* 2004 Jul;139(7):760–5.
316. Alazmi W, Bustamante M, O'Loughlin C, Gonzalez J, Raskin JB. The association of *Streptococcus bovis* bacteremia and gastrointestinal diseases: a retrospective analysis. *Dig. Dis. Sci.* 2006 Apr;51(4):732–6.
317. Milbrandt E. A novel source of enterococcal endocarditis. *Clin Cardiol.* 1998 Feb;21(2):123–6.
318. Roses DF, Richman H, Localio SA. Bacterial endocarditis associated with colorectal carcinoma. *Annals of Surgery.* 1974 Feb;179(2):190.
319. Christakis GB, Perlorentzou SP, Aslanidou M, Savva L, Zarkadis IK. Bacteremia caused by *Pantoea agglomerans* and *Enterococcus faecalis* in a patient with colon cancer. *J BUON.* 2007 Jun;12(2):287–90.
320. Wang X, Allen TD, May RJ, Lightfoot S, Houchen CW, Huycke MM. *Enterococcus faecalis* induces aneuploidy and tetraploidy in colonic epithelial cells through a bystander effect. *Cancer Res.* 2008 Dec 1;68(23):9909–17.
321. Balamurugan R, Rajendiran E, George S, Samuel GV, Ramakrishna BS. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. *J. Gastroenterol. Hepatol.* 2008 Aug;23(8 Pt 1):1298–303.
322. Huycke MM, Moore DR. In vivo production of hydroxyl radical by *Enterococcus faecalis* colonizing the intestinal tract using aromatic hydroxylation. *Free Radic. Biol. Med.* 2002 Sep 15;33(6):818–26.
323. Wang X, Huycke MM. Extracellular superoxide production by *Enterococcus faecalis* promotes chromosomal instability in mammalian cells. *Gastroenterology.* 2007 Feb;132(2):551–61.
324. Sears CL. The toxins of *Bacteroides fragilis*. *Toxicon.* 2001 Nov;39(11):1737–46.
325. Toprak NU, Yagci A, Gulluoglu BM, Akin ML, Demirkalem P, Celenk T, et al. A possible role of *Bacteroides fragilis* enterotoxin in the aetiology of colorectal cancer. *Clin. Microbiol. Infect.* 2006 Aug;12(8):782–6.
326. Goodwin AC, Destefano Shields CE, Wu S, Huso DL, Wu X, Murray-Stewart TR, et al. Polyamine catabolism contributes to enterotoxigenic *Bacteroides fragilis*-induced colon tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* 2011 Sep 13;108(37):15354–9.
327. Åkerblom J. The role and genetic control of non-coding RNAs in bacterial infection. 2010.
328. Housseau F, Sears CL. Enterotoxigenic *Bacteroides fragilis* (ETBF)-mediated colitis in Min (*Apc*^{+/-}) mice: A human commensal-based murine model of colon carcinogenesis. *Cell cycle (Georgetown, Tex.).* 2010 Jan 1;9(1):3.

329. Obiso RJ Jr, Lyerly DM, Van Tassell RL, Wilkins TD. Proteolytic activity of the *Bacteroides fragilis* enterotoxin causes fluid secretion and intestinal damage in vivo. *Infect. Immun.* 1995 Oct;63(10):3820–6.
330. Sears CL, Myers LL, Lazenby A, Van Tassell RL. Enterotoxigenic *Bacteroides fragilis*. *Clin. Infect. Dis.* 1995 Jun;20 Suppl 2:S142–148.
331. Sears CL. Enterotoxigenic *Bacteroides fragilis*: a Rogue among Symbiotes. *Clinical Microbiology Reviews.* 2009 Apr 14;22(2):349–69.
332. Muranski P, Boni A, Antony PA, Cassard L, Irvine KR, Kaiser A, et al. Tumor-specific Th17-polarized cells eradicate large established melanoma. *Blood.* 2008 Jul 15;112(2):362–73.
333. Martin-Orozco N, Muranski P, Chung Y, Yang XO, Yamazaki T, Lu S, et al. T helper 17 cells promote cytotoxic T cell activation in tumor immunity. *Immunity.* 2009 Nov 20;31(5):787–98.
334. Swidsinski A, Khilkin M, Kerjaschki D, Schreiber S, Ortner M, Weber J, et al. Association between intraepithelial *Escherichia coli* and colorectal cancer. *Gastroenterology.* 1998 Aug;115(2):281–6.
335. Lax A. The *Pasteurella multocida* Toxin: A New Paradigm for the Link Between Bacterial Infection and Cancer. *Current topics in microbiology and immunology*. 2012 Jun 15.
336. Horiguchi Y, Nakai T, Kume K. Effects of *Bordetella bronchiseptica* dermonecrotic toxin on the structure and function of osteoblastic clone MC3T3-e1 cells. *Infect. Immun.* 1991 Mar;59(3):1112–6.
337. Churin Y, Al-Ghoul L, Kepp O, Meyer TF, Birchmeier W, Naumann M. *Helicobacter pylori* CagA protein targets the c-Met receptor and enhances the motogenic response. *J. Cell Biol.* 2003 Apr 28;161(2):249–55.
338. Caprioli A, Falbo V, Roda LG, Ruggeri FM, Zona C. Partial purification and characterization of an *Escherichia coli* toxic factor that induces morphological cell alterations. *Infect. Immun.* 1983 Mar;39(3):1300–6.
339. Knust Z, Schmidt G. Cytotoxic Necrotizing Factors (CNFs)-A Growing Toxin Family. *Toxins (Basel).* 2010 Jan;2(1):116–27.
340. Lax AJ. New genotoxin shows diversity of bacterial attack mechanisms. *Trends in Molecular Medicine.* 2007 Mar;13(3):91–3.
341. Boquet P. The cytotoxic necrotizing factor 1 (CNF1) from *Escherichia coli*. *Toxicon.* 2001 Nov;39(11):1673–80.
342. Etienne-Manneville S, Hall A. Rho GTPases in cell biology. *Nature.* 2002 Dec 12;420(6916):629–35.
343. Kadhum HJ, Finlay D, Rowe MT, Wilson IG, Ball HJ. Occurrence and characteristics of cytotoxic necrotizing factors, cytolethal distending toxins and other virulence factors in *Escherichia coli* from human blood and faecal samples. *Epidemiology and Infection*. 2007 Aug 3.
344. Fiorentini C, Matarrese P, Straface E, Falzano L, Fabbri A, Donelli G, et al. Toxin-induced activation of Rho GTP-binding protein increases Bcl-2 expression and influences mitochondrial homeostasis. *Exp. Cell Res.* 1998 Jul 10;242(1):341–50.
345. Thomas W, Ascott ZK, Harmey D, Slice LW, Rozengurt E, Lax AJ. Cytotoxic Necrotizing Factor from *Escherichia coli* Induces RhoA-Dependent Expression of the Cyclooxygenase-2 Gene. *Infection and Immunity.* 2001 Nov 1;69(11):6839–45.
346. Mager DL. Bacteria and cancer: cause, coincidence or cure? A review. *J Transl Med.* 2006;4:14.
347. Samaras V, Rafailidis PI, Mourtzoukou EG, Peppas G, Falagas ME. Chronic bacterial and parasitic infections and cancer: a review. *J Infect Dev Ctries.* 2010 May;4(5):267–81.
348. Fabbri A, Travaglione S, Fiorentini C. *Escherichia coli* Cytotoxic Necrotizing Factor 1 (CNF1): Toxin Biology,

in Vivo Applications and Therapeutic Potential. *Toxins* (Basel). 2010 Feb;2(2):283–96.

349. Dubois D, Delmas J, Cady A, Robin F, Sivignon A, Oswald E, et al. Cyclomodulins in Urosepsis Strains of *Escherichia coli*. *J Clin Microbiol*. 2010 Jun;48(6):2122–9.
350. Bronowski C, Smith SL, Yokota K, Corkill JE, Martin HM, Campbell BJ, et al. A subset of mucosa-associated *Escherichia coli* isolates from patients with colon cancer, but not Crohn's disease, share pathogenicity islands with urinary pathogenic *E. coli*. *Microbiology* (Reading, Engl.). 2008 Feb;154(Pt 2):571–83.
351. Guerra L, Cortes-Bratti X, Guidi R, Frisan T. The Biology of the Cytolethal Distending Toxins. *Toxins*. 2011 Mar 7;3(3):172–90.
352. Blazkova H, Krejcikova K, Moudry P, Frisan T, Hodny Z, Bartek J. Bacterial intoxication evokes cellular senescence with persistent DNA damage and cytokine signalling. *J. Cell. Mol. Med*. 2010 Jan;14(1-2):357–67.
353. Guerra L, Albiñ A, Tronnersjö S, Yan Q, Guidi R, Stenerlöw B, et al. Myc is required for activation of the ATM-dependent checkpoints in response to DNA damage. *PLoS ONE*. 2010;5(1):e8924.
354. Nagamine CM, Sohn JJ, Rickman BH, Rogers AB, Fox JG, Schauer DB. *Helicobacter hepaticus* infection promotes colon tumorigenesis in the BALB/c-Rag2(-/-) Apc(Min/+) mouse. *Infect. Immun*. 2008 Jun;76(6):2758–66.
355. Marchès O, Ledger TN, Boury M, Ohara M, Tu X, Goffaux F, et al. Enteropathogenic and enterohaemorrhagic *Escherichia coli* deliver a novel effector called Cif, which blocks cell cycle G2/M transition. *Mol. Microbiol*. 2003 Dec;50(5):1553–67.
356. Jubelin G, Chavez CV, Taieb F, Banfield MJ, Samba-Louaka A, Nobe R, et al. Cycle inhibiting factors (CIFs) are a growing family of functional cyclomodulins present in invertebrate and mammal bacterial pathogens. *PLoS ONE*. 2009;4(3):e4855.
357. Cui J, Yao Q, Li S, Ding X, Lu Q, Mao H, et al. Glutamine deamidation and dysfunction of ubiquitin/NEDD8 induced by a bacterial effector family. *Science*. 2010 Sep 3;329(5996):1215–8.
358. Homburg S, Oswald E, Hacker J, Dobrindt U. Expression analysis of the colibactin gene cluster coding for a novel polyketide in *Escherichia coli*. *FEMS Microbiol. Lett*. 2007 Oct;275(2):255–62.
359. Dubois D, Baron O, Cougnoux A, Delmas J, Pradel N, Boury M, et al. ClbP is the prototype of a peptidase subgroup involved in the biosynthesis of nonribosomal peptides. *J Biol Chem*. 2011 Jul 27.
360. Fischbach MA, Walsh CT. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chem. Rev*. 2006 Aug;106(8):3468–96.
361. Menzella HG, Reeves CD. Combinatorial biosynthesis for drug development. *Curr. Opin. Microbiol*. 2007 Jun;10(3):238–45.
362. Xu Y, Kersten RD, Nam S-J, Lu L, Al-Suwailem AM, Zheng H, et al. Bacterial Biosynthesis and Maturation of the Didemnin Anti-cancer Agents. *Journal of the American Chemical Society*. 2012 Apr 6.
363. Walsh CT, Gehring AM, Weinreb PH, Quadri LE, Flugel RS. Post-translational modification of polyketide and nonribosomal peptide synthases. *Curr Opin Chem Biol*. 1997 Oct;1(3):309–15.
364. Fischbach MA, Walsh CT. Antibiotics for emerging pathogens. *Science*. 2009 Aug 28;325(5944):1089–93.
365. Koglin A, Walsh CT. Structural insights into nonribosomal peptide enzymatic assembly lines. *Nat Prod Rep*. 2009 Aug;26(8):987–1000.
366. Walsh CT, Fischbach MA. Natural products version 2.0: connecting genes to molecules. *J. Am. Chem. Soc*. 2010 Mar 3;132(8):2469–93.

367. Broadhurst R.W., Nietlispach D., Wheatcroft M.P., Leadlay P.F., Weissman K.J. The Structure of Docking Domains in Modular Polyketide Synthases. *Chemistry & Biology*. 2003;10(8):723–31.
368. Yeh E, Lin H, Clugston SL, Kohli RM, Walsh CT. Enhanced Macrocyclizing Activity of the Thioesterase from Tyrocidine Synthetase in Presence of Nonionic Detergent. *Chemistry & Biology*. 2004 Nov;11(11):1573–82.
369. Putze J, Hennequin C, Nougayrède J-P, Zhang W, Homburg S, Karch H, et al. Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infect. Immun.* 2009 Nov;77(11):4696–703.
370. Fialkina SV, Bondarenko VM, Naboka IL, Miroshnichenko EA. [Revealing the genetic determinants of Pks-pathogenicity island in clinical strains of enterobacteria]. *Zh. Mikrobiol. Epidemiol. Immunobiol.* 2011 Oct; (5):3–7.
371. Hayashi T. Microbiology. Breaking the barrier between commensalism and pathogenicity. *Science*. 2006 Aug 11;313(5788):772–3.
372. Dubois D, Prasadarao NV, Mittal R, Bret L, Roujou-Gris M, Bonnet R. CTX-M beta-lactamase production and virulence of *Escherichia coli* K1. *Emerging Infect. Dis.* 2009 Dec;15(12):1988–90.
373. Johnson JR, Johnston B, Kuskowski MA, Nougayrede J-P, Oswald E. Molecular epidemiology and phylogenetic distribution of the *Escherichia coli* pks genomic island. *J. Clin. Microbiol.* 2008 Dec;46(12):3906–11.
374. Krieger JN, Dobrindt U, Riley DE, Oswald E. Acute *Escherichia coli* prostatitis in previously health young men: bacterial virulence factors, antimicrobial resistance, and clinical outcomes. *Urology*. 2011 Jun;77(6):1420–5.
375. Coppé J-P, Patil CK, Rodier F, Sun Y, Muñoz DP, Goldstein J, et al. Senescence-Associated Secretory Phenotypes Reveal Cell-Nonautonomous Functions of Oncogenic RAS and the p53 Tumor Suppressor. *PLoS Biol.* 2008 Dec 2;6(12):e301.
376. Kunz-Schughart LA, Knuechel R. Tumor-associated fibroblasts (part I): Active stromal participants in tumor development and progression? *Histol. Histopathol.* 2002 Apr;17(2):599–621.
377. Kunz-Schughart LA, Knuechel R. Tumor-associated fibroblasts (part II): Functional impact on tumor tissue. *Histol. Histopathol.* 2002 Apr;17(2):623–37.
378. Rawlings ND, Morton FR, Kok CY, Kong J, Barrett AJ. MEROPS: the peptidase database. *Nucleic Acids Res.* 2008 Jan;36(Database issue):D320–325.
379. Jacoby GA. AmpC beta-lactamases. *Clin. Microbiol. Rev.* 2009 Jan;22(1):161–182, Table of Contents.
380. Delfosse V, Girard E, Birck C, Delmarcelle M, Delarue M, Poch O, et al. Structure of the archaeal pab87 peptidase reveals a novel self-compartmentalizing protease family. *PLoS ONE*. 2009;4(3):e4712.
381. Goffin C, Ghuysen J-M. Biochemistry and Comparative Genomics of SxxK Superfamily Acyltransferases Offer a Clue to the Mycobacterial Paradox: Presence of Penicillin-Susceptible Target Proteins versus Lack of Efficiency of Penicillin as Therapeutic Agent. *Microbiol Mol Biol Rev.* 2002 Dec;66(4):702–38.
382. Chen Y, Minasov G, Roth TA, Prati F, Shoichet BK. The deacylation mechanism of AmpC beta-lactamase at ultrahigh resolution. *J. Am. Chem. Soc.* 2006 Mar 8;128(9):2970–6.
383. Kevany BM, Rasko DA, Thomas MG. Characterization of the complete zwittermicin A biosynthesis gene cluster from *Bacillus cereus*. *Appl. Environ. Microbiol.* 2009 Feb;75(4):1144–55.
384. Reimer D, Pos KM, Thines M, Grün P, Bode HB. A natural prodrug activation mechanism in nonribosomal peptide synthesis. *Nat. Chem. Biol.* 2011 Dec;7(12):888–90.
385. Baker SJ, Ding CZ, Akama T, Zhang Y-K, Hernandez V, Xia Y. Therapeutic potential of boron-containing

- compounds. *Future Med Chem.* 2009 Oct;1(7):1275–88.
386. Anacor - Boron Chemistry Platform Overview [Internet]. [cited 2012 Oct 12]. Available from: http://www.anacor.com/boron_technology_overview.php
387. Buzzoni V, Blazquez J, Ferrari S, Calò S, Venturelli A, Costi MP. Aza-boronic acids as non-beta-lactam inhibitors of AmpC-beta-lactamase. *Bioorg. Med. Chem. Lett.* 2004 Aug 2;14(15):3979–83.
388. Goodsell DS, Morris GM, Olson AJ. Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit.* 1996 Feb;9(1):1–5.
389. Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today.* 2004 May 15;9(10):430–1.
390. Usher KC, Blaszczyk LC, Weston GS, Shoichet BK, Remington SJ. Three-dimensional structure of AmpC beta-lactamase from *Escherichia coli* bound to a transition-state analogue: possible implications for the oxyanion hypothesis and for inhibitor design. *Biochemistry.* 1998 Nov 17;37(46):16082–92.
391. Stein U, Walther W, Arlt F, Schwabe H, Smith J, Fichtner I, et al. MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat. Med.* 2009 Jan;15(1):59–67.
392. Liu Y, Li Q, Zhu L. Expression of the hepatocyte growth factor and c-Met in colon cancer: correlation with clinicopathological features and overall survival. *Tumori.* 2012 Jan;98(1):105–12.
393. Sierra JR, Tsao M-S. c-MET as a potential therapeutic target and biomarker in cancer. *Ther Adv Med Oncol.* 2011 Nov;3(1 Suppl):S21–35.
394. Eder JP, Vande Woude GF, Boerner SA, LoRusso PM. Novel therapeutic inhibitors of the c-Met signaling pathway in cancer. *Clin. Cancer Res.* 2009 Apr 1;15(7):2207–14.
395. Davalos AR, Coppe J-P, Campisi J, Desprez P-Y. Senescent cells as a source of inflammatory factors for tumor progression. *Cancer Metastasis Rev.* 2010 Jun;29(2):273–83.
396. Rodier F, Campisi J. Four faces of cellular senescence. *J. Cell Biol.* 2011 Feb 21;192(4):547–56.
397. Singh R, George J, Shukla Y. Role of senescence and mitotic catastrophe in cancer therapy. *Cell Division.* 2010 Jan 21;5(1):4.
398. Hornsby PJ. Cellular aging and cancer. *Critical Reviews in Oncology/Hematology.* 2011 Aug;79(2):189–95.
399. Campisi J. Cellular senescence as a tumor-suppressor mechanism. *Trends Cell Biol.* 2001 Nov;11(11):S27–31.
400. Liu D, Hornsby PJ. Senescent Human Fibroblasts Increase the Early Growth of Xenograft Tumors via Matrix Metalloproteinase Secretion. *Cancer Res.* 2007 Jan 4;67(7):3117–26.
401. Quintar AA, Doll A, Leimgruber C, Palmeri CM, Roth FD, Maccioni M, et al. Acute inflammation promotes early cellular stimulation of the epithelial and stromal compartments of the rat prostate. *Prostate.* 2010 Aug;70(11):1153–65.
402. Redon CE, Dickey JS, Nakamura AJ, Kareva IG, Naf D, Newshean S, et al. Tumors induce complex DNA damage in distant proliferative tissues in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 2010 Oct 19;107(42):17992–7.
403. Pullen NA, Anand M, Cooper PS, Fillmore HL. Matrix metalloproteinase-1 expression enhances tumorigenicity as well as tumor-related angiogenesis and is inversely associated with TIMP-4 expression in a model of glioblastoma. *J. Neurooncol.* 2012 Feb;106(3):461–71.
404. Maciejewska-Rodrigues H, Karouzakis E, Strietholt S, Hemmatazad H, Neidhart M, Ospelt C, et al. Epigenetics and rheumatoid arthritis: the role of SENP1 in the regulation of MMP-1 expression. *J. Autoimmun.* 2010 Aug;35(1):15–22.

405. Prager GW, Poettler M. Angiogenesis in cancer. Basic mechanisms and therapeutic advances. *Hamostaseologie*. 2012;32(2):105–14.
406. Lu C, Han HD, Mangala LS, Ali-Fehmi R, Newton CS, Ozburn L, et al. Regulation of Tumor Angiogenesis by EZH2. *Cancer Cell*. 2010 Aug 17;18(2):185–97.
407. Gorantla B, Asuthkar S, Rao JS, Patel J, Gondi CS. Suppression of the uPAR–uPA System Retards Angiogenesis, Invasion, and In Vivo Tumor Development in Pancreatic Cancer Cells. *Mol Cancer Res*. 2011 Jan 4;9(4):377–89.
408. Rosso MD. uPAR in angiogenesis regulation. *Blood*. 2011 Apr 14;117(15):3941–3.
409. Koukourakis GV, Sotiropoulou-Lontou A. Targeted therapy with bevacizumab (Avastin) for metastatic colorectal cancer. *Clin Transl Oncol*. 2011 Oct;13(10):710–4.
410. Barascu A, Chalony CL, Pennarun G, Genet D, Imam N, Lopez B, et al. Oxidative stress induces an ATM-independent senescence pathway through p38 MAPK-mediated lamin B1 accumulation. *The EMBO Journal*. 2012 Jan 13.
411. Akar CA, Feinstein DL. Modulation of inducible nitric oxide synthase expression by sumoylation. *J Neuroinflammation*. 2009;6:12.
412. Ribet D, Hamon M, Gouin E, Nahori M-A, Impens F, Neyret-Kahn H, et al. *Listeria monocytogenes* impairs SUMOylation for efficient infection. *Nature*. 2010 Apr 22;464(7292):1192–5.
413. Orth K, Xu Z, Mudgett MB, Bao ZQ, Palmer LE, Bliska JB, et al. Disruption of Signaling by *Yersinia* Effector YopJ, a Ubiquitin-Like Protein Protease. *Science*. 2000 Nov 24;290(5496):1594–7.
414. Hotson A, Chosed R, Shu H, Orth K, Mudgett MB. *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins in planta. *Mol. Microbiol*. 2003 Oct;50(2):377–89.
415. Jeon YJ, Yoo HM, Chung CH. ISG15 and immune diseases. *Biochim. Biophys. Acta*. 2010 May;1802(5):485–96.
416. Marasa BS, Srikanthan S, Martindale JL, Kim MM, Lee EK, Gorospe M, et al. MicroRNA profiling in human diploid fibroblasts uncovers miR-519 role in replicative senescence. *Aging (Albany NY)*. 2010 Jun;2(6):333–43.
417. Shivshankar P, Boyd AR, Le Saux CJ, Yeh I-T, Orihuela CJ. Cellular senescence increases expression of bacterial ligands in the lungs and is positively correlated with increased susceptibility to pneumococcal pneumonia. *Aging Cell*. 2011 Oct;10(5):798–806.
418. Shivshankar P. Modulation of Bacterial Pathogenesis by Oppressive Aging Factors: Insights into Host-Pneumococcal Interaction Strategies. *ISRN Inflammation*. 2012;2012:1–11.
419. Boyd AR, Orihuela CJ. Dysregulated Inflammation as a Risk Factor for Pneumonia in the Elderly. *Aging and Disease*. 2011 Dec;2(6):487.
420. Plottel CS, Blaser MJ. Microbiome and malignancy. *Cell Host Microbe*. 2011 Oct 20;10(4):324–35.
421. Krouwer VJ, Hekking LH, Langelaar-Makkinje M, Regan-Klapisz E, Post JA. Endothelial cell senescence is associated with disrupted cell-cell junctions and increased monolayer permeability. *Vascular cell*. 2012 Aug 28;4(1):12.
422. McCOY WC, MASON JM 3rd. Enterococcal endocarditis associated with carcinoma of the sigmoid; report of a case. *J Med Assoc State Ala*. 1951 Dec;21(6):162–6.
423. McFarland LV. Alternative treatments for *Clostridium difficile* disease: what really works? *J Med Microbiol*. 2005 Jan 2;54(2):101–11.

V Annexe

Noncovalent complexes of an inactive mutant of CTX-M-9 with the substrate piperacillin and the corresponding product.

Leyssene D, Delmas J, Robin F, **Cougnoux A**, Gibold L, Bonnet R.

Antimicrob Agents Chemother. 2011 Dec;55(12):5660-5. Epub 2011 Sep 19.

Characterization of the inhibitor-resistant SHV β -lactamase SHV-107 in a clinical *Klebsiella pneumoniae* strain coproducing GES-7 enzyme.

Manageiro V, Ferreira E, **Cougnoux A**, Albuquerque L, Caniça M, Bonnet R.

Antimicrob Agents Chemother. 2012 Feb;56(2):1042-6. Epub 2011 Nov 14.

Noncovalent Complexes of an Inactive Mutant of CTX-M-9 with the Substrate Piperacillin and the Corresponding Product[∇]

David Leysse, ^{1,2} Julien Delmas, ^{1,2} Frédéric Robin, ^{1,2} Antony Cougnoux, ^{1,2}
Lucie Gibold, ^{1,2} and Richard Bonnet ^{1,2*}

CHU Clermont-Ferrand, Laboratoire de Bactériologie, Clermont-Ferrand, F-63003, France, ¹ and Clermont Université, Université d'Auvergne, JE 2526, Evolution des Bactéries Pathogènes et Susceptibilité de l'Hôte, BP 10448, F-63000 Clermont-Ferrand 1, France²

Received 7 March 2011/Returned for modification 17 April 2011/Accepted 9 September 2011

We determined the crystal structure of an inactive Ser70Gly mutant of CTX-M-9 in complex with the bulky penicillin piperacillin at precovalent and posthydrolytic stages in the catalytic process. The structures obtained at high resolution were compared with the corresponding structures for the small penicillin benzylpenicillin and the bulky cephalosporin cefotaxime. The findings highlight the key role of the configuration of the carbon adjacent to the acylamino group of the side chain of β -lactams in the precovalent recognition of substrates.

The production of β -lactamases is the predominant cause of resistance to β -lactam antibiotics in Gram-negative bacteria. These enzymes are divided on the basis of amino acid sequences into four classes designated A to D, with class A β -lactamases being the most prevalent in *Enterobacteriaceae* (1).

The oximino cephalosporins such as cefotaxime have been used since the 1980s to overcome these resistance mechanisms because of their relative stability to serine β -lactamases and their broad antibacterial spectrum. These β -lactams typically contain a bulky group such as a 2-(2-aminothiazole-4-yl)-2-oximino substituent at position C-7 of the cephalosporin nucleus, which prevents hydrolysis by serine β -lactamases. However, the intensive use of these molecules was quickly followed by the emergence of extended-spectrum β -lactamases (ESBLs), which can accommodate and hydrolyze such large β -lactams. The CTX-M enzymes of the class A family, first described in 1989, have spread worldwide since the late 1990s and are now the most prevalent extended-spectrum β -lactamases (2, 3, 5, 6, 25, 29).

The hydrolysis process catalyzed by these enzymes is based on an acid-base catalytic mechanism, as for the other class A enzymes. It comprises three major steps (8, 9, 14, 19): the formation of a precovalent encounter complex with the substrate; the nucleophilic attack on the β -lactam ring by the active serine residue resulting in an acyl-enzyme; and the nucleophilic attack on this covalent complex by the catalytic water molecule, which produces a transient noncovalent product-enzyme complex and subsequently results in the release of both the product and the active enzyme.

Several studies have looked at the crystal structure of CTX-M enzymes in complexes with β -lactams or transition-state analogs (8, 10, 11, 26). Surprisingly, CTX-M-9 efficiently binds and hydrolyzes the large oximino cephalosporin cefotaxime despite its small active site. Recently, the structures of benzylpenicillin and cefotaxime CTX-M-9 Michaelis com-

plexes revealed that the accommodation of the bulky cephalosporin cefotaxime can be accommodated by a widening of the catalytic pocket, which is not observed for the small molecule benzylpenicillin (10). In this work, we determined the crystal structure of an inactive Ser70Gly mutant of CTX-M-9 in complex with the hydrolyzed and nonhydrolyzed forms of piperacillin to investigate the recognition of this bulky penicillin by CTX-M-type ESBLs. The structure shows that the size of the substrate is not the exclusive determinant in the widening of the catalytic pocket.

MATERIALS AND METHODS

Enzyme expression and purification. The replacement of the active serine (Ser70) by an inactive Gly residue was performed in CTX-M-9 by site-directed mutagenesis using a CTX-M-9-producing pET-9a-based expression vector (11). DNA sequence was checked by gene sequencing. The mutated protein, designated CTX-M-9-S70G, was expressed in *Escherichia coli* BL21(DE3) (Novagene) as previously described (7). Briefly, *E. coli* was inoculated in 2 \times YT medium (Obiogene, Irvine, CA) supplemented with 30 μ g/ml kanamycin for aerobic growth at 37°C up to an A_{600} of 0.8. Expression was induced for 36 h at 30°C with 0.2 mM isopropyl- β -D-thiogalactopyranoside (Sigma Chemical Co., St. Louis, MO). Bacteria were harvested with 20 mM morpholineethanesulfonic acid (MES)-NaOH (pH 6.0). The cell pellet was disrupted by one cycle of freezing and thawing, followed by sonication. After centrifugation (10,000 \times g for 10 min and 18,000 \times g for 60 min at 4°C), the clarified supernatant was loaded onto an ion-exchange CM-Fast Flow column (100 ml; Amersham Pharmacia Biotech, Orsay, France) equilibrated with 20 mM MES-NaOH (pH 6.0). Proteins were eluted with a linear NaCl gradient (0 to 150 mM). The β -lactamase-containing elution peak was loaded on a gel filtration Superose 12 column (25 ml; Amersham Pharmacia Biotech) and eluted with a 5 mM Tris-HCl (pH 7.0) and 50 mM NaCl buffer. The protein was concentrated by ultrafiltration to 10 mg/ml for crystallization. The protein concentration was estimated by the Bio-Rad protein assay (Bio-Rad, Richmond, CA). Homogeneity was estimated to be more than 98% by SDS-PAGE.

Crystallization. Crystals of CTX-M-9-S70G were obtained by vapor diffusion in hanging drops, using microseeding techniques with CTX-M-9 crystals. To a solution of 10 mg/ml protein in 5 mM Tris-HCl (pH 7.0) and 50 mM NaCl was added an equal volume of 1.2 M potassium phosphate buffer (pH 8.2) as previously described (10). Crystals appeared within 24 to 48 h after equilibration at 20°C and were soaked overnight at 20°C with hydrolyzed and native β -lactams at a concentration of 50 mM in 1.2 M potassium phosphate buffer (pH 8.2). Piperacillin was purchased from Dakota Pharm. Hydrolyzed piperacillin was obtained by 24 h of incubation with the wild-type CTX-M-9 at 30°C. Crystals were cryoprotected with 30% sucrose in 1.2 M potassium phosphate (pH 8.2) and flash frozen in liquid nitrogen.

* Corresponding author. Mailing address: Laboratoire de Bactériologie, 28 place H. Dunant, 63001 Clermont-Ferrand, France. Phone: 33 4 73 75 49 20. Fax: 33 4 73 75 49 22. E-mail: rbonnet@chu-clermontferrand.fr.

[∇] Published ahead of print on 19 September 2011.

TABLE 1. Data collection and refinement statistics^a

Parameter	CTX-M-9 S70G	CTX-M-9 S70G/piperacillin	CTX-M-9 S70G/hydrolyzed piperacillin
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a, b, c (Å)	45.07, 106.29, 47.54	45.13, 106.34, 47.61	45.10, 106.61, 47.54
α, β, γ (°)	90.00, 102.17, 90.00	90.00, 101.90, 90.00	90.00, 101.92, 90.00
R_{merge}	0.047 (0.180)	0.048 (0.950)	0.064 (0.334)
$I/\sigma I$	22.4 (5.4)	14.7 (5.9)	16.6 (3.5)
Completeness (%)	96.2 (91)	94.5 (90.1)	99.8 (98.4)
Redundancy	4.9 (4.8)	5.0	4.9
Resolution (Å)	42.6–1.5	19.47–1.50	19.20–1.50
No. of reflections	74,760	62,848	66,447
$R_{\text{work}}/R_{\text{free}}$	0.156/0.184	0.145/0.175	0.144/0.176
RMSD			
Bond length (Å)	0.008	0.010	0.011
Bond angle (°)	1.536	1.623	1.287

^a Values in parentheses are for the highest-resolution shell. One crystal was used for each data set.

Data collection and structure determination. Data were collected on beam line ID23-1 of the European Synchrotron Radiation Facility at Grenoble, France. Reflections were indexed, integrated, and scaled using the HKL software package (21). The space group of crystals was $P2_1$, with two molecules in the asymmetric unit. Phases were calculated by molecular replacement with the MolRep program (28) of the CCP4 package (24), using the apoenzyme structure of CTX-M-9 (9) with water molecules and ions removed. Marvin version 5.3.8 was used for drawing compound structures (ChemAxon). The structures were refined with program Refmac5 (27) of the CCP4 package and Coot (12). Cross-validation was used throughout, and 5% of the data were used for the R_{free} calculation. The stereochemical quality of the models was monitored with the Procheck program (15), and the figures were generated by Chimera (22). Between 91.3 and 91.9% of residues, excluding proline and glycine, were in the favored region, and 7.4 to 8.3% of residues were in the allowed region of the Ramachandran plot (17).

Molecular modeling. The starting structure for simulations was taken from the ultrahigh-X-ray structure of CTX-M-9 (9). The GROMACS package and the geometric and charge parameters of the OPLSAA force field were used to carry out all energy minimizations and molecular dynamics simulations (16). Topology and RESP charges of hydrolyzed piperacillin were obtained from *ab initio* calculations using GAMESS (HF/6-31G* theory level) with RED-II software. Steered molecular dynamics (SMD) simulation was performed as previously described (10).

Protein structure accession numbers. Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 3HRE, 3Q07, and 3Q1F.

RESULTS AND DISCUSSION

Overall crystal structures. The residue Ser70 of CTX-M-9 is responsible for the nucleophilic attack against the carbonyl carbon of β -lactams, which forms a covalent bond with the substrate resulting in the acyl-enzyme. Ser70 side chain can also bind the carboxylate function resulting from β -lactam hydrolysis. A Ser70Gly mutant of CTX-M-9 was therefore constructed to capture intact and hydrolyzed forms of piperacillin in the binding site as noncovalent adducts. As expected, no hydrolytic activity of this mutant against piperacillin was detected by using a computerized microacidimetric method (7).

The structures were determined at high resolution (Table 1). Two monomers were observed in the crystallographic asymmetric unit. The final R_{work} and R_{free} varied from 14.4 to 15.4% for the structure in complex with the intact form of piperacillin

and from 17.5 to 18.4% for the hydrolyzed form, respectively. Electronic density maps confirmed the presence of the Ser70Gly substitution and of the catalytic water molecule. This water molecule was hydrogen bonded to the side chain of Glu166 and Asp170 (2.5 to 2.9 Å) as observed in CTX-M-9 apoenzyme (2P74). Electronic density maps also revealed the presence of ligands in each of the two active sites of the crystallographic asymmetric unit. Densities compatible with the dioxopiperazine cycle identified the piperacillin molecules. Analysis of the structure of the crystal soaked with intact piperacillin identified the intact and hydrolyzed forms of piperacillin in the active site. However, the intact form, which harbors a four-membered β -lactam, was the predominant adduct (80% occupancy). Analysis of the structure of the crystal soaked with hydrolyzed piperacillin identified only the hydrolyzed molecule in the active site. The $C\alpha$ atom root mean square deviation (RMSD) calculated from the overlay of CTX-M-9-S70G apoenzyme structure with the corresponding complex structures varied from 0.08 to 0.18 Å, showing that the insertion of piperacillin in the catalytic pocket did not modify the overall structure of the enzyme. In addition, the hydrogen bonds between the key residues of the binding site were similar to those previously observed in CTX-M enzymes (Table 2).

Binding of piperacillin into the CTX-M-9-S70G active site. The electrostatic interactions, previously identified as critical for hydrolytic activity (8, 13, 14, 18, 26), were observed between piperacillin and key residues of the catalytic pocket (Fig. 1A). The oxygen atom of the β -lactam ring was located in an oxyanion hole formed by Ser70N and Ser237N. C-3 carboxylate function interacted with conserved residues Ser130 and Thr235 of the active site. C-6 amide function interacted with the residues at positions 104, 132, and 237. The other atoms of the C-6 substituent did not establish additional electrostatic interactions with the CTX-M-9-S70G protein. The dioxopiperazine substituent of piperacillin adopted a positioning similar to that of the cefotaxime methoximino group. The latter bulky group can be accommodated only if the catalytic pocket is opened and the conserved hydrogen bond between the residues N170

TABLE 2. Key interactions in the complex structures with β -lactams

Type of distance	Complex	Distance (Å)											
		CTX-M-9 S70G		CTX-M-9 S70G/ benzylpenicillin		CTX-M-9 S70G/ S70G/piperacillin		CTX-M-9 S70G/ hydrolyzed benzylpenicillin		CTX-M-9 S70G/ hydrolyzed piperacillin			
		Chain A	Chain B	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B	Chain A	Chain B		
Distance between key residues of the binding site ^a	H ₂ O cat-Glu166 (Oε1)	2.57	2.50	2.62	2.67	2.54	2.50	2.62	2.65	2.50	2.47		
	H ₂ O cat-Asn170 (Oδ1)	2.69	2.74	2.83	2.94	2.80	2.89	2.80	2.78	2.79	2.89		
	Lys73 (Nζ)-Asn132 (Oδ1)	2.70	2.69	2.68	2.73	2.69	2.67-2.76	2.65-2.75	2.68	2.68	2.68-2.78		
	Lys73 (Nζ)-Ser130 (Oγ)	2.89	2.84	2.90	2.88	2.86	2.93	2.92-3.05	2.84	2.88	2.92		
	Ser130 (Oγ)-Lys234 (Nζ)	2.84	2.82	2.86	2.88	2.77	2.82	2.83	2.82	2.81	2.86		
	Glu166 (Oε2)-Asn170 (Nδ2)	2.74	2.76	2.74	2.78	2.79	2.81	2.73	2.81	2.81	2.83		
	Pro167 (O)-Asn170 (N)	3.07	3.17	3.04	3.09	3.16	3.26	3.12	3.20	3.20	3.27		
	Pro167 (O)-Asn170 (Nδ2)	3.23	3.07	3.02	2.94	3.03	3.00	3.00	2.99	3.09	3.03		
	Asn170 (O)-Asp240 (N)	3.05	3.05	3.15	3.11	3.22	3.29	3.15	3.14	3.20	3.28		
	Distance between key residues and amide function of β -lactams	N-12 (β -lactam)-Ser237 (O)	NA ^b	NA	2.85	2.89	2.95	2.97	2.84	2.83	2.95	3.00	
O-14 (β -lactam)-Asn104 (Nδ2)		NA	NA	3.02	3.10	3.03	2.99	3.08	3.12	3.07	3.00		
O-14 (β -lactam)-Asn132 (Nδ2)		NA	NA	2.90	2.93	2.91	2.90	3.01	3.01	2.90	2.91		
Distance between key residues and conserved C-3 carboxylate function of β -lactams	O-11 (β -lactam)-Ser130 (Oγ)	NA	NA	2.48^c	2.52	3.03	2.88	4.35	4.22	5.45	5.55		
	O-11 (β -lactam)-Thr235 (Oδ)	NA	NA	2.56	2.62	2.70	2.92	2.61	2.65	4.82	5.01		
	O-11 (β -lactam)-Lys234 (Nζ)	NA	NA	3.03	3.07	3.65	3.74	4.53	4.44	6.74	6.89		
	O-10 (β -lactam)-Ser237 (Oγ)	NA	NA	4.36	4.00	2.79	2.62	4.51	4.51	3.03	3.03		
Distance between key residues and carbonyl function of the β -lactam ring or the carboxylate resulting from hydrolysis	O-8 (β -lactam)-Gly70 (N)	NA	NA	2.83	2.83	2.82	2.88	2.77	2.76	2.83	2.84		
	O-8 (β -lactam)-Ser237 (N)	NA	NA	2.83	2.85	2.88	2.95	2.82	2.76	2.86	2.86		
	O-8' (β -lactam)-Lys73 (Nζ)	NA	NA	NA	NA	NA	NA	2.66	2.83	2.66	2.46		
	O-8' (β -lactam)-Ser130 (Oγ)	NA	NA	NA	NA	NA	NA	2.49	2.43	3.00	2.98		

^a Two distances in monomers A and B. In cases with two distances shown within one monomer, multiple conformations are present.

^b NA, not applicable.

^c Boldface indicates a significant difference in enzyme-ligand distances between complexes.

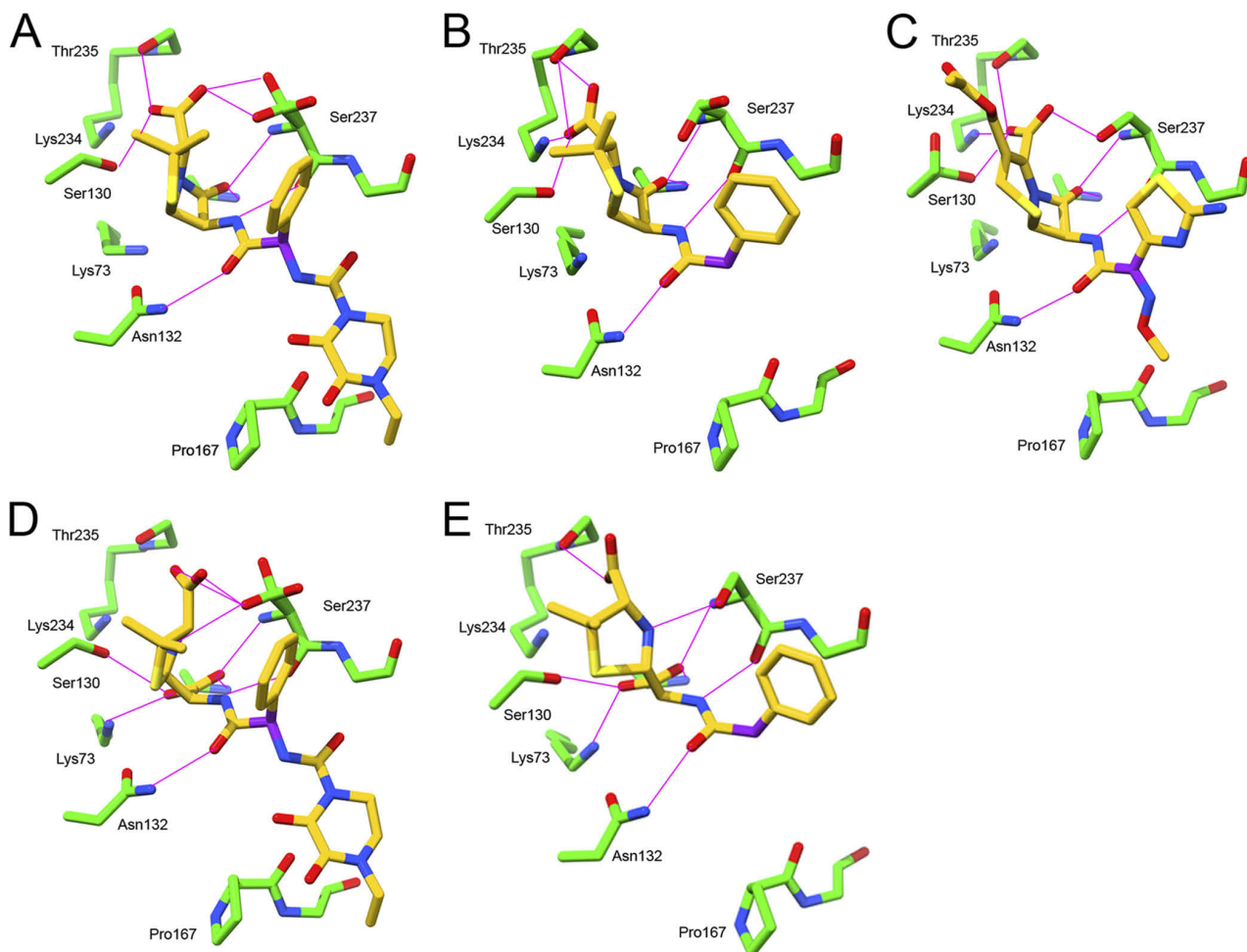


FIG. 1. Key polar interactions observed with crystallographic structures between CTX-M-9-S70G and piperacillin (A), benzylpenicillin (B), cefotaxime (C), hydrolyzed piperacillin (D), and hydrolyzed benzylpenicillin (E). Carbon atoms are in green for the protein or in yellow for the adduct; oxygen atoms are red, and nitrogen atoms are blue. The α -carbon of the carbonyl of the side chain of β -lactams is magenta.

and D240 is broken (10). The binding of piperacillin did not induce such conformational change, although the dioxopiperazine substituent is bigger than the methoximino group (volume, 265 versus 153 \AA^3). This paradox can be explained by the configuration of the α -carbon of the carbonyl of the C-6/C-7 side chain harboring these two substituents. The imino sp^2 of the α -carbon of the side chain of cefotaxime imposes a planar configuration on the carbons harboring the aminothiazolidine methoximino groups, which are difficult to accommodate and open the CTX-M-9 binding site. The sp^3 hybridization of the α -carbon of the side chain of piperacillin confers a tetrahedral configuration to the C bearing the benzyl dioxopiperazine group, in which the H atom is oriented toward the β_3 strand and allows the accommodation of the substituents without notable constraint (Fig. 2). The steric hindrance of the C-6/7 substituent of β -lactams is reported as an important factor in the resistance to β -lactamases (11, 18, 23). Our crystallographic data reveal that the presence of a bulky side chain is not sufficient to explain a restricted recognition of substrate by the class A enzymes; topographic constraints must also be

imposed on substituents of the side chain. In oximino cephalosporins, this constraint is supplied by the imino group, in which the α -carbon adjacent to the acylamino group adopts a planar hybridization (sp^2), which puts the substituents in steric hindrance with the β_3 strand and Ω loop (Fig. 2). In contrast, the accommodation of the C-6 substituent of piperacillin occurs without notable constraint. Accordingly, the effect linked to the modifications of the C-6 substituent in penicillins is slight when considering the K_m values (25 and 20 μM for benzylpenicillin and piperacillin, respectively [2]).

However, the positioning of the C-6 side chain of the piperacillin toward the β_3 strand, to avoid steric hindrance between the dioxopiperazine cycle and the residue Pro167 of the Ω loop, induces an 0.7- \AA shift in the opposite direction of its thiazolidine ring toward Asn132, in comparison with benzylpenicillin. Consequently, the interactions of the C-3 carboxylate functions of piperacillin and benzylpenicillin were different. Piperacillin C-3 carboxylate was at a hydrogen bond distance of Ser237O γ (2.6 to 2.8 \AA), unlike benzylpenicillin C-3 carboxylate (4.0 to 4.4 \AA). Interestingly, no hydrogen bond was

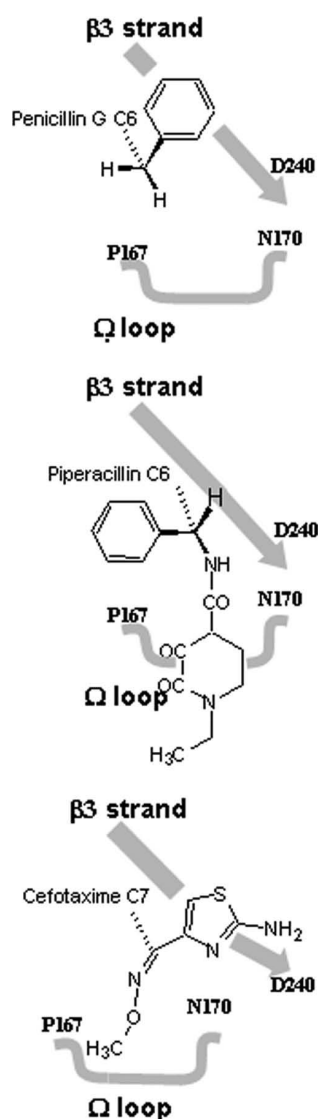


FIG. 2. Schematic representation of the α -carbon of the carbonyl of the C-6/C-7 side chain of benzylpenicillin, cefotaxime, and piperacillin, in relation to the β 3 strand and Ω loop of CTX-M-9. The α -carbon of cefotaxime adopts planar (sp^2) hybridization; the α -carbon of penicillins adopts a tetrahedral (sp^3) hybridization.

established between the C-3 carboxylate function of piperacillin and lysine 234, in contrast to penicillin and cefotaxime (3.7 versus 3.0 and 3.2 Å, respectively; Fig. 1A and 1B). Lysine 234 is conserved among the class A β -lactamases; it belongs to the KTG motif. This cationic residue binds the C-3 carboxylate common to most β -lactamase substrates and is important for the stabilization of the transition state in acylation, much more so than for the complementarity between the enzyme and the substrate ground states, which is responsible for the initial binding (4). The protonated lysine of the KTG motif is also important in penicillin-binding proteins, in which it aids in the proper positioning of the substrate during acyl-enzyme complex formation (30). The absence of a bond between the lysine

234 and the C-3 carboxylate of piperacillin can affect the acylation and may be therefore involved in the decrease of the k_{cat} value observed for piperacillin in comparison to those measured for penicillin and cefotaxime (110 s^{-1} versus 295 s^{-1} and 450 s^{-1} , respectively) (2).

Product release from CTX-M binding site. In the structure in complex with the hydrolysis product of piperacillin, the C-7 carbonyl function of the β -lactam ring was replaced by a carboxylate function, which interacted with active-site residues (Fig. 1D) as observed in the complex structure of CTX-M-9-S70G with hydrolyzed benzylpenicillin (Fig. 1E). Because of electrostatic repulsion between this carboxylate function and the C-3 carboxylate function of the thiazolidine ring, the thiazolidine ring of both hydrolyzed penicillins rotated $\sim 60^\circ$. It disrupted interactions with residues Ser130 (5.5 to 5.6 Å) and Thr235 (4.8 to 5.0 Å) for piperacillin, but the hydrogen bond with the O δ atom of Ser237 (2.9 Å) was conserved. To complete the crystallographic data, the release of hydrolyzed piperacillin was investigated by steered molecular dynamics (SMD) simulations. The thiazolidine ring rotated 180° over the β 3 strand, and the interactions between the C-3 carboxylate group of the thiazolidine ring and the electropositive region of the CTX-M binding site (Ser130, Lys234, Thr235) were disrupted. During this rotation, the C-3 carboxylate group interacted with Ser237 until the hydrolyzed piperacillin exited from the active site. Piperacillin and cefotaxime are therefore expelled from the active site of CTX-M-9 by a closely related mechanism (10). However, the differences in the structures of the penicillin and cephalosporin products may be involved in the higher turnover of CTX-M against cephalosporins than against penicillins ($k_{cat} = 3,000, 450, 295,$ and $110 s^{-1}$, respectively, for cephalothin, cefotaxime, benzylpenicillin, and piperacillin [2]). Indeed, the carbon adjacent to the carboxylate function of cephalosporins, such as cefotaxime, adopts a planar hybridization sp^2 and that of penicillins adopts a tetrahedral hybridization sp^3 . Consequently, the orientation of cephalosporin carboxylate function is in the plane of the 6-membered ring. Because of this, the rotation of the ring observed with penicillin-derived products may create a steric clash between carboxylate of cephalosporins and the Ser237 side chain. This steric conflict, absent in penicillins, can help an efficient auto-expulsion of cephalosporin-derived product.

In conclusion, we determined the crystal structure of CTX-M-9-S70G in complex with intact and hydrolyzed forms of piperacillin. The findings revealed the determinants of the interaction between the CTX-M binding site and piperacillin at pre-covalent and posthydrolytic steps of the catalytic process. From these data, we show that the configuration of the α -carbon of the carbonyl of the side chain of β -lactams is crucial for their accommodation and that the absence of a bond between Lys234 and the C-3 carboxylate function of piperacillin could explain, in part, a lower catalytic activity of the CTX-M enzymes against this penicillin.

ACKNOWLEDGMENTS

This work was supported, in part, by a grant from INRA and Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche (Paris, France).

We thank Rolande Perroux and Marlene Jan for technical assistance. Provision of beam time at the European Synchrotron Radiation Facility (ESRF, Grenoble, France) is gratefully acknowledged.

REFERENCES

1. Ambler, R. P., et al. 1991. A standard numbering scheme for the class A beta-lactamases. *Biochem. J.* **276**:269–270.
2. Bonnet, R. 2004. Growing group of extended-spectrum beta-lactamases: the CTX-M enzymes. *Antimicrob. Agents Chemother.* **48**:1–14.
3. Bradford, P. A. 2001. Extended-spectrum beta-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. *Clin. Microbiol. Rev.* **14**:933–951.
4. Brannigan, J., et al. 1991. The mutation Lys234His yields a class A beta-lactamase with a novel pH-dependence. *Biochem. J.* **278**:673–678.
5. Cantón, R., et al. 2008. Prevalence and spread of extended-spectrum beta-lactamase-producing Enterobacteriaceae in Europe. *Clin. Microbiol. Infect.* **14**(Suppl. 1):144–153.
6. Cantón, R., and T. M. Coque. 2006. The CTX-M beta-lactamase pandemic. *Curr. Opin. Microbiol.* **9**:466–475.
7. Chen, Y., J. Delmas, J. Siro, B. Shoichet, and R. Bonnet. 2005. Atomic resolution structures of CTX-M beta-lactamases: extended spectrum activities from increased mobility and decreased stability. *J. Mol. Biol.* **348**:349–362.
8. Chen, Y., B. Shoichet, and R. Bonnet. 2005. Structure, function, and inhibition along the reaction coordinate of CTX-M beta-lactamases. *J. Am. Chem. Soc.* **127**:5423–5434.
9. Chen, Y., R. Bonnet, and B. K. Shoichet. 2007. The acylation mechanism of CTX-M beta-lactamase at 0.88 Å resolution. *J. Am. Chem. Soc.* **129**:5378–5380.
10. Delmas, J., et al. 2010. Structural insights into substrate recognition and product expulsion in CTX-M enzymes. *J. Mol. Biol.* **400**:108–120.
11. Delmas, J., et al. 2008. Structure and dynamics of CTX-M enzymes reveal insights into substrate accommodation by extended-spectrum beta-lactamases. *J. Mol. Biol.* **375**:192–201.
12. Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**:2126–2132.
13. Gazouli, M., E. Tzelepi, S. V. Sidorenko, and L. S. Tzouvelekis. 1998. Sequence of the gene encoding a plasmid-mediated cefotaxime-hydrolyzing class A beta-lactamase (CTX-M-4): involvement of serine 237 in cephalosporin hydrolysis. *Antimicrob. Agents Chemother.* **42**:1259–1262.
14. Ibuka, A. S., et al. 2003. Crystal structure of extended-spectrum beta-lactamase Toho-1 insights into the molecular mechanism for catalytic reaction and substrate specificity expansion. *Biochemistry* **42**:10634–10643.
15. Laskowski, R. A., D. S. Moss, and J. M. Thornton. 1993. Main-chain bond lengths and bond angles in protein structures. *J. Mol. Biol.* **231**:1049–1067.
16. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model.* **7**:306–317.
17. Lovell, S. C., et al. 2003. Structure validation by C α geometry: phi, psi and Cbeta deviation. *Proteins* **50**:437–450.
18. Massova, I., and S. Mobashery. 1996. Molecular bases for interactions between beta-lactam antibiotics and beta-lactamases. *Acc. Chem. Res.* **30**:162–168.
19. Matagne, A., J. Lamotte-Brasseur, and J. M. Frere. 1998. Catalytic properties of class A beta-lactamases: efficiency and diversity. *Biochem. J.* **330**:581–598.
20. Reference deleted.
21. Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Macromol. Crystallogr. A* **276**:307–326.
22. Pettersen, E. F., et al. 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**:1605–1612.
23. Philippon, A., G. Arlet, and G. A. Jacoby. 2002. Plasmid-determined AmpC-type beta-lactamases. *Antimicrob. Agents Chemother.* **46**:1–11.
24. Potterton, E., S. McNicholas, E. Krissinel, K. Cowtan, and M. Noble. 2002. The CCP4 molecular-graphics project. *Acta Crystallogr. Sect. D* **58**:1955–1957.
25. Rossolini, G. M., M. M. D'Andrea, and C. Mugnaioli. 2008. The spread of CTX-M-type extended-spectrum beta-lactamases. *Clin. Microbiol. Infect.* **14**(Suppl. 1):33–41.
26. Shimamura, T., et al. 2002. Acyl-intermediate structures of the extended-spectrum class A beta-lactamase, Toho-1, in complex with cefotaxime, cephalothin, and benzylpenicillin. *J. Biol. Chem.* **277**:46601–46608.
27. Steiner, R. A., A. A. Lebedev, and G. N. Murshudov. 2003. Fisher's information in maximum-likelihood macromolecular crystallographic refinement. *Acta Crystallogr. D Biol. Crystallogr.* **59**:2114–2124.
28. Vagin, A., and A. Teplyakov. 2000. An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **56**:1622–1624.
29. Walther-Rasmussen, J., and N. Hoiby. 2004. Cefotaximases (CTX-M-ases), an expanding family of extended-spectrum beta-lactamases. *Can. J. Microbiol.* **50**:137–165.
30. Zhang, W., Q. Shi, O. Meroueh, S. B. Vakulenko, and S. Mobashery. 2007. Catalytic mechanism of penicillin-binding protein 5 of *Escherichia coli*. *Biochemistry* **46**:10113–10121.

Characterization of the Inhibitor-Resistant SHV β -Lactamase SHV-107 in a Clinical *Klebsiella pneumoniae* Strain Coproducing GES-7 Enzyme

Vera Manageiro,^a Eugénia Ferreira,^a Antony Cougnoux,^d Luís Albuquerque,^b Manuela Caniça,^a and Richard Bonnet^{c,d}

National Reference Laboratory of Antimicrobial Resistances, Department of Infectious Diseases, National Institute of Health Dr. Ricardo Jorge, Lisbon, Portugal^a; Laboratory of Microbiology, Centro Hospitalar de Coimbra, EPE, Coimbra, Portugal^b; CHU Clermont-Ferrand, Laboratoire de Bactériologie, Clermont-Ferrand, France^c; and Clermont Université, Université d'Auvergne, Evolution des Bactéries Pathogènes et Susceptibilité de l'Hôte, Clermont-Ferrand, France^d

The clinical *Klebsiella pneumoniae* INSRA6884 strain exhibited nonsusceptibility to all penicillins tested (MICs of 64 to >2,048 $\mu\text{g/ml}$). The MICs of penicillins were weakly reduced by clavulanate (from 2,048 to 512 $\mu\text{g/ml}$), and tazobactam restored piperacillin susceptibility. Molecular characterization identified the genes $bla_{\text{GES-7}}$ and a new β -lactamase gene, $bla_{\text{SHV-107}}$, which encoded an enzyme that differed from SHV-1 by the amino acid substitutions Leu35Gln and Thr235Ala. The SHV-107-producing *Escherichia coli* strain exhibited only a β -lactam resistance phenotype with respect to amoxicillin, ticarcillin, and amoxicillin-clavulanate combination. The kinetic parameters of the purified SHV-107 enzyme revealed a high affinity for penicillins. However, catalytic efficiency for these antibiotics was lower for SHV-107 than for SHV-1. No hydrolysis was detected against oxyimino- β -lactams. The 50% inhibitory concentration (IC_{50}) for clavulanic acid was 9-fold higher for SHV-107 than for SHV-1, but the inhibitory effects of tazobactam were unchanged. Molecular dynamics simulation suggested that the Thr235Ala substitution affects the accommodation of clavulanate in the binding site and therefore its inhibitory activity.

β -Lactamases are the most important mechanism of β -lactam drug resistance in Gram-negative bacteria. They hydrolyze and inactivate β -lactams and are divided into four major classes (A to D) on the basis of their sequence (1). Class B is composed of metalloenzymes which require the presence of zinc cations for activity; classes A, C, and D are hydrolases with an active-site serine (1, 5). Class A enzymes comprise several enzyme families, including the clinically relevant β -lactamases TEM-1 and SHV-1 (35). These enzymes confer resistance to penicillins and to narrow-spectrum cephalosporins but are inhibited by β -lactamase inhibitors such as clavulanate, which have been developed to evade the activity of these β -lactamases (17, 38).

However, the broad-spectrum β -lactamases TEM-1, TEM-2, and SHV-1 have acquired point mutations in key positions, often near the active site, that have given rise to (i) TEM/SHV-type extended-spectrum β -lactamases (ESBLs), which hydrolyze oxyimino- β -lactams; (ii) TEM/SHV-type inhibitor-resistant β -lactamases; and (iii) complex mutant TEMs (CMTs), which can combine resistance to β -lactam inhibitors and activity against oxyimino- β -lactams (4, 11, 12, 23, 34). In addition, many recent reports have described the emergence of class A ESBLs belonging to other families, such as GES, VEB, PER, and the worldwide-disseminated CTX-M family (2, 21). The GES family, specifically, GES-1, has also been reported in many areas including Portugal (28).

Since 1992, about 35 inhibitor-resistant TEMs (IRTs) and 11 CMTs have been found (<http://www.lahey.org/Studies/>). However, only six inhibitor-resistant SHVs (IRSs) have been detected clinically so far, mainly in *Klebsiella pneumoniae* strains, in different countries (8, 13, 14, 26, 31, 37) (Table 1). No complex mutant SHV has been encountered, which is surprising since SHV enzymes are usually more susceptible to inactivation by clavulanate than TEM, due to differences in the enzyme active sites (4, 11). In the SHV family, naturally occurring inhibitor-resistant phenotypes harbor substitutions at positions Met69 (SHV-49), Ser130 (SHV-10), Arg187 (SHV-26), and Lys234 (SHV-56, SHV-72, and

SHV-84) (numbering according to the Ambler classification) (1, 11) (Table 1). Further investigations of most enzymes and the corresponding substitutions have been performed. However, the recently reported enzyme SHV-107 (AM941848) (28), which is suspected to belong to the IRS group, has not been characterized.

In this study, we report in a clinical *K. pneumoniae* strain the coexpression of the ESBL GES-7 and a new IRS-type enzyme called SHV-107, which, compared to SHV-1, harbors the substitutions Leu35Gln and Thr235Ala. The kinetic constants of SHV-107 were determined, and molecular modeling was undertaken to investigate the role of Thr235Ala in the resistance to clavulanate.

MATERIALS AND METHODS

Bacterial strains and plasmid. *K. pneumoniae* INSRA6884 was isolated in 2006 from the sputum of a 92-year-old woman in an internal medicine service at the Centro Hospitalar in Coimbra, Portugal (28). The *Escherichia coli* DH5 α $\Delta ampC$ strain and *E. coli* BL21 (DE3) were used for initial cloning experiments and overexpression of the β -lactamase-encoding gene, respectively. *E. coli* C600 was used for mating-out assays. Plasmid pBK-CMV (Stratagene, Amsterdam, The Netherlands) was used for cloning experiments (9).

Antimicrobial susceptibility testing and ESBL detection. An agar dilution method and criteria according to the French Society of Microbiology (SFM) (7) were used to determine the MICs of the antibiotics listed in Table 2.

Analytical IEF. β -Lactamases were characterized by isoelectric focusing (IEF) as previously described (6), with strains *E. coli* C600 (SHV-1, pI

Received 19 October 2010 Returned for modification 8 December 2010

Accepted 23 October 2011

Published ahead of print 14 November 2011

Address correspondence to Manuela Caniça, manuela.canica@insa.min-saude.pt.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/AAC.01444-10

TABLE 1 Comparison of amino acid substitutions, inhibitory properties, and epidemiology of IRS β -lactamases

β -Lactamase	Amino acid at position no.: ^a														IC ₅₀ (μ M) (CLA/TAZ) ^b	Strain ^c	Country (yr) ^d	Reference(s)
	8	35	54	69	130	140	146	187	192	193	234	235	238	240				
SHV-1	I	L	G	M	S	A	A	A	K	L	K	T	G	E	0.17/0.11	<i>K. pneumoniae</i>	Switzerland (1972)	36
SHV-10			Del		G	R			N	V			S	K	6.9/1.3	<i>E. coli</i>	Greece (1997)	37
SHV-26								T							0.48/ND ^e	<i>K. pneumoniae</i>	Taiwan (2001)	8
SHV-49				I											1.5/2.5	<i>K. pneumoniae</i>	France (2004)	13
SHV-56		Q									R				2.5/0.75	<i>K. pneumoniae</i>	France (2008)	14
SHV-72	F					V					R				1.72/0.08	<i>K. pneumoniae</i>	Portugal (2008)	31
SHV-84											R				2.21/0.03	<i>E. coli</i>	Portugal (2010)	26
SHV-107		Q										A			1.53/0.11	<i>K. pneumoniae</i>	Portugal (2009)	28; this study

^a Numbering according to Ambler et al. (1).

^b Values of clavulanate (CLA) and tazobactam (TAZ) for the different IRS β -lactamases. SHV-1, SHV-72, and SHV-84, SHV-10, and SHV-107 values were determined using 200 μ M ticarcillin, 50 mM nitrocefin, and 200 μ M penicillin, respectively, which were incubated with the inhibitor for 5 min at 37°C at pH 7; SHV-49 and SHV-56 values were determined by benzylpenicillin (100 μ M) hydrolysis after 3 min of preincubation at 30°C; and SHV-26 values were determined after 10 min of preincubation of nitrocefin with the clavulanate at room temperature. IC₅₀, 50% inhibitory concentration.

^c Strain producing IRS β -lactamase, in the first report.

^d Country that first reported the β -lactamase, and year (in parentheses).

^e ND, not determined.

7.6), *K. pneumoniae* INSRA5767 (TEM-24, pI 6.5, and SHV-1, pI 7.6), and *E. coli* INSRA7813 (GES-1, pI 5.8) as standards.

DNA amplification and sequencing of β -lactamase genes. β -Lactamase genes *bla*_{SHV}, *bla*_{CXA}, *bla*_{TEM}, *bla*_{CTX}, and *bla*_{GES} were detected by PCR amplification and then purified and sequenced as previously described (28).

***bla*_{GES-7} gene transfer.** Direct transfer of the β -lactam resistance phenotype and selection of transconjugants were performed as described previously (29).

Cloning. The *bla*_{SHV-107} gene was cloned in the plasmid pBK-CMV and transferred as previously described (31). The recombinant pBK-SHV-107 plasmid was transformed by electroporation into the *E. coli* BL21(DE3) strain. Transformants were selected with 30 μ g/ml of kanamycin and 50 μ g/ml of amoxicillin. The orientation of the inserted genes was confirmed as described elsewhere (31).

Purification of β -lactamase. SHV-107 was produced from *E. coli* BL21(DE3) in LB broth, supplemented with 30 μ g/ml kanamycin, 16

TABLE 2 MICs of β -lactam antibiotics for the *K. pneumoniae* clinical strain producing the SHV-107 enzyme, recipient, and transformant^a

Antimicrobial drug	MIC (μ g/ml) for strain:				
	<i>K. pneumoniae</i> INSRA6884 (SHV-107 + GES-7)	<i>E. coli</i> BL21(DE3)	<i>E. coli</i> BL21(DE3) (pBK-SHV-107)	<i>E. coli</i> C600	<i>E. coli</i> C600 (GES-7)
Amoxicillin	2,048	≤2	256	≤2	256
Amoxicillin + CLA ^b	512	≤2	64	≤2	8
Ticarcillin	>2,048	≤2	128	≤2	2,048
Piperacillin	64	1	8	0.5	4
Piperacillin + TAZ ^c	4	1	1	0.5	0.5
Amdinocillin	1	≤0.015	≤0.015	≤0.015	0.25
Cephalothin	128	1	1	4	32
Cefuroxime	64	0.5	1	2	16
Cefoperazone	8	≤0.25	≤0.25	≤0.25	0.5
Ceftazidime	128	0.06	0.06	0.06	32
Ceftazidime + CLA	8	0.03	0.03	≤0.015	0.5
Ceftriaxone	4	≤0.015	≤0.015	≤0.015	0.5
Ceftriaxone + CLA	0.5	≤0.015	≤0.015	≤0.015	≤0.015
Cefotaxime	4	≤0.015	≤0.015	0.03	0.5
Cefotaxime + CLA	0.25	≤0.015	≤0.015	≤0.015	0.03
Aztreonam	2	≤0.015	≤0.015	0.06	1
Aztreonam + CLA	0.125	≤0.015	≤0.015	0.03	0.03
Cefepime	0.25	≤0.015	≤0.015	≤0.015	0.06
Cefoxitin	4	2	2	2	2
Imipenem	0.25	0.25	0.25	≤0.06	≤0.06
Kanamycin	4	≤0.125	≤0.125	≤0.125	≤0.125
Nalidixic acid	4	≤0.003	≤0.003	4	4
Ciprofloxacin	0.03	≤0.003	≤0.003	0.007	0.007
Trimethoprim	0.5	≤0.125	≤0.125	≤0.125	≤0.125

^a *E. coli* BL21(DE3)(pBK-SHV-107) and *E. coli* C600 (GES-7) were a transformant and conjugant, respectively, of *K. pneumoniae* INSRA6884 (harboring SHV-107 plus GES-7 enzyme).

^b CLA, clavulanate at a fixed concentration of 2 μ g/ml.

^c TAZ, tazobactam at a fixed concentration of 4 μ g/ml.

TABLE 3 Comparative kinetic parameters for SHV-1 and SHV-107 β -lactamases

Antibiotic	Kinetic parameters of enzyme			SHV-107 ^c		
	SHV-1 ^b			SHV-107 ^c		
	K_m (μM) ^a	k_{cat} (s^{-1}) ^a	k_{cat}/K_m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	K_m (μM) ^a	k_{cat} (s^{-1}) ^a	k_{cat}/K_m ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)
Penicillin G	23 \pm 0.4	1,937 \pm 82	84	11 \pm 0.9	378 \pm 6	34
Amoxicillin	31 \pm 1	1,044 \pm 10	34	42 \pm 4	657 \pm 10	16
Ticarcillin	11 \pm 3	220 \pm 49	20	23 \pm 12	51 \pm 7	2
Piperacillin	24 \pm 0.5	1,490 \pm 96	62	31 \pm 0.4	412 \pm 87	13
Cephalothin	40 \pm 1	128 \pm 33	3	179 \pm 9	<0.1	<0.001
Ceftazidime	142 \pm 3	<0.1	ND ^d	ND	<0.1	ND
Cefotaxime	257 \pm 21	<0.1	ND	ND	<0.1	ND

^a Values are means \pm standard deviations.

^b The kinetic constants for SHV-1 are from reference 31.

^c *E. coli* BL21(DE3)(pBK-SHV-107) was the transformant producing SHV-107.

^d ND, not determinable, because the hydrolysis rates were too low.

$\mu\text{g/ml}$ ticarcillin, and 0.1 mM IPTG (isopropyl- β -D-thiogalactopyranoside; Sigma Chemical Co., St. Louis, MO), overnight at 30°C. The enzyme was extracted by ultrasonic treatment, and the clarified supernatant was purified as described elsewhere (3, 31).

Determination of β -lactamase kinetic constants. The Michaelis constant (K_m) and catalytic activity (k_{cat}) of SHV-107 and the concentrations of the inhibitors (clavulanate and tazobactam) required to inhibit enzyme activity by 50% (IC_{50} s) were determined as previously described (3, 31). Specific activity and IC_{50} s were monitored with penicillin G (200 μM) as the reporter substrate. The inhibitor kinetics with clavulanate were determined under steady-state conditions for SHV-1 and SHV-107 as previously described (18, 27). The enzyme activity was monitored with a UV-1800 spectrophotometer (Shimadzu, Champs sur Marne, France) in phosphate-buffered saline (PBS) 20 mM (pH 7.4) at 25°C by using 100 μM chromogenic substrate nitrocephin ($\Delta\epsilon$, 17,400 $\text{M}^{-1} \text{cm}^{-1}$ at 482 nm) (Sigma-Aldrich, France) as the reporter substrate. The kinetic constants were determined three times for each substrate tested.

Molecular modeling. The structural model of SHV-107 was constructed from the crystal structure of SHV-1 using Chimera (24). Molecular dynamics simulations (MDSs) were performed as previously described with GROMACS software package, version 4.0 (40), using the parameters of the OPLSAA force field (22) in a periodic cubic box, as previously described (10). The AutoDock 4.0 package was used to perform molecular docking (33). Coordinates for clavulanate structure were obtained using CACTVS (<http://www2.ccc.uni-erlangen.de/software/cactvs/>) and the CORINA server (19). Structure optimization was calculated at the 6-31G** level using the Gamess quantum chemistry software (39). All the torsion angles in clavulanate were set free to perform flexible docking, and the Gasteiger-Hückel atomic charge was assigned. Kollman united atom partial charges were assigned for the receptor, which was kept rigid during the docking study. Water molecules, ions, and ligands were removed, except the catalytic water molecule. The empirical free-energy function and Lamarckian genetic algorithm were used for docking with the following settings: a maximum number of 25,000,000 energy evaluations, an initial population of 150 randomly placed individuals, a maximum number of 27,000 generations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value (number of top individuals that automatically survive) of 1. For the local search, the Solis and Wets algorithm was used with a maximum of 300 iterations per search; 100 independent docking runs were performed for each docking experiment. Results were clustered according to default root mean square deviation (RMSD) criteria.

RESULTS AND DISCUSSION

General characteristics of *K. pneumoniae* INSRA6884. In this study, we characterized a clinical *K. pneumoniae* strain isolated in

a Portuguese hospital coexpressing two different β -lactamases, the IRS enzyme SHV-107 and the ESBL GES-7, of pIs 7.5 and 6.9, respectively.

Phenotypic characterization. The transconjugant *E. coli* C600 (GES-7) acquired resistance only to amoxicillin, ticarcillin, cephalothin, cefuroxime, and ceftazidime, in contrast to the clinical strain, which was resistant to all penicillins and all cephalosporins tested except cefepime and ceftoxitin (Table 2). The addition of clavulanate drastically reduced the MIC value of oxyiminocephalosporins and aztreonam in both the transconjugant and the clinical strains. Like the clinical strain, *E. coli* BL21(DE3)(pBK-SHV-107) acquired resistance to amoxicillin and ticarcillin and exhibited a synergy of amoxicillin and piperacillin in combination with clavulanate (2-fold) and tazobactam (8-fold), respectively (Table 2).

β -lactamase analysis. We report for the first time the GES-7 β -lactamase isolated in Portugal. Although not associated with class 1 integrons (data not shown), the gene encoding this β -lactamase was present in a conjugative plasmid. The deduced amino sequence of *bla*_{GES-7} differed from that of GES-1 (1) by Glu107Lys and Leu125Ala substitutions (Ambler numbering). SHV-107 enzyme, an IRS β -lactamase ($\text{IC}_{50} \geq 1 \mu\text{M}$), is the seventh member of this subgroup, which is characterized by the expression of resistance to clavulanate (Table 1). This enzyme differed from SHV-1 by the amino acid substitutions Leu35Gln and Thr235Ala (Table 1) (28). The first substitution, Leu35Gln, frequently encountered and observed alone in the SHV-11 enzyme (30), is known to have no significant impact on kinetic constants (32). To our best knowledge, the second substitution is the first natural substitution reported at position 235 in class A enzymes (Table 1). In these enzymes, Thr235 residue is part of a conserved element (Lys234-Ser/Thr235-Gly236). It is situated on the β strand of a β -sheet in the α/β domain and forms a major wall of the catalytic cavity site (10). The hydroxyl function of Ser/Thr-235 contributes to the binding of the conserved C₃/C₄ carboxylate function of β -lactams (41).

Enzymatic parameters and properties of β -lactamase SHV-107. After purification by ion exchange and gel filtration, the purity of β -lactamase SHV-107 was estimated to be $\geq 98\%$ on a sodium dodecyl sulfate-polyacrylamide gel. The band of 28.8 kDa corresponds to the molecular mass deduced from the amino acid sequence (data not shown).

TABLE 4 Clavulanate kinetic parameters of SHV-1 and SHV-107 β -lactamases

β -Lactamase	K_i (μM) ^a	k_{inact} (s^{-1}) ^a	k_{inact}/K_i ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	k_{cat} (s^{-1}) ^a	k_{cat}/K_i ($\mu\text{M}^{-1} \cdot \text{s}^{-1}$)	$k_{\text{cat}}/k_{\text{inact}}$
SHV-1	0.21 \pm 0.06	0.050 \pm 0.007	0.23	0.94 \pm 0.13	4.48	19
SHV-107	1.82 \pm 0.17	0.042 \pm 0.009	0.02	0.75 \pm 0.11	0.41	18

^a Values are means \pm standard deviations.

We observed that the substitution Thr235Ala altered SHV-107 kinetic constants against β -lactams without dramatic effects on the resistance to penicillins. Two-fold changes or less between the SHV-1 and SHV-107 β -lactamases were obtained for K_m values for penicillins (11 to 31 μM and 11 to 42 μM , respectively) (Table 3). Catalytic efficiency against penicillins was 2- to 9-fold lower for SHV-107 (k_{cat}/K_m , 2.2 to 34.4 $\mu\text{M}^{-1} \cdot \text{s}^{-1}$) than for SHV-1 (k_{cat}/K_m , 20.0 to 84.2 $\mu\text{M}^{-1} \cdot \text{s}^{-1}$) (Table 3). Amoxicillin, with a k_{cat} value of 657 s^{-1} , was the best substrate for SHV-107. Likewise, previous enzymatic studies of TEM-1 and its isogenic mutant Ser235Ala showed similar results for penicillins, although the mutant had greatly reduced cephalosporinase activity (15, 20). In the SHV β -lactamase family, residue Thr235 is also critical for cephalosporinase activity, because its replacement by Ala235 in SHV-107 leads to severe impairment of the catalytic constant against cephalosporins, as observed in TEM-type enzyme (15, 20). K_m values were 4-fold higher for cephalothin or not determinable for the other cephalosporins, because the hydrolysis rate was too low.

It is noteworthy that amoxicillin, ticarcillin, and piperacillin have similar kinetic parameters, but production of SHV-107 β -lactamase increases the MIC by factors of >128, >64, and 8, respectively. In contrast, Ala235 residue, although it contributes to a loss of activity, seems to have no influence in the MIC of cephalothin. These discrepancies between phenotypic and biochemical values could result from the fact that the MIC of a β -lactam for β -lactamase-producing bacteria correlates not only the catalytic efficiency of the enzyme but also its concentration in the periplasm and the penetration of the antibiotic through the outer membrane, as noted before (25).

IC₅₀ results showed that tazobactam was 14-fold more active than clavulanate against the SHV-107 enzyme, as it has a triazole group at the C₂ β -methyl position, which led to the improvement of IC₅₀s (11) (Table 1). IC₅₀s of clavulanate were 9-fold higher for SHV-107 than for SHV-1, and no loss of sensitivity to tazobactam was observed. The Thr235Ala substitution affected the susceptibility to clavulanate and consequently the susceptibility to the combination of amoxicillin-clavulanate. Similarly, three IRS enzymes, SHV-56, SHV-72, and SHV-84, which presented an amino acid substitution in the vicinity of position 235, showed decreased susceptibility to clavulanate (14, 26, 31). Molecular dynamics simulations suggested that the Lys234Arg substitution observed in this mutant probably affects the positioning of the Ser130 side chain, a key element in the inhibition reaction mediated by clavulanate (31).

SHV-1 and SHV-107 inhibitor kinetics with clavulanate were further investigated (Table 4). K_i for clavulanate was almost 9-fold higher for SHV-107 than for SHV-1 (1.82 μM versus 0.21 μM). k_{inact} (inactivation constant) and k_{cat} values were closely similar for both enzymes. Consequently, the second-order rate constants, k_{inact}/K_i and k_{cat}/K_i , were 10-fold lower for SHV-107 than for SHV-1, because of the increase in K_i .

Molecular modeling. We investigated the recognition of clavulanate by the SHV-107 binding site using molecular modeling tools to explain the role of the Thr235Ala substitution in resistance against clavulanate. The introduction of the Thr235Ala substitution caused no overall or large-scale deviation of dynamic properties. Likewise, the survey of residues and water molecules located in the vicinity of position 235 revealed no dramatic change in this area (data not shown). For SHV-1, all docking experiments provided a similar conformation and positioning of clavulanate in the catalytic pocket (RMSD value for clavulanate atoms of <1 Å). The oxygen atom of the β -lactam carbonyl function was at a hydrogen bond distance of the Ser70 and Ala237 N atoms (2.9 and 2.8 Å, respectively), which form the oxyanion hole. The C₃ carboxylate function of clavulanate established electrostatic interaction with Ser130 and Thr235 (2.7 to 2.9 Å). For SHV-107, 98% of clavulanate-docking experiments provided similar positionings of clavulanate in the SHV-107 catalytic pocket (RMSD value for clavulanate atoms of <1 Å). However, the positioning of clavulanate in SHV-107 was clearly different from that in SHV-1. Clavulanate was flipped (Fig. 1). Consequently, the oxygen atom of the β -lactam carbonyl function did not interact with the oxyanion hole. Instead, Ser70 and Ala237 N atoms and the Ala237 O atom shared hydrogen bonds with the C₂ substituent of clavulanate (2.8 to 3.0 Å), and the β -lactam ring was accommodated in the vicinity of position 244. It emerges from these results that Thr235 favors the interaction of β -lactam carbonyl function with the oxyanion hole. This interaction is a key factor in active serine acylation by clavulanate and therefore in the inhibition process mediated by this inhibitor. The Thr235Ala substitution may therefore affect the correct orientation of clavulanate in the binding site and consequently the inhibitory activity of clavulanate. This possibility is supported by the increase in clavulanate K_i with SHV-107.

Conclusion. Our findings show that mutations at Thr235 of the SHV-1 β -lactamase contribute to decreased susceptibility to β -lactamase inhibitors. Given the increasing number of new enzymes conferring resistance to β -lactamase inhibitors, further investigations should be undertaken on IRS enzymes. This emer-

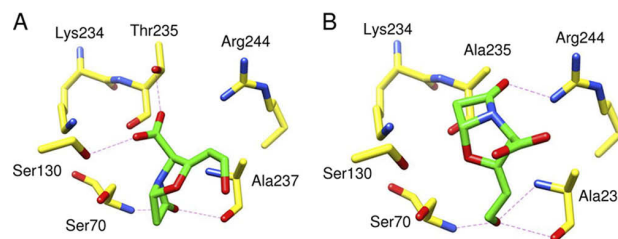


FIG 1 Clavulanate positioning in SHV-1 (A) and SHV-107 (B) binding sites after docking experiments using AutoDock. Carbon atoms are in yellow for the protein or in green for clavulanate; oxygen atoms are red and nitrogen atoms blue. Hydrogen bonds are indicated by dashed lines.

gence may be the consequence of the widespread use of penicillins with β -lactamase inhibitors (J01CR) in different European countries, a practice that needs to be reassessed and modified (16).

ACKNOWLEDGMENTS

We thank Deolinda Louro for technical assistance.

This work was supported financially by the project POCTI/ESP/43037 from Fundação para a Ciência e a Tecnologia, Lisbon, Portugal, awarded to M. Caniça, and by a grant from INRA and Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche (Paris, France), awarded to R. Bonnet. V. Manageiro was supported by grant SFRH/BD/32578/2006 from Fundação para a Ciência e a Tecnologia, Lisbon, Portugal.

REFERENCES

1. Ambler RP, et al. 1991. A standard numbering scheme for class A β -lactamases. *Biochem. J.* 276:269–272.
2. Bonnet R. 2004. Growing group of extended-spectrum β -lactamases: the CTX-M enzymes. *Antimicrob. Agents Chemother.* 48:1–14.
3. Bonnet R, et al. 2001. Novel cefotaximase (CTX-M-16) with increased catalytic efficiency due to substitution Asp-240→Gly. *Antimicrob. Agents Chemother.* 45:2269–2275.
4. Bonomo RA, Rice LB. 1999. Inhibitor resistant class A β -lactamases. *Front. Biosci.* 4:e34–e41.
5. Bush K, Jacoby JA. 2010. An updated functional classification of β -lactamases. *Antimicrob. Agents Chemother.* 54:969–976.
6. Caniça MM, et al. 1997. Properties of IRT-14 (TEM-45), a newly characterized mutant of TEM-type β -lactamases. *Antimicrob. Agents Chemother.* 41:374–378.
7. Cavallo JD, et al. 2009. Recommendations. 2009. Comité de l'antibiogramme de la Société Française de Microbiologie, Paris, France.
8. Chang F-Y, Siu LK, Fung C-P, Huang M-H, Ho M. 2001. Diversity of SHV and TEM β -lactamase in *Klebsiella pneumoniae*: gene evolution in northern Taiwan and two novel β -lactamases, SHV-25 and SHV-26. *Antimicrob. Agents Chemother.* 45:2407–2413.
9. Delmas J, Robin F, Carvalho F, Mongaret C, Bonnet R. 2006. Prediction of the evolution of ceftazidime resistance in extended-spectrum β -lactamase CTX-M-9. *Antimicrob. Agents Chemother.* 50:731–738.
10. Delmas J, et al. 2008. Structure and dynamics of CTX-M enzymes reveal insights into substrate accommodation by extended-spectrum β -lactamases. *J. Mol. Biol.* 375:192–201.
11. Drawz SM, Bonomo RA. 2010. Three decades of β -lactamase inhibitors. *Clin. Microbiol. Rev.* 23:160–201.
12. Du Bois SK, Marriott MS, Amyes SG. 1995. TEM- and SHV-derived extended-spectrum β -lactamases: relationship between selection, structure, and function. *J. Antimicrob. Chemother.* 35:7–22.
13. Dubois V, et al. 2004. SHV-49, a novel inhibitor resistant β -lactamase in a clinical isolate of *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* 48:4466–4469.
14. Dubois V, et al. 2008. Molecular and biochemical characterization of SHV-56, a novel inhibitor-resistant β -lactamase from *Klebsiella pneumoniae*. *Antimicrob. Agents Chemother.* 52:3792–3794.
15. Dubus A, Wilkin JM, Raquet X, Normark S, Frère JM. 1994. Catalytic mechanism of active-site serine β -lactamases: role of the conserved hydroxy group of the Lys-Thr(Ser)-Gly triad. *Biochem. J.* 301:485–494.
16. Ferech M, et al. 2006. European Surveillance of Antimicrobial Consumption (ESAC): outpatient antibiotic use in Europe. *J. Antimicrob. Chemother.* 58:401–407.
17. Finlay J, Miller L, Poupard JA. 2003. A review of the antimicrobial activity of clavulanate. *J. Antimicrob. Chemother.* 52:18–23.
18. Frère JM, Dormans C, Duyckaerts C, De Graeve J. 1982. Interaction of β -iodopenicillanate with the beta-lactamases of *Streptomyces albus* G and *Actinomadura* R39. *Biochem. J.* 207:437–444.
19. Gasteiger J, Rudolph C, Sadowski J. 1990. Automatic generation of 3D-atomic coordinates for organic molecules. *Tetrahedron Comput. Methods* 3:537–547.
20. Imtiaz U, Manavathu EK, Lerner SA, Mobashery S. 1993. Critical hydrogen bonding by serine 235 for cephalosporinase activity of TEM-1 β -lactamase. *Antimicrob. Agents Chemother.* 37:2438–2442.
21. Jacoby GA, Munoz-Price LS. 2005. The new β -lactamases. *N. Engl. J. Med.* 352:380–391.
22. Jorgensen WL, Maxwell DS, Tirado-Rives J. 1996. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* 118:11225–11236.
23. Knox JR. 1995. Extended-spectrum and inhibitor-resistant TEM-type β -lactamases: mutations, specificity, and three-dimensional structure. *Antimicrob. Agents Chemother.* 39:2593–2601.
24. Kuzin AP, et al. 1999. Structure of the SHV-1 β -lactamase. *Biochemistry* 38:5720–5727.
25. Lakaye B, Dubus A, Lepage S, Gros Lambert S, Frère J-M. 1999. When drug inactivation renders the target irrelevant to antibiotic resistance: a case study with β -lactams. *Mol. Microbiol.* 31:89–101.
26. Manageiro V, Ferreira E, Albuquerque L, Bonnet R, Caniça M. 2010. Biochemical study of a new inhibitor-resistant β -lactamase, SHV-84, produced by a clinical *Escherichia coli* strain. *Antimicrob. Agents Chemother.* 54:2271–2272.
27. Matagne A, et al. 1995. Kinetic study of interaction between BRL 42715, β -lactamases, and D-alanyl-D-alanine peptidases. *Antimicrob. Agents Chemother.* 39:227–231.
28. Mendonça N, Ferreira E, Louro D, ARSIP participants, Caniça M. 2009. Molecular epidemiology and antimicrobial susceptibility of extended- and broad-spectrum β -lactamase-producing *Klebsiella pneumoniae* isolated in Portugal. *Int. J. Antimicrob. Agents* 34:29–37.
29. Mendonça N, Ferreira E, Louro D, Caniça M. 2006. Occurrence of a novel SHV-type enzyme (SHV-55) among isolates of *Klebsiella pneumoniae* from Portuguese origin in a comparison study for extended-spectrum β -lactamase-producing evaluation. *Diagn. Microbiol. Infect. Dis.* 56:415–420.
30. Mendonça N, Nicolas-Chanoine M, Caniça M. 2009. Diversity of the *bla*_{SHV} genes. *Diagn. Microbiol. Infect. Dis.* 65:439–446.
31. Mendonça N, et al. 2008. The Lys234Arg substitution in the enzyme SHV-72 is a determinant for resistance to clavulanate inhibition. *Antimicrob. Agents Chemother.* 52:1806–1811.
32. Nuesch-Inderbinen MT, Kayser FH, Hachler H. 1997. Survey and molecular genetics of SHV β -lactamases in *Enterobacteriaceae* in Switzerland: two novel enzymes, SHV-11 and SHV-12. *Antimicrob. Agents Chemother.* 41:943–949.
33. Osterberg F, Morris GM, Sanner MF, Olson AJ, Goodsell DS. 2002. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in AutoDock. *Proteins* 46:34–40.
34. Page MG. 2008. Extended-spectrum β -lactamases: structure and kinetic mechanism. *Clin. Microbiol. Infect.* 14:63–74.
35. Paterson DL, Bonomo RA. 2005. Extended-spectrum β -lactamases: a clinical update. *Clin. Microbiol. Rev.* 18:657–686.
36. Pitton JS. 1972. Mechanism of the bacterial resistance to antibiotics, p 15–93. In Adrian RH, et al (ed), *Reviews of physiology*, vol 65. Springer, Berlin, Germany.
37. Prinarakis EE, Miriagou V, Tzelepi E, Gazouli M, Tzouveleki LS. 1997. Emergence of an inhibitor-resistant β -lactamase (SHV-10) derived from an SHV-5 variant. *Antimicrob. Agents Chemother.* 41:838–840.
38. Reading C, Cole M. 1977. Clavulanate: a β -lactamase-inhibiting β -lactam from *Streptomyces clavuligerus*. *Antimicrob. Agents Chemother.* 11:852–857.
39. Schmidt MW, Jr, et al. 2004. General atomic and molecular electronic structure system. *J. Comput. Chem.* 14:1347–1363.
40. Van Der Spoel D, et al. 2005. GROMACS: fast, flexible, and free. *J. Comput. Chem.* 26:1701–1718.
41. Zafaralla G, Manavathu EK, Lerner SA, Mobashery S. 1992. Elucidation of the role of arginine-244 in the turnover processes of class A β -lactamases. *Biochemistry* 31:3847–3852.

Résumé : La colibactine est une toxine largement distribuée chez *Escherichia coli*. Sa synthèse est assurée par des enzymes codés par un îlot génomique appelé *pks*. Elle provoque des cassures double brin de l'ADN, des mutations, d'important remaniements chromosomiques et favorise l'émergence de tumeurs intestinales en modèle murin. Par ailleurs, les *E. coli* producteurs colonisent fréquemment les tumeurs de patients atteints de cancer colorectal. Nos travaux montrent que les bactéries productrices de colibactine induisent la sénescence cellulaire et stimulent de façon indirecte la prolifération cellulaire *in vitro* et la croissance tumorale *in vivo*. L'action pro-proliférative des cellules rendues sénescents par les *E. coli* producteurs de colibactine est liée à la production du facteur de croissance HGF. L'étude de la signalisation cellulaire responsable montre l'implication du facteur de transcription c-Myc, l'activation de la transcription d'un microARN qui en ciblent la peptidase SENP1, et une modification de la SUMOylation des protéines de l'hôte, notamment p53, un effecteur connu de la sénescence cellulaire. Cette voie de signalisation et les transcripts codant HGF ont été analysés dans des tumeurs de patients atteints de cancer colorectal colonisés ou non par des *E. coli* producteurs de colibactine. Les résultats obtenus soutiennent les résultats obtenus *in vitro* et dans le modèle murin. L'ensemble suggère que les bactéries productrices de colibactine favorisent l'émergence d'un micro-environnement tumoral sénescents susceptible de favoriser la croissance tumorale via la sécrétion de HGF. En parallèle, nous avons caractérisé sur le plan structural et fonctionnel la protéine ClbP de l'îlot *pks*. Les résultats obtenus montrent que ClbP est une peptidase à serine active dont le site actif est extracytoplasmique et indispensable à l'activité biologique de l'îlot *pks*. Des inhibiteurs « drug-like » de ClbP ont été identifiés à l'aide d'approches structurales, biochimiques, cellulaires et microbiologiques. Ces molécules se lient au site actif de ClbP avec une affinité nanomolaire et bloquent les activités génotoxiques et pro-tumorales des *E. coli* producteurs de colibactine. ClbP constitue donc une cible thérapeutique potentielle permettant de bloquer les effets délétères des *E. coli* producteurs de colibactine.

Mots clés : colibactine, cancer colorectal, carcinogenèse, peptidase, inhibiteurs,

Abstract: The colibactin toxin is widely distributed in *Escherichia coli*. Its synthesis is performed by enzymes encoded by the genomic island *pks*. It causes DNA double-strand breaks, mutations, chromosomal rearrangements in host cells and contributes to tumorigenesis in a mouse model. In addition, colibactin-producing *E. coli* are frequently isolated from tumors of patients with colorectal cancer. Our work shows that colibactin-producing bacteria induce cellular senescence and, consequently, can indirectly stimulate cell proliferation *in vitro* and tumor growth *in vivo*. The pro-proliferative effect mediated by these senescent cells is due to the secretion of growth factors, in particular HGF. The cell signaling responsible for cellular senescence shows the involvement of the transcription factor c-Myc, the transcription of a microRNA targeting the peptidase SENP1, and a modification of protein SUMOylation, including p53, a well-known effector of cellular senescence. This signaling pathway and HGF-encoding transcripts were analyzed in tumors of patients with colorectal cancer colonized or not by colibactin-producing *E. coli*. The results support the findings obtained *in vitro* and in the mouse model. Taken together, the results suggest that, in tumors, colibactin-producing bacteria promote the emergence of a senescent microenvironment, which can stimulate tumor growth via the secretion of HGF. In parallel, we determined the structure and function of the *pks*-encoded protein ClbP. The results show that ClbP is an active serine peptidase, whose active site is extracytoplasmic and essential to the biological activity of *pks* island. "Drug-like" inhibitors of ClbP were identified using structural, biochemical, cellular and microbiological approaches. These molecules bind to the active site of ClbP with nanomolar affinity and block the genotoxic and pro-tumoral activities of colibactin-producing *E. coli*. ClbP is therefore a potential therapeutic target to block the deleterious effects of bacteria-producing colibactin.

Key-words: colibactin, peptidase, inhibitors, tumor-promotion