



HAL
open science

Structure de déchets lignocellulosiques : effets sur la colonisation, les communautés microbienne et les performances de méthanisation, caractérisés par des approches fonctionnelles et haut-débit

Nelly Badalato

► To cite this version:

Nelly Badalato. Structure de déchets lignocellulosiques : effets sur la colonisation, les communautés microbienne et les performances de méthanisation, caractérisés par des approches fonctionnelles et haut-débit. Sociologie. AgroParisTech, 2014. Français. NNT : 2014AGPT0002 . tel-04143976

HAL Id: tel-04143976

<https://pastel.hal.science/tel-04143976v1>

Submitted on 28 Jun 2023

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.



Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

L'Institut des Sciences et Industries du Vivant et de l'Environnement (AgroParisTech)

Spécialité : Sciences de l'environnement

présentée et soutenue publiquement par

Nelly BADALATO

le 30 janvier 2014

Structure de déchets lignocellulosiques : effets sur la colonisation, les communautés microbiennes et les performances de méthanisation, caractérisés par des approches fonctionnelles et haut-débit

Directeur de thèse : **Jean-Jacques PERNELLE**

Co-encadrement de la thèse : **Ariane BIZE**

Jury

Mme Pascale MOSONI, Chargée de Recherche, INRA
M. Mathieu SPERANDIO, Professeur, INSA
M. Jean-Jacques GODON, Directeur de Recherche, INRA-LBE
Mme Marion LECLERC, Chargée de Recherche, INRA
M. Théodore BOUCHEZ, ICPEF, Irstea
M. Jean-Jacques PERNELLE, Directeur de Recherche, Irstea
Mme Ariane BIZE, IPEF, Irstea

Rapporteur
Rapporteur
Examinateur
Examinateur
Membre invité
Directeur de thèse
Co-encadrante

RESUME

La méthanisation des composés lignocellulosiques présente un fort intérêt en raison de leur haut potentiel énergétique et de leur abondance, notamment dans les ordures ménagères résiduelles. Toutefois, leur complexité de structure et de composition rend ces matériaux difficilement dégradables en conditions anaérobies et l'utilisation de prétraitements est généralement requise afin d'améliorer leurs rendements de biodégradation. Outre l'effet de ces prétraitements sur la biodégradation de ces composés, la colonisation des lignocelluloses par les micro-organismes cellulolytiques est une étape clé pour l'efficacité de sa dégradation. Dans ce cadre, le travail de thèse a pour objectifs de mieux comprendre le déterminisme de la colonisation de déchets, d'établir le lien entre la colonisation des déchets lignocellulosiques et l'efficacité de leur dégradation et enfin de caractériser plus finement les mécanismes et interactions mises en jeu au sein de la biomasse. Afin de répondre à ces questions, une approche transversale a été développée, combinant des modèles de cultures de souches pures et des systèmes de méthanisation en laboratoire par des communautés complexes. Des approches intégratives ont été appliquées à l'étude de ces systèmes, couplant des analyses haut-débit (métagénomique/(méta)protéomique), un suivi physico-chimique de la biodégradation et des caractérisations physico-chimiques des composés lignocellulosiques étudiés. L'ensemble des résultats met en évidence le rôle des propriétés chimiques, micro- et macro-structurales des composés lignocellulosiques dans leur récalcitrance, leur performances de dégradation et la réponse du compartiment microbien.

La réalisation de la première étude de protéomique totale et quantitative sur la souche pure cellulolytique *Clostridium cellulolyticum*, modèle des *Clostridia* cellulolytiques mésophiles, a permis de mettre en évidence que la vitesse maximale de biodégradation du mouchoir en papier est supérieure à celle du coton et que cette dégradation est associée à un profil métabolique particulier, à une colonisation plus rapide et plus étendue et à une modulation quantitative du système cellulosique. D'autre part, une étude sur un système plus réaliste pour l'étude de la méthanisation des déchets lignocellulosiques a confirmé la bonne concordance entre ce système et le système modèle utilisé et a également permis de mettre en évidence les effets substrats sur la structure des communautés microbiennes avec la dominance de la classe *Bacteroidia* en présence de mouchoir en papier et la forte proportion de la classe *Spirochaetes* en présence de coton. Enfin l'étude des effets de broyages très fins de la paille de blé et du carton plat ont mis en évidence les limites de ces prétraitements sur les performances de leur dégradation, avec l'effet positif modéré du broyage fin de la paille. Ils ont également montré la sensibilité des communautés microbiennes aux changements de surface du

substrat, qui se manifeste par l'émergence de communautés parfois différentes en fonction du prétraitement mécanique appliqué.

En conclusion, ce travail a permis de traiter sous un angle nouveau les questions liées à la récalcitrance des déchets lignocellulosiques en abordant à la fois les aspects structuraux, écologiques et fonctionnels. Ces résultats alimentent le corps de connaissances fondamentales sur les bioprocédés. Ils confirment que les matériaux lignocellulosiques sont particuliers parmi les déchets non-dangereux et qu'une exploitation plus large de leur potentiel énergétique nécessiterait la mise en œuvre de procédés spécifiquement adaptés.

Mots-clés : colonisation de la cellulose, communautés microbiennes, déchets, lignocellulose, méthanisation, métaprotéomique, prétraitements.

ABSTRACT

Lignocellulosic materials have a high energy potential and are abundant, especially in municipal solid waste and their methanization is a promising waste-to-energy bioprocess. However, owing to their highly complex and heterogeneous structure, they are recalcitrant to anaerobic conditions and the use of pre-treatments is usually required to improve their biodegradation yields. Besides, lignocellulose colonization by cellulolytic microorganisms is a key step for an efficient biodegradation. In this context, the PhD work aimed to better understand the factors affecting waste colonization, to establish the link between lignocellulosic waste colonization and its biodegradation efficiency and to characterize more precisely the mechanisms and interactions within the biomass. A transversal approach was developed, combining cultures of model pure strains and lab-scale methanization microcosms with a complex biomass. Integrated approaches were applied to these studies, combining high-throughput analyses (metagenomics/(meta) proteomics), physico-chemical monitoring of bioconversion and finally physico-chemical characterization of substrates. The main results highlight the important role of lignocellulosic materials chemical and micro- and macro -structural features for their recalcitrance, their biodegradation efficiency and the response of the microbial compartment.

The first global quantitative proteomic study on the cellulolytic model *Clostridium cellulolyticum* was conducted. Results showed an increased biodegradation rate of the facial tissue compared to cotton. This enhanced biodegradation was associated to a particular metabolic profile, a faster and more extensive colonization and finally a quantitative modulation of the cellulosic system. On the other hand, study of lignocellulosic waste methanization confirmed the good agreement between this more realistic system and the above-described model system. It also provided new information about the effects of substrate on microbial community structure. Noticeably, *Bacteroidia* members predominated in the presence of tissue and a high proportion of *Spirochaetes* members was observed in the presence of cotton. Finally, study of the effects of wheat straw and cardboard dry grinding revealed the limitations of these pretreatments on biodegradation efficiency. Main key points were a moderate positive effect of wheat straw fine grinding, and the sensitivity of the microbial communities to substrate surface characteristics, as evidenced by the emergence of different microbial communities according to the applied mechanical pretreatment.

In conclusion, this work brings new perspectives to the study of lignocellulosic waste recalcitrance by addressing both the structural, functional and ecological aspects. These results contribute to the core fundamental knowledge on bioprocesses. They confirm that the lignocellulosic materials are specific among non- hazardous waste and require the implementation of adapted specific processes.

Keywords : anaerobic digestion, cellulose colonization, lignocellulose, lignocellulose pretreatments, metaproteomics, microbial communities, municipal solid waste.

REMERCIEMENTS

Ce projet n'aurait pu être mené à bien sans le concours d'Irstea et de la région Ile-de-France que je remercie. J'adresse mes profonds remerciements à Ariane Bize pour m'avoir encadrée pendant ces trois années de thèse. Merci de ce soutien et cette confiance ainsi que pour toute l'aide apportée à la réalisation de ce travail. Je remercie aussi Théodore Bouchez pour m'avoir accueillie dans son équipe et d'avoir soutenu ce travail tout au long de ces trois années. Mes remerciements vont également à Jean-Jacques Pernelle d'avoir été mon directeur de thèse et pour la relecture de ce manuscrit.

Je remercie les membres du jury, Jean-Jacques Godon, Pascale Mosoni, Mathieu Sperandio et Marion Leclerc d'avoir évalué ce travail.

Egalement les membres du comité de pilotage de la thèse, Fabrice Béline, Nicolas Bernet, Romain Briandet, Mickaël Desvaux, Alain Guillot, Anne-Marie Pourcher et Xavier Rouau que je remercie d'avoir guidé ce travail, pour leur aide et leurs conseils. Merci encore d'avoir assisté à toutes les réunions malgré la distance, la neige ou les grèves. Je remercie également Laurent Mazéas et Olivier Chapleur d'avoir suivi ce travail et pour m'avoir conseillée.

Tous mes remerciements aux membres de l'équipe BIOMIC. A Celine Madigou et Chrystelle Bureau pour leur aide, leurs conseils et leur disponibilité. Laetitia Cardona et Carolina Hoyos pour leur amitié, leur aide et les journées marathon au labo. Elie Desmond pour m'avoir aidé à décrypter la bioinfo, Charlotte Richard et Hugues Carlei pour m'avoir initiée à la préparation des boues de digesteur et des réacteurs, Arnaud Bridier pour ses conseils et son aide en microscopie, Gregory Marandat pour m'avoir transmis son expérience du séquençage et Liping Hao pour m'avoir fait découvrir la culture chinoise. Je n'oublie pas l'équipe Chimie que je remercie, en particulier Nina Pourrette et Angeline Guenne pour leur disponibilité et leur efficacité même à la dernière minute. Merci également à Anne Goubet et Lauriane Juzan pour leurs conseils en microbiologie.

Ma reconnaissance s'adresse également à Roger Martins pour sa bonne humeur et ses dépannages informatiques, Nathalie Camus pour l'organisation in extremis des visioconférences et toutes les équipes des services généraux. Je remercie également tous les HBANIens avec qui j'ai partagé un bureau ou un coin de paillasse, un protocole ou un goûter et sans qui ces années n'auraient pas été aussi agréables.

Je n'oublie pas toutes les personnes qui m'ont apportées leur aide durant la réalisation de ce travail. En particulier Diana Sousa et Andreia Salvador pour m'avoir transmis leur expérience de la culture de

souches pures ainsi que toutes les personnes de l'équipe BRIDGE de l'université de Minho au Portugal qui m'ont si bien accueillie : Obrigado !

Je remercie également tous les partenaires pour leur collaboration, en particulier Thierry Meylheuc pour la réalisation des photographies en microscopie électronique, Gérard Mortha et Marie-Noëlle Bellon-Fontaine pour leur disponibilité et pour m'avoir initiée à d'autres disciplines.

Je remercie également ma famille et mes amis proches pour leur appui constant.

SOMMAIRE

Publications et communications	14
Abréviations	16
Introduction générale	18
SYNTHESE BIBLIOGRAPHIQUE	20
I. Structure et composition des matériaux lignocellulosiques	22
I.1. Structure et composition des lignocelluloses.....	22
I.1.1. La cellulose.....	23
I.1.2. Les hémicelluloses	24
I.1.3. Les lignines.....	25
I.2. Les matériaux fabriqués à partir de lignocelluloses	25
II. Ecosystèmes de dégradation des lignocelluloses et techniques d'études	27
II.1. Ecosystèmes anaérobies de bioconversion des lignocelluloses.....	29
II.1.1. Le plus petit écosystème anaérobie modèle : le tube digestif des termites	30
II.1.2. Un autre écosystème anaérobie modèle : le rumen	32
II.1.3. Digesteurs anaérobies à fortes concentrations en lignocelluloses	34
II.2. Aperçu des techniques d'études de la diversité taxonomique et fonctionnelle des écosystèmes. 35	
II.2.1. La culture en souches pures	36
II.2.2. Les techniques d'écologie moléculaires non haut-débit	36
II.2.1. Les techniques haut-débit	39
III. Aspects moléculaires de la dégradation anaérobie de la cellulose	40
III.1. Bactéries cellulolytiques anaérobies isolées	40
III.2. Systèmes enzymatiques de dégradation anaérobies des lignocelluloses	42
III.2.1. Diversité et mécanismes enzymatiques	42
III.2.2. Les cellulosomes : des complexes multi-enzymatiques particuliers	46
III.3. <i>C. cellulolyticum</i> : modèle des <i>Clostridia</i> cellulolytiques mésophiles.....	49
III.3.1. Compléments sur le cellulosome de <i>C. cellulolyticum</i>	49
III.3.2. Le catabolisme de la cellulose chez <i>C. cellulolyticum</i>	51
IV. Adhésion et dynamiques de colonisation des composés lignocellulosiques	54
IV.1. Colonisation des substrats lignocellulosiques.....	54
IV.1.1. Exemples d'études réalisées sur souches pures.....	55
IV.1.2. Exemples d'étude en biomasses complexes	57
IV.2. Adhésion bactérienne à la cellulose.....	58
V. Prétraitements des composés lignocellulosiques	60
V.1. Aperçu des stratégies de prétraitements.....	60
V.2. Les prétraitements dans un contexte de méthanisation	62
VI. Organisation du travail de thèse	64
RESULTATS EXPERIMENTAUX	68
Chapitre 1	70
Avant-propos	70
I. Dry grinding of lignocellulosic substrates shapes microbial communities during methanization 74	
Abstract	74

Keywords	74
Introduction.....	74
Materials and methods	76
Substrate characterizations and grinding	76
Methanization microcosms.....	76
Analytical methods	77
Microbial community analyses	77
Results and discussion.....	78
Waste characteristics.....	78
Anaerobic digestion performance	81
Microbial community structures.....	83
Conclusions.....	88
Acknowledgments	88
Points-clés.....	90
Chapitre 2.....	92
Avant-propos	92
Compléments sur les aspects méthodologiques des analyses protéomiques	95
II. Integrated quantitative whole-proteome analyses of the biofuel-producing <i>Clostridium cellulolyticum</i>.....	98
Abstract	98
Keywords.....	98
Introduction.....	99
Results	100
Substrate characteristics.....	100
Colonization and metabolic patterns.....	102
Overview of the proteome datasets	106
Label-free protein quantification.....	108
Gene set enrichment analyses of functional categories.....	109
Carbohydrate-active enzymes (CAZymes)	111
Central carbon metabolism.....	113
Bacterial adhesion and chemotaxis	116
Discussion	116
Materials and methods	119
Substrate materials and characterizations	119
Strain and culture conditions.....	120
Imaging.....	120
Physico-chemical analyses of the incubation samples.....	121
Total proteome and exoproteome extraction and MS/MS analyses.....	121
Protein database search and data analysis.....	122
Label-free quantification.....	123
Statistical analyses of the label-free quantification proteome datasets	123
Acknowledgments	125
Supplementary materials	125
Point-clés	136
Complément concernant les expériences réalisées sur cultures de souches pures	138
Chapitre 3.....	142
Avant-propos	142
III. Overview of methanization from distinct cellulosic materials based on comparative metaproteogenomics	146
Abstract	146

Keywords	146
Introduction.....	146
Materials and methods	147
Reactor setup and analytical methods.....	147
Carbon mass balance calculations	148
Microbial community analyses	148
Metagenomic analyses	149
Metaproteomic analyses	149
Results and discussion	151
General overview of the experimental design.....	151
Carbon mass balance over the course of degradation	153
Microbial community structures.....	154
Metagenomic analyses for filter paper and facial tissue	158
Coupled metaproteogenomics analyses.....	161
Targeted functional analyses	166
Conclusions.....	169
Acknowledgments	169
Points-clés.....	170
DISCUSSION GENERALE	172
I. Origine de la récalcitrance	174
II. Impact de la récalcitrance sur le compartiment microbien	176
II.1. Effets de la récalcitrance sur la structure des communautés	176
II.2. Effets de la récalcitrance sur la colonisation microbienne.....	177
III. Éléments sur les stratégies de (pré)traitement des déchets.....	179
CONCLUSION ET PERSPECTIVES	180
Références.....	184

LISTE DES FIGURES

Figure 1. Structure et composition des parois végétales et des principaux biopolymères.	23
Figure 2. Schéma général de la bioconversion des lignocelluloses en conditions anaérobies.	29
Figure 3. Aperçu des techniques d'étude de la diversité taxonomique et fonctionnelle.	37
Figure 4. Structure modulaire de quelques cellulases.	42
Figure 5. Exemples de mécanismes d'hydrolyse de la liaison β -glycosidique.....	44
Figure 6. Structure générale des cellulosomes.	46
Figure 7. Modèle présentant les mécanismes de régulation des CAZymes chez <i>C. cellulolyticum</i>	50
Figure 8. Schéma simplifié du transport et de la fermentation des produits d'hydrolyse de la cellulose chez <i>C. cellulolyticum</i>	51
Figure 9. Schéma de la formation d'un biofilm pendant la dégradation de cellulose.	55
Figure 10. Stratégie expérimentale mise en place pendant la thèse.	65
Figure 11. Aperçu de quelques fractions de paille de blé obtenues après broyage par voie sèche.	71
Figure 12. Dispositif expérimental réalisé pour l'étude de la dégradation de paille de blé et de carton plat broyés par une biomasse complexe.....	71
Figure 13. Dispositif expérimental réalisé pour l'étude de la dégradation de trois substrats cellulosiques par <i>C. cellulolyticum</i>	93
Figure 14. Aperçu du schéma expérimental pour l'analyse protéomique.	95
Figure 15. Suivi de la production d'acétate (gauche) et composition finale du biogaz (droite) lors de la co-culture de <i>M. barkeri</i> avec <i>C. cellulolyticum</i> ou <i>A. cellulolyticus</i>	138
Figure 16. Suivi de la croissance de <i>M. barkeri</i> en co-culture avec une bactérie cellulolytique par PCR quantitative.	139
Figure 17. Dispositif expérimental réalisé pour l'étude de la dégradation de trois substrats cellulosiques par une biomasse complexe.	142
Figure 18. Aperçu du schéma expérimental pour l'analyse métagénomique.	143
Figure 19. Photos de colonisation de mouchoir en papier, papier filtre et coton par une biomasse complexe obtenues après 8 jours d'incubations.	144
Fig. I-1. Scanning electron micrographs of wheat straw (A-F) and cardboard (G-J) grinded batches. .	80
Fig. I-2. Cumulative methane production during anaerobic batch experiments.	81
Fig. I-3. PCA biplots of bacterial (left) and archaeal (right) ARISA profiles from day 0 and 31.	83
Fig. I-4. Overview of bacterial (top) and archaeal (bottom) diversity from 16S rDNA pyrosequencing.	86
Fig. I-5. Quantitative real-time PCR of bacterial and archaeal (top) and <i>Bacteroidetes</i> and <i>Firmicutes</i> populations (bottom).	87
Fig. II-1. Molecular weight density distributions of the cellulosic materials.....	102
Fig. II-2. Illustrative microscopy images of substrate colonization.	104
Fig. II-3. Gas and liquid phase analyses of the incubations related to the fermentation metabolic patterns.	106
Fig. II-4. Overview of the proteomic results.	108
Fig. II-5. Heat map of CAZymes showing significant substrate or substrate by time interaction effects in the pellet or supernatant fraction.	112
Fig. II-6. Individual proteins with significant substrate or substrate-time interaction effects mapped over <i>C. cellulolyticum</i> central carbon metabolism and fermentation pathways adapted from the Biocyc website.	116
Fig. III-1 Carbon mass balances for tissue (A), paper (B) and cotton (C) incubations.	153

Fig. III-2. ARISA profiles obtained with a primer set targeting the bacteria.	154
Fig. III-3. Taxonomic distributions obtained by 16S rRNA gene tag pyrosequencing, at the genus level.	156
Fig. III-4. COG repartition of the annotated proteins based on the MG-RAST metagenome analysis results. A: tissue (T3). B: paper (P1).	159
Fig. III-5. Taxonomic composition based on the MG-RAST metagenome analysis results (Identified proteins). A: tissue (T3). B: paper (P1).	160
Fig. III-6. Molecular weight profiles of the proteins extracted from sample pellets.	161
Fig. III-7. Venn diagrams of the identified proteins for each substrate type (panel A, results from the replicates were pooled) and for each replicate (panels B to D).....	163
Fig. III-8. Number of non redundant proteins from each condition retrieved at each step along the MG-RAST pipeline.....	164
Fig. III-9. Number of non redundant proteins from each condition related to the <i>Spirochaetes</i> class and the <i>Ruminococcaceae</i> family.....	165
Fig. III-10. FISH-CLSM biofilm images from filter paper (left) and facial tissue (right) incubations at day 18.....	166
Fig. III-11. Quantification from the extracted DNAs of genes related to taxonomy or to specific gene functions.....	167
Fig. SII-1. Infrared spectra obtained for cotton, facial tissue and filter paper.	125
Fig. SII-2. Gel-like images of proteins from total proteomes (A) and exoproteomes (B).....	134
Fig. SII-3. Proteins with a significant effect based on AUC and SC approaches.	134

LISTE DES TABLEAUX

Tableau 1. Caractéristiques de quelques bactéries cellulolytiques anaérobies isolées et séquencées.	40
Tableau 2. Caractéristiques de quelques biofilms de bactéries cellulolytiques anaérobies isolées.....	56
Tableau 3. Effets de quelques facteurs physico-chimiques sur l'adhésion de souches pures cellulolytiques à la cellulose.	58
Table I-1. Raw materials characteristics.....	79
Table I-2. Batch reactors and particle characteristics.	79
Table I-3. Final methane and degradation yield obtained in batch experiments with the grinded substrates.	82
Table II-1. Summary of the substrate's main characteristics.....	101
Table II-2. Statistical indicators related to peptide and protein identification.	107
Table II-3. Gene Set Enrichment Analyses (GSEA) for major functional categories.	110
Table III-1. Overview of the experimental design.	152
Table III-2. Number of identified OTUs and diversity index.....	155
Table III-3. MG-RAST metagenome analysis indicators.....	158
Table III-4. Comparison of protein identified with two different databases	162
Table III-5. COG classification of the sequences from the metagenomic database used for protein identification.	162
Table SII-1. Quantitative and statistical results from the label-free quantitative approaches based on AUC and SC for all the identified <i>C. cellulolyticum</i> CAZymes.	126

PUBLICATIONS ET COMMUNICATIONS

Conférences d'audience internationale avec comité de sélection

Badalato, N., Pourette, N., Mazéas, L., Bouchez, T., Bize, A. Microbial community shifts associated to the methanization of cellulosic waste with distinct properties. 13th World Congress on Anaerobic Digestion, Santiago de Compostela (Espagne), Juin **2013**. Poster et acte*

Rincker, M.-N., Diara, A., Peu, P., **Badalato, N.**, Girault, R., Carrère, H., Bassard, D., Paus, A., Ribeiro, T., Béline, F. Anaerobic respirometry as a tool to evaluate the effect of pretreatment on anaerobic digestion efficiency. 13th World Congress on Anaerobic Digestion, Santiago de Compostela (Espagne), Juin **2013**. Communication orale et acte*

Badalato, N., Bridier, A., Bouchez, T., Bize, A. Biofilm dynamics on lignocellulosic waste during municipal solid waste methanization. Biofilms 5, Paris (France), Décembre **2012**. Poster

Badalato, N., Rouau, X., Mazéas, L., Bouchez, T., Bize, A. Effect of lignocellulosic waste dry grinding on methanization and microbial populations. 8th International ORBIT Conference, Rennes (France), Juin **2012**. Communication orale et acte

*Actes avec comité de lecture

Conférences d'audience nationale avec comité de sélection

Bize, A., Fan, L., Guillot, A., **Badalato, N.**, Cardona, C., Monnet, V., Madigou, C., Bureau, C., Chapleur, O., Mazéas, L., He, P., Bouchez, T. Approches métabolomiques pour la caractérisation des bioprocédés de valorisation des déchets : présentation de la méthodologie et des résultats obtenus pour la digestion thermophile de déchets celluloses. AdebioTech – BiomInnov, Romainville (France), Décembre **2013**. Poster

Badalato, N., Castella, C., Pons, C., Bouchez, T., Bize, A. Evaluation de la colonisation de déchets lignocellulosiques par des souches pures en lien avec les performances de dégradation. Journées thématiques du Réseau National Biofilm : les biofilms au service des biotechnologies, Narbonne (France), Janvier **2012**. Poster

ABRÉVIATIONS

AGV :	Acides Gras Volatil
ARISA :	Automated Ribosomal Inter Spacer Analysis
AUC :	Area Under Curve
BMP :	Biochemical Methane Potential
CAZy/CAZymes :	Carbohydrate-Active enZymes (database)
CBM :	Carbohydrate Binding Module
CE :	Carbohydrate Esterase
CI :	Crystallinity Index
CID :	Carbone Inorganique Dissous
COD (français) :	Carbone Organique Dissous
COD (anglais) :	Chemical Oxygen Demand
COG :	Clusters of Orthologous Groups
DIC :	Dissolved Inorganic Carbon
DOC :	Dissolved Organic Carbon
DP :	Degree of Polymerization
EC :	Enzyme Comission number
FISH :	Fluorescent In Situ Hybridization
GH :	Glycoside Hydrolase
GSEA :	Gene Set Enrichment Analysis
ISDND :	Installation de Stockage des Déchets Non Dangereux
MSW :	Municipal Solid Waste
SLH :	S-Layer Homology
PL :	Polysaccharide Lyase
qPCR :	quantitative Polymerase Chain Reaction
SC :	Spectral Counting
SDS-PAGE :	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
TS :	Total solids
VFA :	Volatils Fatty Acid
VS :	Volatil solids

INTRODUCTION GENERALE

La méthanisation hors-sol est un bioprocédé de dépollution et de valorisation énergétique des déchets biodégradables qui consiste en une digestion anaérobie de ces déchets. Elle permet la production de biogaz, valorisable sur le plan énergétique (chaleur, électricité), une réduction importante de la masse des déchets et dans certains cas, la fraction résiduelle peut faire l'objet d'une valorisation organique, en épandage. La Directive européenne 2001/77/CE du 27 septembre 2001 stipule que le biogaz est une énergie nouvelle et renouvelable (ENR) et dans une optique de limitation des impacts environnementaux, la loi Grenelle II du 13 juillet 2010 stipule que 23% de l'énergie produite en France devra être d'origine renouvelable d'ici 2020. En 2010, la part du biogaz était de 2,7% des ENR (ATEE, 2011). Les principaux gisements de déchets biodégradables valorisables par méthanisation hors-sol proviennent des ordures ménagères, des stations d'épurations (boues), de certains secteurs industriels (agro-alimentaire, papeterie ...) et du secteur agricole (lisiers, résidus agricoles ...). Le secteur industriel correspond au plus gros gisement traité par méthanisation hors-sol avec plus de 7 millions de tonnes de matières brutes traitées en 2011 (ATEE, 2011).

En France, le gisement des ordures ménagères résiduelles (hors collecte sélective) était de 20 millions de tonnes en 2007 (ADEME, 2009). Actuellement, la grande majorité de ces déchets sont stockés dans des installations de stockage des déchets non-dangereux (ISDND) (avec ou sans valorisation du biogaz), ou incinérés (ADEME, 2009). La méthanisation des ordures ménagères résiduelles (la fraction biodégradable) peut donc contribuer à une stratégie de gestion des déchets adaptée ainsi qu'à la production d'ENR et s'inscrit bien dans le cadre des objectifs environnementaux fixés. En 2011, 9 installations de méthanisation hors-sol des ordures ménagères étaient opérationnelles et 2 étaient en construction (AMORCE, 2011). En pratique, ces installations ne traitent pas uniquement des ordures ménagères et peuvent les traiter conjointement à d'autres déchets, tels que des déchets verts et biodéchets, et des boues de station d'épuration. Ces installations ont traité 200 000 tonnes de déchets en 2010, ce qui ne représente en moyenne que 50% de leur capacité maximale (AMORCE, 2011). La production moyenne d'énergie pour cette même année a été de 36 GWh, contre 114 GWh/an théorique pour l'ensemble des installations (opérationnelles et en construction) (AMORCE, 2011). Il s'agit donc d'une source d'énergie limitée mais non-négligeable avec un fort potentiel de développement.

Outre les problématiques socio-économiques liées à la gestion des déchets, les potentialités de développement de la méthanisation hors-sol pourraient être grandement augmentées par l'amélioration des cinétiques et rendements de dégradation de la fraction lignocellulosique des déchets. En effet, les déchets lignocellulosiques représentent une part importante, jusqu'à plus de 50% (ADEME, 2009) des ordures ménagères, et des déchets agricoles. La masse des déchets lignocellulosiques susceptible d'entrer dans la filière de méthanisation n'a jamais été évaluée précisément. Dans le cas des ordures ménagères, la fraction lignocellulosique correspond principalement aux papiers et cartons, aux textiles sanitaires, dont la proportion augmente fortement, et à quelques déchets verts (ADEME, 2009). Il est aujourd'hui admis que l'hydrolyse de cette fraction constitue une étape limitante lors de la digestion anaérobie des déchets organiques solides (Jeihanipour *et al.*, 2011).

Dans ce contexte général, ce travail de thèse s'insère en partie dans le projet « Codigestion de déchets organiques : optimisation des rendements de dégradation des fractions lignocellulosiques et des conditions de dégradation de micropolluants organiques » financé par le programme R2DS de la région Ile-de-France. Le sujet de thèse traite de la colonisation des déchets lignocellulosiques en conditions anaérobies en lien avec l'efficacité de leur dégradation. En particulier le déterminisme de cette étape sera abordé, notamment concernant les caractéristiques physico-chimiques des substrats avec ou sans prétraitements. L'ensemble du processus sera pris en compte afin de mieux comprendre leurs effets sur l'ensemble de la bioconversion de ces composés.

Les enjeux de la thèse s'inscrivent donc dans un contexte environnemental et économique de fort potentiel pour la valorisation des déchets et également dans un domaine scientifique très actif concernant la digestion anaérobie et la biodégradation des composés lignocellulosiques.

Ce manuscrit est organisé en trois grandes parties et comprend :

- Une synthèse bibliographique de l'état de l'art des domaines scientifiques couverts par le sujet de thèse
- Les principaux résultats expérimentaux présentés sous formes d'articles en préparation pour publication
- Une discussion générale permettant de confronter les conclusions de l'ensemble des résultats expérimentaux et de présenter leurs implications dans une perspective d'application à la méthanisation des déchets lignocellulosiques

SYNTHESE BIBLIOGRAPHIQUE

I. Structure et composition des matériaux lignocellulosiques

Le terme lignocellulose désigne un composé naturel d'origine végétal constitué de multiples espèces chimiques et très structuré. Son abondance, ses propriétés chimiques et structurales ainsi que son rôle dans la nature en font un composé très largement étudié et dont l'exploitation est au cœur de nombreux procédés de fabrication, telle que l'industrie papetière, ou de valorisation, telle que la production d'énergie. La complexité de ce composé et ses propriétés spécifiques le rendent néanmoins difficilement biodégradables, ce qui constitue une limite à sa valorisation par de nombreux procédés biologiques, telle que la méthanisation. Ce chapitre a pour objectif de décrire les principales caractéristiques de ce composé afin de mettre en évidence les origines des enjeux liés à sa biodégradation.

I.1. Structure et composition des lignocelluloses

La paroi des cellules végétales est une matrice complexe principalement constituée de trois biopolymères qui sont la cellulose, les hémicelluloses et les lignines (Figure 1). D'autres espèces chimiques sont aussi présentes, telles que les pectines et d'autres petites molécules (terpénoïdes, stéroïdes ...). Cependant cette composition, et notamment la teneur en chaque composé, est très variable selon l'espèce végétale, le tissu et l'âge du végétal. La teneur globale en cellulose peut par exemple varier entre 30 et 40% chez la plupart des céréales, tandis qu'elle varie entre 40 et 50% chez la plupart des essences d'arbres (Zhao *et al.*, 2012). De plus, les parois des cellules en phase de croissance sont principalement constituées de cellulose, hémicelluloses et pectines, la lignification n'intervenant qu'en fin de croissance dans certains types de tissu. Différentes fonctions biologiques sont associées à ces composés, qui constituent une forme de stockage de l'énergie (cellulose, hémicelluloses et pectines), qui assurent la structure du végétal et apportent une protection contre les agressions extérieures (lignines).

L'agencement et la structuration de ces différents composés au sein des lignocelluloses apparaissent également très complexes. Elles mettent en jeu divers type de liaisons, comme des liaisons ester ou éther entre les lignines et les hémicelluloses et des liaisons hydrogènes entre la cellulose et les hémicelluloses. Un autre niveau de complexité réside également dans la composition et la structure des biopolymères constitutifs.

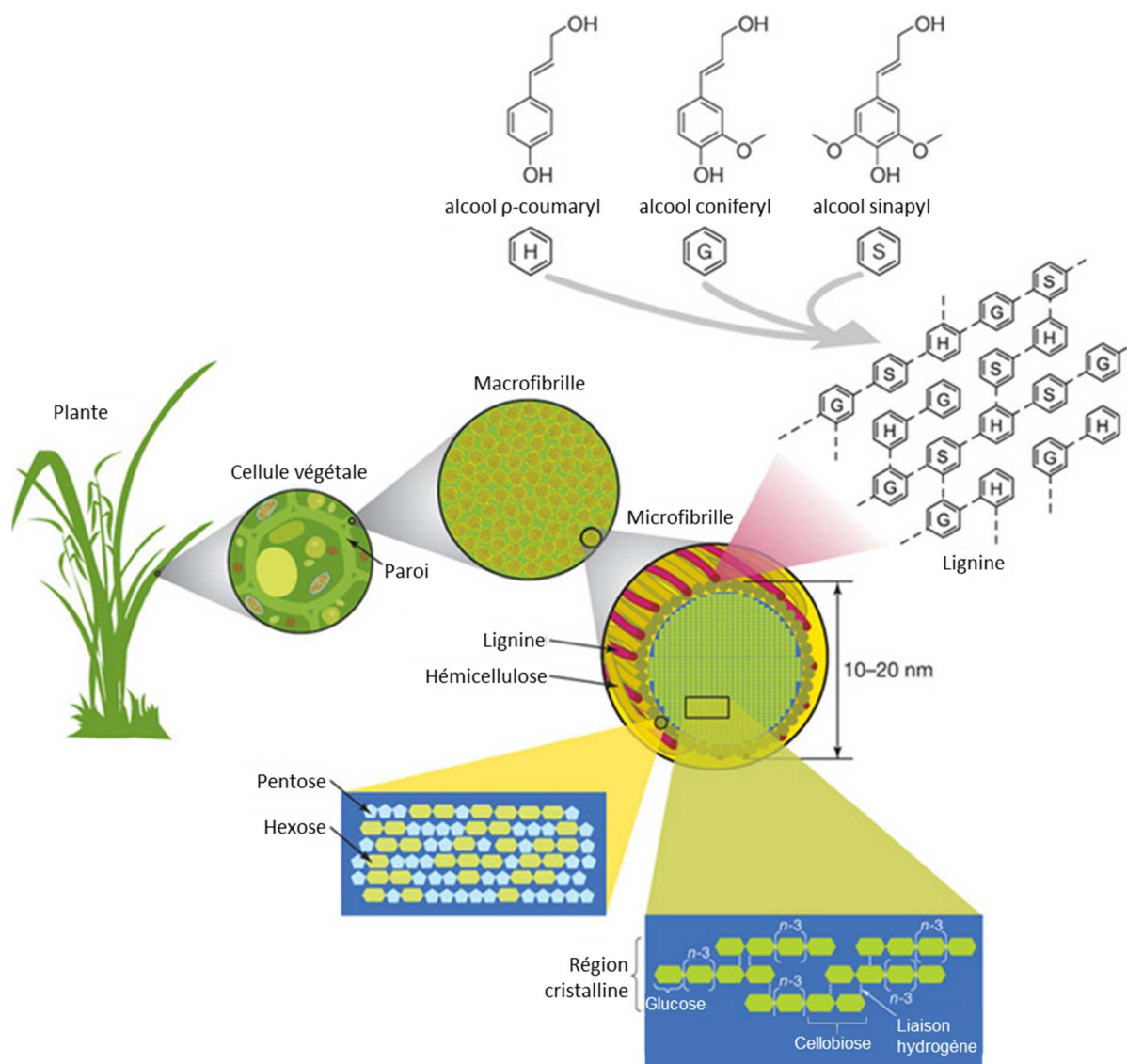


Figure 1. Structure et composition des parois végétales et des principaux biopolymères.
Adapté de Rubin (2008).

1.1.1. La cellulose

La cellulose est le biopolymère naturel le plus abondant en masse dans la nature. Le motif de base est le cellobiose qui est formé par l'association de deux β -D-glucopyranoses dans leur conformation chaise, liés par une liaison β 1-4 glycosidique. Contrairement au cellobiose, la cellulose est insoluble dans l'eau. Plus généralement, les oligomères de glucose unis par des liaisons β 1-4, appelés cellodextrines, sont solubles en milieu aqueux jusqu'à un degré de polymérisation de 7. La cellulose a un degré de polymérisation moyen généralement compris entre 2 500 et 14 000 unités de glucose (McNeil *et al.*, 1984; Xiao *et al.*, 2001). Chaque chaîne de cellulose possède une extrémité réductrice et une extrémité non réductrice. Le caractère réducteur est conféré par l'équilibre hémiacétal – aldéhyde réducteur du carbone anomérique qui est libre, contrairement à l'autre extrémité où ce carbone est engagé dans la liaison glycosidique.

La cellulose possède une structure complètement linéaire. Ses nombreuses fonctions hydroxyles libres permettent la formation de liaisons hydrogènes qui aboutissent à l'association de plusieurs chaînes de cellulose en macromolécules de cellulose. Ces macromolécules s'associent également pour former une microfibrille. Ces microfibrilles s'entrelacent les unes avec les autres en structure caténaire pour former les macrofibrilles, qui s'associent à leur tour pour former une fibre rigide de cellulose. Au sein des microfibrilles, la cellulose peut être présente sous deux formes, amorphe ou cristalline, alternant le long des microfibrilles. La proportion de ces régions amorphes et cristallines varie selon l'origine de la cellulose. Le degré de cristallinité, qui traduit la proportion de régions cristallines, peut varier entre 60 et 90%. La forme cristalline native de la cellulose naturelle est appelée cellulose I et est caractérisée par un arrangement parallèle des chaînes de cellulose. La cellulose cristalline II est caractérisée par une orientation antiparallèle de chaînes de cellulose. Cette dernière forme est la plus thermodynamiquement stable mais est généralement obtenue suite à des traitements (mercerisage) de la cellulose I. Dans la cellulose I, il existe également deux types de cellulose suballomorphe appelées cellulose I α et I β qui sont caractérisées par le type de maille cristalline. La cellulose I β est la plus stable et est la plus représentée dans la cellulose naturelle. Il est également à noter que la cellulose n'est pas un biopolymère exclusivement d'origine végétale et peut également être biosynthétisée par certaines bactéries (Ross *et al.*, 1991).

I.1.2. Les hémicelluloses

Les hémicelluloses regroupent un ensemble de biopolymères. Il s'agit généralement d'hétéropolysaccharides, constitués de différents monomères généralement des pentoses ou hexoses. Les chaînes principales, formées de polymères de xylose (xylanes), de mannose (mannanes) ou de galactose (galactanes), peuvent porter des chaînes latérales. Les hétéropolysaccharides sont alors nommés d'après la nature de ces substituants. L'arabinose (ex : arabinoxylanes), le glucose (ex : glucomannanes) et l'acide glucuronique (ex : glucuronoxylanes) sont par exemple des substituants très répandus parmi les hémicelluloses. Les monomères sont liés entre eux par des liaisons β 1-4, β 1-3 ou β 1-6. En comparaison avec la cellulose, les hémicelluloses ont un degré de polymérisation assez faible, par exemple de 100 à 200 monomères (Xiao *et al.*, 2001). Leur composition varie également en fonction de l'espèce végétale. Ainsi, chez les céréales, les xylanes, qui sont des polymères principalement constitués de chaînes linéaires de xylose liés en β 1-4, sont les plus représentés.

I.1.3. Les lignines

Les lignines sont des biopolymères phénoliques amorphes et hydrophobes. Ils sont constitués de trois monolignols dont la teneur varie en fonction de l'espèce végétale et du tissu. Ces monolignols sont l'alcool coniférylique, l'alcool coumarylique et l'alcool sinapylique et ils sont tous biosynthétisés à partir de l'acide aminé phénylalanine. La polymérisation des lignines s'effectue par la condensation radicalaire des monolignols catalysée par des oxydases. Leur association se fait sans ordre ni répétition notable.

I.2. Les matériaux fabriqués à partir de lignocelluloses

Diverses sources de lignocellulose naturelle sont utilisées pour la fabrication de nombreux matériaux, dont certains sont très abondants dans les déchets ménagers (papiers, cartons ...). En particulier, le bois est la principale source de fibres vierges utilisées par l'industrie papetière. La fabrication de papier et de carton passe tout d'abord par la production des pâtes à papier. Ce procédé consiste à séparer les fibres de bois les unes des autres. En fonction des critères requis liés à l'utilisation des matériaux (blancheur, résistance, opacité) et de la source des fibres, la mise en pâte peut être effectuée par des procédés chimiques, mécaniques ou par une combinaison des deux procédés, les procédés chimico-mécaniques.

Dans les procédés chimiques, les fibres sont individualisées à l'aide de réactifs chimiques, acides (procédés au sulfite) ou alcalins (procédés kraft à la soude et au sulfite de sodium), par dissolution de la lignine. Ces procédés, en particulier le procédé kraft, sont actuellement les procédés majoritaires de production de pâte à papier. Le procédé kraft, qui a pour objectif d'éliminer la lignine, consiste à traiter des copeaux de bois par de la soude à haute température et sous une forte pression. Ces pâtes à papier sont par exemple utilisées pour la fabrication des papiers d'impression et d'écriture.

Les procédés mécaniques utilisés pour la fabrication des pâtes sont nombreux et différents selon le matériau de départ (copeaux ou rondins de bois). Parmi ces procédés, le procédé chimico-thermo-mécanique consiste à réaliser un traitement chimique avec de la soude et du sulfite de sodium ainsi qu'un traitement thermique, avant le défibrage mécanique par un défibreux. Ces pâtes sont principalement employées pour la fabrication de la ouate de cellulose et des cartons plats.

Lors des procédés chimico-mécaniques, les copeaux de bois sont préalablement traités par un agent chimique, avant d'être soumis à un défibrage mécanique. Ces pâtes sont utilisées essentiellement pour la fabrication des cartons ondulés.

Dans certains cas, selon l'usage des matériaux, un procédé de blanchiment est également appliqué à la pâte. Ce procédé a pour objectif d'augmenter la blancheur des pâtes à papier en utilisant des réactifs chimiques.

Ainsi, les propriétés des matériaux lignocellulosiques manufacturés résultent à la fois de la source naturelle employée (bois, coton ...) et des traitements appliqués lors du processus de fabrication.

La structure des composés lignocellulosiques apparaît donc très complexe aussi bien au niveau de l'organisation que de la composition. Nombre des caractéristiques des lignocelluloses, tel que le taux de régions cristallines, le degré de polymérisation des chaînes de cellulose ou la présence de lignine, contribuent à la récalcitrance de ces composés et peuvent être affectés par des prétraitements visant à améliorer la réactivité biologique des composés lignocellulosiques dans les bioprocédés. Cette récalcitrance peut s'exprimer à la fois sur des aspects cinétiques (accessibilité moindre) et/ou sur le rendement (faible biodégradabilité).

La biodégradation des lignocelluloses joue également un rôle important dans le cycle biogéochimique du carbone puisque le carbone peut y être séquestré pendant des centaines d'années. En comparaison avec les liaisons chimiques impliquées dans la formation d'autres biopolymères (ADN, ARN ou protéines), les liaisons glycosidiques impliquées dans la formation des lignocelluloses sont les plus stables (Wolfenden *et al.*, 1998). La demi-vie de la cellulose, à pH neutre et à 25°C en l'absence de catalyseurs, a été estimée à 4,7 millions d'années (Wolfenden *et al.*, 1998). Le rôle des catalyseurs biologiques apparaît donc essentiel pour le recyclage du carbone lignocellulosique. Ces catalyseurs sont à la base de l'utilisation de cette ressource par les organismes, pour la plupart regroupés au sein d'écosystèmes organisés et spécialisés. Une littérature abondante existe sur ces écosystèmes, et ces connaissances disponibles constituent un cadre de référence important pour l'interprétation du travail de thèse.

II. Ecosystèmes de dégradation des lignocelluloses et techniques d'études

La biodégradation des lignocelluloses résulte principalement de l'action de micro-organismes, tels que les champignons, les protistes et les bactéries (Lynd *et al.*, 2002). Pendant longtemps, la capacité de biodégradation des lignocelluloses de certains animaux (ruminants, insectes ...) a été attribuée uniquement aux micro-organismes symbiotiques présents dans leur tractus digestif, mais certains insectes seraient capables de participer eux-mêmes à la déconstruction de ces composés (Watanabe, 2001). Enfin, peu d'espèces d'archées ont à ce jour été décrites comme pouvant participer à la dégradation des lignocelluloses (Blumer-Schuette *et al.*, 2008; Graham *et al.*, 2011). Il est aujourd'hui établi que certains micro-organismes sont spécialisés dans la déstructuration d'une partie des lignocelluloses. Ainsi il est possible de distinguer les micro-organismes cellulolytiques qui sont capables de dégrader la cellulose et les hémicelluloses, les micro-organismes hémicellulolytiques, capables de dégrader les hémicelluloses, et les micro-organismes lignolytiques, capables de dégrader les lignines. Une distinction doit cependant être faite entre la capacité de dégradation et l'utilisation des produits de dégradation. En effet, certains micro-organismes cellulolytiques sont capables de dégrader les hémicelluloses afin d'avoir accès à la cellulose mais n'utilisent pas les glucides solubles libérés par cette hydrolyse comme source de carbone. Tel est par exemple le cas de *Clostridium sufflavum* qui n'utilise pas le xylose et *Clostridium thermocellum* qui n'utilise pas le mannose (Ng *et al.*, 1977; Nishiyama *et al.*, 2009).

Dans l'environnement, diverses espèces de micro-organismes cellulolytiques sont regroupées au sein d'écosystèmes qui hébergent également de nombreux micro-organismes non cellulolytiques. Ces consortia microbiens réalisent, en plus de la déstructuration des lignocelluloses, la bioconversion de ces composés en biomasse, molécules solubles, eau et gaz. Diverses interactions sont alors mises en jeu entre les micro-organismes cellulolytiques et également entre les micro-organismes cellulolytiques et non-cellulolytiques. Ces interactions peuvent être de type synergique, lorsque l'activité d'au moins un des deux partenaires augmente l'activité ou la survie de l'autre partenaire, ou antagoniste, lorsque l'activité d'au moins un des deux partenaires réduit l'activité ou la survie de l'autre partenaire. Parmi les micro-organismes non cellulolytiques jouant un rôle important pour la biodégradation des lignocelluloses figurent les micro-organismes saccharolytiques. Ces derniers, incapables de dégrader directement les lignocelluloses, consomment les glucides solubles libérés grâce à l'action des micro-organismes cellulolytiques (Russell, 1985). Quelques exemples de ces interactions seront détaillés dans la section suivante. Ces écosystèmes ne sont pas figés, et la structure des consortia dépend des interactions entre les organismes mais également des conditions environnementales et de la nature des composés lignocellulosiques digérés (Wu and Conrad, 2001).

Les composés lignocellulosiques étant naturellement très abondants, les biotopes où ils s'accumulent sont nombreux et divers. Cette diversité recouvre des caractéristiques telles que le taux d'humidité, la teneur en oxygène, le potentiel redox, la température, le pH et la biodisponibilité des nutriments (Lynd *et al.*, 2002). Bien que tous ces facteurs aient un impact important sur la structure des communautés qui composent ces écosystèmes, les différences les plus marquées sont observées entre les environnements aérobies et anaérobies. Cet effet est en grande partie lié à la différence des métabolismes mis en jeu et par conséquent à la nature des produits de la bioconversion des lignocelluloses.

Dans les environnements aérobies, la bioconversion des composés lignocellulosiques aboutit à la formation de dioxyde de carbone et d'eau, *via* des réactions d'oxydo-réduction. La couche superficielle du sol est le principal écosystème impliqué (Lynd *et al.*, 2002). En conditions anaérobies, les principaux produits de cette bioconversion sont du méthane, du dioxyde de carbone et des molécules solubles telles que des alcools et acides organiques (Leschine, 1995; Lynd *et al.*, 2002). Les sédiments, le sol et le tractus digestif des animaux supérieurs, principalement les herbivores et les xylophages, sont les écosystèmes anaérobies majeurs (Leschine, 1995). La biodégradation des lignocelluloses en conditions aérobies est généralement plus efficace qu'en conditions anaérobies, en grande partie grâce à une solubilisation plus importante des lignines (Leschine, 1995). A l'échelle globale, la majeure partie des composés lignocellulosiques est dégradée en conditions aérobies (Leschine, 1995). En réalité, la frontière entre environnements aérobies et anaérobies est floue. Les systèmes digestifs des animaux sont par exemple régulièrement alimentés en oxygène, par diffusion depuis la circulation sanguine de l'animal et par les aliments ingérés (Weimer *et al.*, 2009). L'oxygène est rapidement réduit ou consommé par des micro-organismes anaérobies facultatifs, permettant ainsi le retour à des conditions anaérobies (Weimer *et al.*, 2009). Certains micro-organismes anaérobies facultatifs sont saccharolytiques et peuvent donc jouer un rôle direct important dans la biodégradation des lignocelluloses (Russell, 1985).

La suite de ce chapitre sera focalisée sur la description des écosystèmes anaérobie de digestion des lignocelluloses, plus directement en lien avec le travail réalisé pendant la thèse.

II.1. Ecosystèmes anaérobies de bioconversion des lignocelluloses

La bioconversion des lignocelluloses en conditions anaérobies fait intervenir une succession de réactions qui se produisent en continu au sein des écosystèmes (Figure 2). Comme décrit précédemment, la première étape de ce processus fait intervenir les micro-organismes cellulolytiques ou hémicellulolytiques qui dégradent les composés lignocellulosiques en monomères et oligomères solubles. Ces composés sont fermentés en acides et alcools organiques, tels que l'acétate et l'éthanol, et en dihydrogène et dioxyde de carbone, par les micro-organismes cellulolytiques eux-mêmes et par les micro-organismes saccharolytiques. Ces acides et solvants organiques, hors acétate, sont ensuite convertis en acétate, dihydrogène et dioxyde de carbone par les micro-organismes acétogènes. La dernière étape est la production de méthane à partir des produits de l'acétogénèse selon deux voies métaboliques principales. La méthanogénèse, exclusivement réalisée par des archées, peut être acétoclaste lors de la décarboxylation de l'acétate, ou hydrogéntrophe lors de la réduction du dioxyde de carbone. Il est à noter que le travail de thèse sera concentré sur la déconstruction et la fermentation des parties cellulosiques et hémicellulosiques, puisqu'à ce jour les données sur la fermentation des composés issus de la dépolymérisation des lignines en conditions de méthanisation sont très limitées (Kaiser and Hanselmann, 1982).

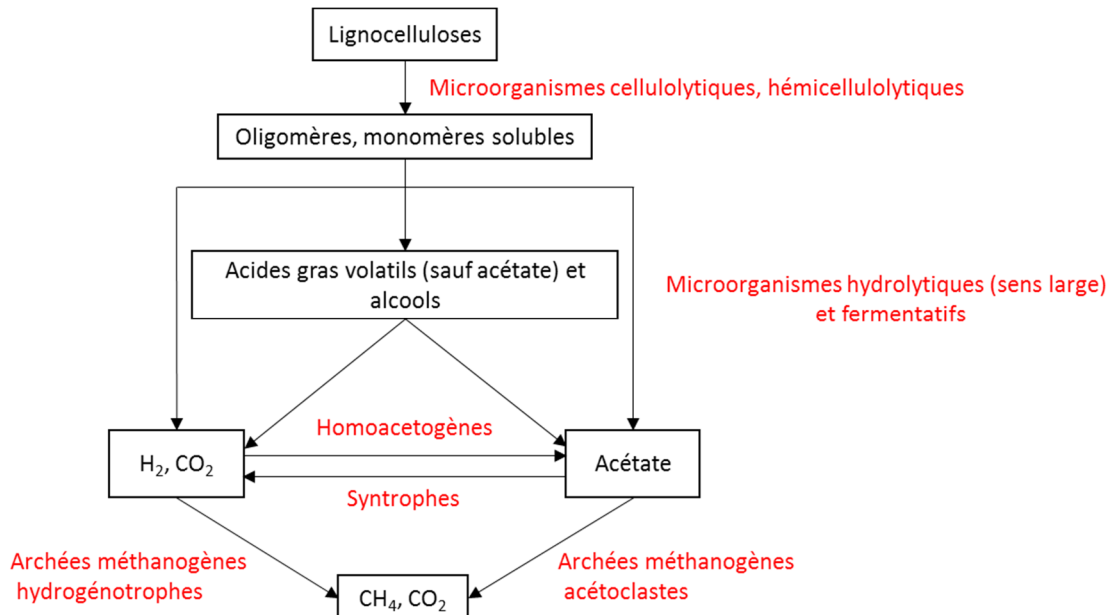


Figure 2. Schéma général de la bioconversion des lignocelluloses en conditions anaérobies.
Adapté de Tsavkelova & Netrusov (2012).

Outre les réactions de fermentation, les réactions de transfert d'électrons inter-espèces sont essentielles au processus de méthanisation. Parmi ces réactions figurent les réactions de transfert d'hydrogène inter-espèces. Dans les environnements anaérobies, l'hydrogène est un élément important puisque son accumulation a un effet inhibiteur sur la production d'acétate et peut conduire à des blocages dans le processus de méthanisation, en particulier lors de la digestion des lignocelluloses (Leschine, 1995; Sasaki *et al.*, 2012). Les réactions de transfert d'hydrogène inter-espèces nécessitent des interactions syntrophiques entre les micro-organismes producteurs et les micro-organismes consommateurs d'hydrogène. Les bactéries homoacétogènes par exemple, consomment le dihydrogène pour réduire le dioxyde de carbone en acétate (Figure 2). Cette réaction est rendue possible grâce à leur interaction syntrophique avec les archées réalisant la méthanogénèse hydrogénotrophe. La consommation continue de dihydrogène par les archées maintient de faibles pressions partielles en dihydrogène nécessaires à la croissance des bactéries homoacétogènes. De nombreuses autres réactions transverses mettant en jeu des interactions syntrophiques, telle que l'oxydation syntrophique de l'acétate, ont également lieu mais ne seront pas détaillées ici (Stams, 1994).

La digestion anaérobie en général et des matériaux lignocellulosiques en particulier est donc un processus complexe dont la compréhension passe notamment par l'étude des écosystèmes, et en particulier des consortia microbiens impliqués. Parmi ces écosystèmes, le tube digestif des termites et le rumen ont fait l'objet de nombreuses études. Plus précisément, les consortia microbiens et les interactions écologiques figurent parmi les mieux caractérisées au sein des écosystèmes associés à la digestion de lignocelluloses et seront présentés ci-après.

II.1.1. Le plus petit écosystème anaérobie modèle : le tube digestif des termites

Le tube digestif des termites est un système hétérogène et très structuré formé de plusieurs compartiments. En particulier, l'intestin moyen et l'intestin postérieur, comprenant la panse, sont des lieux importants de la digestion. Le pH est proche de la neutralité bien que les premières parties de l'intestin postérieur soient très alcalines chez certaines espèces (Brune *et al.*, 1995). Il existe également un gradient d'oxygénation, contrairement au rumen, qui fait de la périphérie de la panse un microenvironnement aérobie. Ce gradient permet la dégradation des composés récalcitrants en conditions aérobies, tels que les composés aromatiques issus de la dégradation des lignines (Harazono *et al.*, 2003) et la bioconversion du dihydrogène en méthane en conditions anaérobies par un consortium microbien complexe. Celui-ci comprend des bactéries, des archées, des champignons et dans certains cas également des protistes. L'intestin postérieur est le compartiment le plus documenté sur le consortium microbien.

Les protistes ne sont pas toujours présents dans les tubes digestifs des termites. Généralement, les termites sont séparés en deux catégories : les termites dits “inférieurs” qui ont gardé des protistes contrairement aux termites dits “supérieurs”. Les protistes sont les plus abondants en terme de biomasse mais moins nombreux que les bactéries en nombre de cellules. Ces protistes sont des flagellés appartenant aux phyla *Parabasalia* et *Preaxostyla* (Ohkuma, 2008). Bien que moins nombreux que les bactéries, les protistes sont un groupe très actifs pour la digestion des composés lignocellulosiques. Plusieurs gènes codant des cellulases et des hémicellulases ont été identifiés (Ohkuma, 2008). Les protistes sont également considérés comme les principaux producteurs de dihydrogène.

Les bactéries sont le groupe microbien le plus important en nombre de cellules. La diversité bactérienne est conservée au sein d’un même genre de termites mais diffère selon les genres. La communauté est largement dominée par les bactéries appartenant au phylum *Spirochaetes* mais les phyla *Fibrobacteres*, *Bacteroidetes*, *Firmicutes* sont également représentés. (Hongoh *et al.*, 2005; Warnecke *et al.*, 2007). Les activités cellulolytique et hémicellulolytique bactériennes des termites “supérieurs” semblent être majoritairement portées par les bactéries appartenant au genre *Treponema* (phylum *Spirochaetes*) (Warnecke *et al.*, 2007). De nouveaux groupes phylogénétiques, tel que le candidat phylum TG1 (Termite Group 1), ont également été détectés chez des termites “inférieurs” et peuvent représenter une part importante de la communauté (Hongoh *et al.*, 2005).

Par ailleurs, des associations symbiotiques entre des bactéries et des protistes ont été décrites. Certaines populations bactériennes sont associées à la surface (épibiontes) ou dans le cytoplasme (endobiontes) des protistes. Ces interactions semblent largement répandues puisque la plus grande proportion de la communauté bactérienne comprend les bactéries associées à un protiste. Parmi les exemples, les bactéries appartenant au candidat phylum TG1 ont été décrites comme des endobiontes en association avec des protistes appartenant au genre *Trichonympha* (phylum *Parabasalia*) et certaines bactéries appartenant au phylum *Spirochaetes* ont été décrites comme des épibiontes en association avec le protiste *Metadevescovina cuspidate* (phylum *Parabasalia*) (Ohkuma, 2008). Le rôle de ces symbiontes n’est pas encore totalement défini mais ils pourraient participer à la motilité du protiste et intervenir dans la consommation de l’hydrogène et la fixation de l’azote (Hongoh *et al.*, 2008).

Enfin, les archées méthanogènes sont également présentes dans le tubes digestif des termites et joueraient un rôle important dans la consommation du dihydrogène. La plupart des archées décrites appartiennent au genre *Methanobrevibacter* (ordre *Methanobacteriales*) et sont aussi décrites comme associées à des protistes, tels que *Dinenympha parva* (phylum *Preaxostyla*) (Tokura *et al.*, 2000).

Les termites participent également à la relation symbiotiques via des structures du tube digestif spécialisées et cette tripartie, termite-protistes-procaryotes, semble être la clé de l'efficacité de la digestion des matériaux lignocellulosiques. Le rumen est également très étudié, en grande partie pour des raisons liées à l'optimisation de la production animale et est donc également présenté en détail ci-après.

II.1.2. Un autre écosystème anaérobie modèle : le rumen

Le rumen, poche pré-gastrique présente chez certains mammifères herbivores, est l'organe dans lequel se déroule principalement la dégradation et la fermentation des composés lignocellulosiques ingérés. Ses principales caractéristiques sont : un milieu anaérobie, une température de 38°C, un pH proche de la neutralité et un temps de rétention des aliments qui varie entre 48 et 72h. Le consortium microbien impliqué, majoritairement des micro-organismes anaérobies stricts, est composé de bactéries, de champignons, de protistes et d'archées. Des virus sont également présents, mais leur diversité et leur rôle étant encore peu connus, ce point ne sera pas abordé dans cette section. La répartition de ces différents groupes microbiens chez un même animal dépend en grande partie de son régime alimentaire (Belanche *et al.*, 2012). Il semble cependant que les bactéries soient majoritaires en nombre de cellules, avec une concentration de l'ordre de 10^{10} cellules/ml de liquide ruminal (Morgavi *et al.*, 2012). Le principal rôle de ce consortium microbien est de dégrader et fermenter les composés lignocellulosiques ingérés par l'animal en fournissant en retour des métabolites solubles tels que l'acétate et le propionate qui serviront de source d'énergie pour l'animal. Ce consortium microbien s'installe dès la naissance, toutefois des changements de communautés s'opèrent entre la naissance et l'âge adulte (Jami *et al.*, 2013).

Les bactéries ont un rôle essentiel, puisque ce groupe est connu comme étant le plus actif dans la digestion des composés lignocellulosiques ou dans les écosystèmes ruminiaux (Lee *et al.*, 2000). Les espèces cellulolytiques dominantes sont *Ruminococcus albus* et *Ruminococcus flavefaciens*, appartenant au phylum *Firmicutes* et *Fibrobacter succinogenes* appartenant au phylum *Fibrobacteres*. Le phylum *Bacteroidetes* est également très représenté dans les écosystèmes ruminiaux : il s'agit majoritairement de bactéries hémicellulolytiques ou saccharolytiques, telles que celles du genre *Prevotella* (Matsui *et al.*, 2000). Le phylum *Spirochaetes*, dominé par des bactéries du genre *Treponema*, joue également un rôle important dans l'utilisation des produits de dégradation des lignocelluloses (Stanton and Canale-Parola, 1980). Une étude récente a montré que les animaux d'un même troupeau (ayant reçu la même alimentation) partagent les mêmes genres bactériens et que la variabilité inter-individuelle de la population bactérienne résidait principalement dans les différences d'abondances de ces groupes (Jami and Mizrahi, 2012). Bien qu'essentielles, les bactéries

cellulolytiques ne représentent qu'environ 10% du total des bactéries du rumen (Russell *et al.*, 2009). L'utilisation des glucides solubles par les bactéries non cellulolytiques développant des interactions synergiques avec les bactéries cellulolytiques, peut en partie expliquer la forte concentrations en bactéries non cellulolytiques (Russell, 1985; Wells *et al.*, 1995). Plusieurs interactions de ce type ont en effet été mises en évidence par des études *in vitro*, comme par exemple l'interaction entre *Prevotella ruminicola* et *F. succinogenes* pour la dégradation des hémicelluloses et entre *Selenomonas ruminantium* et *R. albus* pour la dégradation de la cellulose (Chen and Weimer, 2001; Fondevila and Dehority, 1996). Cependant ces interactions ne sont pas généralisables aux associations entre bactéries cellulolytiques et saccharolytiques, puisqu'aucune synergie n'est par exemple détectée dans l'association entre *P. ruminicola* et *R. albus* (Fondevila and Dehority, 1996). Les espèces cellulolytiques citées peuvent également être impliquées dans des interactions de type antagoniste, dont un exemple est l'inhibition de *R. flavefaciens* par *R. albus* (Chen and Weimer, 2001). Dans ces phénomènes de compétition, les principaux facteurs déterminant sont la spécificité de substrat, la vitesse de croissance ou la production de molécules inhibitrices (Chen and Weimer, 2001; Chen *et al.*, 2004).

Les protistes décrits dans les écosystèmes ruminiaux sont majoritairement des ciliés (Morgavi *et al.*, 2010). Bien que peu nombreux, les protistes sont dominants en termes de biomasse. Le rôle de ces organismes a longtemps été controversé. Il semblerait qu'ils participent à la fois au renouvellement de la biomasse en absorbant de la matière organique, en particulier des bactéries, mais également à la dégradation des lignocelluloses et à la fermentation des glucides solubles (Akin and Benner, 1988; Lee *et al.*, 2000). Ce dernier rôle a été souvent discuté en raison de la difficulté à obtenir des cultures axéniques. L'une des techniques les plus utilisées pour étudier le rôle des protistes est la défaunation, qui consiste à supprimer les protistes du rumen par des méthodes physiques ou chimiques (Mosoni *et al.*, 2011). Les études au cours desquelles cette technique a été utilisée ont toutes montré que l'absence de protistes n'affecte pas la survie de l'animal. Cependant, les effets sur la structure et l'activité de la population microbienne restante ne sont pas clairs (Morgavi *et al.*, 2010).

Les champignons du rumen peuvent également participer activement à la biodégradation des composés lignocellulosiques (Russell *et al.*, 2009; Trinci *et al.*, 1994). En comparaison avec les bactéries, les champignons possèdent le plus fort potentiel fibrolytique puisque leur activité de déconstruction des fibres de lignocelluloses, et en particulier la solubilisation des lignines, est la plus efficace (Kuhad *et al.*, 1997). Néanmoins, les composés issus de la solubilisation des lignines semblent très peu métabolisés (Akin and Benner, 1988; McSweeney and Dulieu, 1994). Les espèces majoritaires appartiennent aux genres *Neocallimastix* et *Piromyces* et de nombreuses interactions sont à nouveau possibles entre ces champignons et les bactéries cellulolytiques. Par exemple, en co-culture, *R. flavefaciens* et *Neocallimastix frontalis* développent une interaction synergique lors des premières

étapes de la biodégradation de la cellulose, suivie par une interaction antagoniste qui aboutit à l'inhibition de *N. frontalis* (Bernalier *et al.*, 1993). En revanche aucune interaction n'a été décrite entre *F. succinogenes* et *N. frontalis* (Bernalier *et al.*, 1993).

Enfin, les archées méthanogènes présentes dans les écosystèmes ruminiaux appartiennent généralement au genre *Methanobrevibacter* et la voie hydrogénotrophe est, de façon cohérente, dominante (Balch *et al.*, 1979; Morgavi *et al.*, 2010). Pour des raisons environnementales (le méthane est un gaz à effet de serre puissant), la production de méthane par ces écosystèmes n'est pas souhaitée et de nombreuses études visent à inhiber la méthanogénèse (Morgavi *et al.*, 2010; Mosoni *et al.*, 2011). Les méthanogènes sont impliquées dans de nombreuses interactions syntrophiques avec les bactéries, les champignons et les protistes producteurs de dihydrogène (Morgavi *et al.*, 2010). Par exemple, il a été montré *in vitro* que les activités de *N. frontalis* et de *R. flavefaciens* sur la cellulose sont augmentées en présence de *Methanobrevibacter smithii* (Rychlik and May, 2000; Teunissen *et al.*, 1992).

Le rumen est considéré depuis longtemps comme un écosystème très efficace pour la digestion des composés lignocellulosiques (Weimer *et al.*, 2009). En plus de la présence d'un consortium microbien adapté et efficace, la rumination, la mastication et l'action de la salive jouent un rôle déterminant dans l'efficacité du rumen. Ces actions mécaniques et chimiques correspondent en fait à des prétraitements des matières lignocellulosiques ingérées. Une section sera dédiée aux effets de tels prétraitements à la fin de ce chapitre. Du point de vue des bioprocédés, le rumen peut être comparé à un réacteur très régulé fonctionnant en semi-continu et en conditions mésophiles (Godon *et al.*, 2013; Weimer *et al.*, 2009). L'efficacité des écosystèmes ruminiaux et leurs caractéristiques physico-chimiques assez facilement transposables, en font un système qui a inspiré et qui continue d'inspirer la conception et la conduite des bioprocédés de digestion anaérobie des composés lignocellulosiques (Bayané and Guiot, 2010; Weimer *et al.*, 2009).

II.1.3. Digesteurs anaérobies à fortes concentrations en lignocelluloses

Contrairement aux écosystèmes ruminiaux, les écosystèmes participant à la dégradation anaérobie des composés lignocellulosiques dans les bioprocédés de méthanisation hors-sol n'ont fait l'objet de recherches intensives que depuis quelques décennies, le consortium microbien ayant pendant longtemps été considéré comme une « boîte noire » et les outils d'écologie moléculaires n'étant pas utilisés si largement qu'actuellement. Contrairement aux études sur le rumen, la plupart des études sur ces consortia sont réalisées en microcosmes (ou mésocosmes) inoculés par des biomasses complexes issues de divers environnements avec des conditions opératoires proches de celles des digesteurs. Les micro-organismes composant ces écosystèmes sont également très variés. Cette diversité dépend de l'origine de l'inoculum, mais est aussi fortement liée aux conditions d'opérations,

incluant la température et le type de substrat digéré, ainsi qu'à l'évolution des paramètres physico-chimiques pendant la dégradation, tels que le pH et la concentration en acides organiques. L'influence précise de ces paramètres ne sera pas détaillée ici (Chapleur, 2012).

Les études réalisées ont montré que lors du traitement de composés lignocellulosiques fortement concentrés en conditions mésophiles, les phyla bactériens *Firmicutes* et *Bacteroidetes* sont les plus représentés (Li *et al.*, 2009; Schlüter *et al.*, 2008; van der Lelie *et al.*, 2012). Au sein du phylum *Firmicutes*, les bactéries sont majoritairement apparentées à des bactéries cellulolytiques, selon des données de phylogénie moléculaire, telles que *Clostridium thermocellum*, *Clostridium cellulolyticum* et *Acetivibrio cellulolyticus*. En conditions thermophiles, les *Firmicutes* dominent largement, avec des représentants cellulolytiques proches de *C. thermocellum* et *Caldicellulosiruptor saccharolyticus*.

Concernant les archées, l'ordre *Methanomicrobiales* est généralement dominant, et plus particulièrement le genre *Methanoculleus*. A l'instar du rumen, des interactions syntrophiques entre bactéries cellulolytiques et archées méthanogènes isolées de ces écosystèmes ont été mises en évidence. Par exemple, la co-culture entre *Clostridium clariflavum* et *Methanothermobacter thermautotrophicus* a montré que l'archée est capable d'utiliser le dihydrogène produit par la bactérie cellulolytique lors de la fermentation de la cellulose, ce qui a permis d'augmenter le taux d'utilisation de la cellulose (Sasaki *et al.*, 2012). Cependant, ces études n'ont pas mis en évidence une présence significative d'eucaryotes dans ces écosystèmes.

Toutes ces études ont nécessité l'utilisation de techniques de biologie adaptées à l'étude d'écosystèmes complexes, qui ont évolué au cours des années et ont permis de mieux caractériser la diversité taxonomique et fonctionnelle de ces écosystèmes.

II.2. Aperçu des techniques d'études de la diversité taxonomique et fonctionnelle des écosystèmes

Un grand nombre de techniques de microbiologie et d'écologie moléculaire ont été mises en œuvre lors du travail de thèse. Afin de pouvoir situer les techniques utilisées dans un cadre plus général, cette section présente un rapide aperçu des techniques les plus utilisées pour l'étude de la diversité taxonomique et fonctionnelle des écosystèmes. Des informations plus détaillées sur chaque technique utilisée pendant le travail de thèse seront abordés dans un autre chapitre, notamment en ce qui concerne la métagénomique et la métaprotéomique.

Ces techniques peuvent être divisées en trois grands groupes, les techniques cultures dépendantes, les techniques d'écologie moléculaire non haut-débit puis les techniques à haut-débit (Figure 3).

II.2.1. La culture en souches pures

L'enrichissement, l'isolation et la culture de micro-organismes ont longtemps été des techniques de prédilection pour étudier leur rôle au sein des écosystèmes. Dans ce cadre, les techniques et milieux développés par Hungate ont joué un rôle déterminant dans la caractérisation phénotypique et métabolique des micro-organismes anaérobies. En effet, les premiers micro-organismes à avoir été cultivés selon ces méthodes sont des bactéries et protistes du rumen (Hungate, 1950). Ces méthodes ont ensuite été appliquées à l'étude d'autres environnements, tels que le sol ou les sédiments, et à d'autres micro-organismes, et ont ainsi permis de caractériser un grand nombre de micro-organismes impliqués dans la digestion anaérobie des composés lignocellulosiques. D'autres techniques, telles que l'observation microscopique et la caractérisation par cytométrie en flux, permettent également d'avoir un aperçu de la structure des écosystèmes. Néanmoins, utilisées sans marqueurs moléculaires, ces techniques permettent principalement de caractériser la diversité de morphotype. Toutes ces techniques ont cependant une importante limite. Il est en effet établi que généralement moins de 1% des micro-organismes présents dans les écosystèmes peuvent être cultivés en particulier les micro-organismes très spécialisés, en raison de la présence d'interactions biotiques ou abiotiques essentielles à leur survie mais encore inconnues ou difficiles à reproduire en laboratoire. Grâce à l'essor des techniques de biologie moléculaire, l'accès à la diversité de ces micro-organismes non-cultivés est possible, et permet également l'étude des micro-organismes dans leur environnement naturel.

II.2.2. Les techniques d'écologie moléculaires non haut-débit

Parmi les outils développés depuis les années 1980, les techniques moléculaires d'analyse de la diversité taxonomique sont souvent basées sur l'étude des séquences des ARN ribosomiques (ARNr) ou de leurs gènes. Pour les bactéries et archées il s'agit des ARNr 16S et pour les eucaryotes généralement des ARNr 18S. L'intérêt de ces marqueurs réside dans leur universalité, leur fonction conservée, leur abondance dans les cellules et leur variabilité de séquence d'une espèce à l'autre, ce qui permet d'estimer les vitesses d'évolution. Ces marqueurs sont utilisés depuis plusieurs décennies et ont significativement contribué à la construction de l'arbre phylogénétique du monde vivant tel qu'il est actuellement représenté (Cavalier-Smith, 2004).

Les techniques d'empreintes moléculaires sont parmi les plus utilisées. Elles permettent d'obtenir des profils de diversité des communautés étudiées. Rapides, elles sont réalisées par amplification par PCR des gènes codant les ARNr, ou de régions entre ces gènes, par migration électrophorétique des

fragments amplifiés et analyse des profils obtenus. Selon la technique utilisée, les fragments sont différenciés par leur variabilité de taille (en paires de bases) ou de séquence. Aujourd'hui beaucoup de ces techniques sont dites automatisées car la migration des fragments se fait par électrophorèse capillaire, plus rapide et plus résolutive que l'électrophorèse traditionnelle. Par exemple, l'une de ces techniques, l'ARISA (Automated Ribosomal Inter Spacer Analysis), est basée sur la variabilité de taille de la région intergénique (ITS) entre les gènes ARNr 16S et 23S (Ranjard *et al.*, 2001). L'amplification de cette région est réalisée grâce à des amorces s'hybridant sur les régions conservées des gènes ARNr 16S et 23S. Ces approches ont néanmoins des limites. La première est commune à toutes les techniques basées sur la PCR et concerne l'introduction de biais d'amplification (spécificité des amorces, efficacité de l'amplification, formation de chimères). De plus, en raison de la variabilité de ces marqueurs au sein d'une même cellule, les fragments amplifiés peuvent posséder des caractéristiques différentes, ce qui aboutit à la formation de multiples signaux pour un même génome. Enfin, la résolution n'est souvent pas suffisante pour distinguer des fragments dont la mobilité électrophorétique est proche. Les profils de communautés obtenus par ces approches constituent donc des empreintes moléculaires d'un écosystème donné à un temps donné et ne permettent pas, seules, d'identifier précisément les micro-organismes présents.

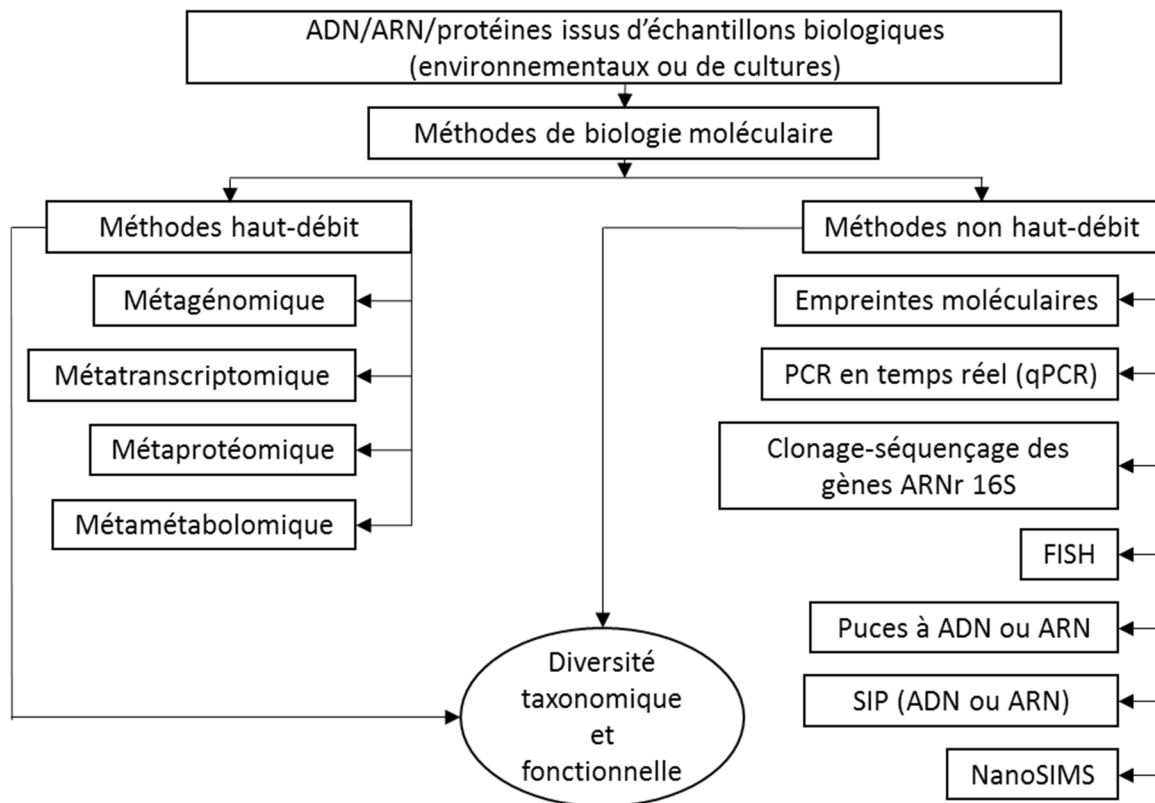


Figure 3. Aperçu des techniques d'étude de la diversité taxonomique et fonctionnelle.

Les techniques d'empreintes moléculaires ne permettant pas de quantifier les groupes microbiens présents, la PCR quantitative (qPCR) s'avère être un outil utile et complémentaire en écologie microbienne. Cette technique est basée sur l'amplification par PCR de gènes, ARNr ou fonctionnels, et la détection en temps réel des fragments amplifiés. Cette détection est rendue possible grâce à l'utilisation de marqueurs fluorescents, liés à une sonde complémentaire d'une partie de la séquence ciblée ou se liant à l'ADN double brin. Deux technologies sont généralement utilisées, respectivement TaqMan et SYBRGreen. La quantification des fragments amplifiés est réalisée grâce à l'utilisation d'une gamme standard pour laquelle les quantités de molécules sont connues. Cette technique peut également être appliquée à la quantification des ARN, en incluant au préalable une étape de rétro-transcription afin d'obtenir les ADN complémentaires (RT-qPCR).

Certaines techniques d'étude de la diversité fonctionnelle basées sur l'utilisation, en amont, d'isotopes stables, telles que le SIP (Stable Isotope Probing) (Li *et al.*, 2009) et l'ensemble des techniques de NanoSIMS (Nano Secondary Ion Mass Spectrometry) (Musat *et al.*, 2011), ne seront pas détaillées ici.

Une autre technique couramment utilisée est le FISH (Fluorescent *in situ* hybridization) (Amann *et al.*, 2001). Contrairement aux autres techniques moléculaires décrites ici, le FISH est réalisé sur des cellules intègres et permet la visualisation d'échantillons préservés par fixation. Le principe de base est l'hybridation d'oligonucléotides de séquence précise, marqués par une molécule fluorescente, à une séquence cible d'ADN ou d'ARN. Le marquage fluorescent de la sonde permet la visualisation et quantification des cellules ciblées par microscopie optique. Les cibles peuvent être les ARNr 16S, pour une identification taxonomique des cellules, ou des gènes de fonctions pour une identification fonctionnelle (Moraru *et al.*, 2010). L'un des atouts de cette technique réside dans la possibilité d'identifier et de quantifier des micro-organismes dans leur milieu naturel et de mettre en évidence leurs interactions avec l'environnement et les autres micro-organismes. Cette approche nécessite cependant de connaître les séquences cibles afin d'utiliser des sondes spécifiques. De nombreuses sondes ciblant des groupes microbiens majeurs ont déjà été développées mais le développement de sondes ciblant des groupes plus spécifiques de micro-organismes est un travail délicat. En plus de la difficulté à concevoir une sonde sensible et spécifique, obtenir un signal de qualité peut également s'avérer problématique, en particulier lorsque les micro-organismes ciblés sont peu actifs ou abondants. Pour contourner cette limitation, de nombreuses variantes de cette technique ont été développées, telles que le CARD-FISH ou le TSA-FISH, en intégrant des étapes d'amplification du signal (Kawakami *et al.*, 2010; Pernthaler *et al.*, 2002).

II.2.1. Les techniques haut-débit

Depuis la dernière décennie et l'essor des technologies à haut-débit, les techniques d'études globales sont devenues des outils majeurs en écologie microbienne. Ces techniques sont représentées par les approches de meta-omique : métagénomique, métatranscriptomique, métaprotéomique et métabolomique. Elles sont nommées d'après les composants cellulaires ou biologiques ciblés, respectivement l'ADN, les ARN, les protéines et les métabolites. Les deux premières techniques sont basées sur le séquençage haut-débit des acides nucléiques, l'annotation fonctionnelle des gènes de fonction ou des ARN messagers, en général par homologie de séquence, et l'assignation taxonomique des séquences d'ADNr 16S ou d'ARNr 16S. Les deux dernières techniques sont basées sur la mesure précise de la masse de molécules par spectrométrie de masse, la détermination de la séquence en acides aminés ou de la formule brute associées à cette masse et l'identification des protéines ou métabolites, en général par comparaison avec les séquences ou formules brutes présentes dans les banques de données. Ces techniques fournissent une vaste quantité d'information, et présentent des complémentarités puisqu'elles ne donnent pas accès aux mêmes informations.

Ainsi la métagénomique permet de réaliser l'inventaire des micro-organismes et des fonctions présents dans un écosystème, sans permettre une discrimination entre les éléments actifs ou inactifs. La métatranscriptomique permet l'étude des micro-organismes actifs, des fonctions activées, potentiellement exprimées, et de leurs mécanismes de régulation. La métaprotéomique et la métabolomique permettent d'étudier respectivement les fonctions réellement exprimées, possiblement liées aux activités métaboliques de l'écosystème, et les voies métaboliques empruntées. Grâce à la précision des instruments, toutes ces méthodes peuvent être utilisées pour des analyses qualitatives ou quantitatives. Les approches quantitatives peuvent toutefois être délicates à mettre en œuvre pour des écosystèmes complexes, en raison de différents biais qui ne seront pas détaillés ici.

L'ensemble de ces techniques d'étude de la diversité fonctionnelle et taxonomique des écosystèmes a permis de mieux comprendre les mécanismes cellulaires mis en jeu lors de la digestion anaérobie de composés lignocellulosiques et d'explorer plus finement la diversité des micro-organismes qui réalisent cette bioconversion. Le chapitre suivant est dédié aux mécanismes cellulaires impliqués dans la biodégradation des lignocelluloses et la fermentation des produits de dégradation par les bactéries cellulolytiques anaérobies.

III. Aspects moléculaires de la dégradation anaérobie de la cellulose

Le chapitre précédent a permis de souligner le rôle essentiel joué par les bactéries dans la bioconversion des composés lignocellulosiques en conditions anaérobies. Cette bioconversion met en jeu de nombreux mécanismes cellulaires, de la dégradation enzymatique extracellulaire des lignocelluloses au transport, dans la cellule, des produits de dégradation, qui sont ensuite métabolisés dans le cytoplasme. Ces mécanismes, en particulier l'étape d'hydrolyse reposant sur les cellulases et les cellulosomes, sont très étudiés en raison des enjeux économiques liés à leurs multiples applications biotechnologiques et de leur originalité dans le monde du vivant. Les connaissances sur ces mécanismes, en particulier celles liées à *Clostridium cellulolyticum*, organisme modèle utilisé lors de ce travail de thèse, constituent une base fondamentale pour la compréhension de la bioconversion des composés lignocellulosiques en conditions anaérobies. Ces connaissances proviennent principalement de l'étude des bactéries cellulolytiques isolées par culture.

III.1. Bactéries cellulolytiques anaérobies isolées

De nombreuses bactéries cellulolytiques anaérobies strictes ont été isolées et caractérisées. Elles sont pour la plupart mésophiles mais de nombreuses bactéries thermophiles, parfois même hyperthermophiles, ont également été isolées (Blumer-Schuette *et al.*, 2008).

Tableau 1. Caractéristiques de quelques bactéries cellulolytiques anaérobies isolées et séquencées.

Phylum-Classe	Ordre-Famille	Espèce	Origine	T (°C)	C.G.	M	S
Firmicutes – Clostridia	<i>Clostridiales</i> – <i>Ruminoco-</i> <i>ccaceae</i>	1. <i>Clostridium cellulolyticum</i>	Compost	25-45	+	+	+
		2. <i>Clostridium thermocellum</i>	Sols	50-68	+	+	+
		3. <i>Clostridium cellulovorans</i>	Digesteur	20-40	-	-	+
		4. <i>Ruminococcus flavefaciens</i>	Rumen	35-45	+	-	-
	<i>Thermoanaero-</i> <i>bacterales</i> – Family III	5. <i>Caldicellulosiruptor obsidiansis</i>	Source chaude	55-85	+	-	-
Fibrobacteres – Fibrobacteria	<i>Fibrobacterales</i> – <i>Fibro-</i> <i>bacteraceae</i>	6. <i>Fibrobacter succinogenes</i>	Rumen	23-44	-	-	-

Origine : source d'isolation, T : gamme de températures optimale pour la croissance, C.G. : coloration de Gram, M : motilité, S : formation de spores. Références : 1. (Petitdemange *et al.*, 1984), 2. (McBee, 1954), 3. (Sleat *et al.*, 1984), 4. (Ayers, 1958), 5. (Hamilton-Brehm *et al.*, 2010), 6. (Hungate, 1950).

Les écosystèmes à partir desquels ces bactéries ont été isolées sont notamment des tractus digestifs des animaux pour les organismes mésophiles et des sources chaudes pour les organismes thermophiles. Les caractéristiques physiologiques de ces bactéries isolées jusqu'à présent sont très variées : coloration de Gram positive ou négative, cellules motiles ou non, formant ou non des spores. Les principales caractéristiques de bactéries cellulolytiques anaérobies strictes isolées et parmi les mieux caractérisées sont présentées dans le tableau ci-dessus (Tableau 1).

La phylogénie des bactéries cellulolytiques anaérobies strictes isolées, basée sur les séquences des ARNr 16S, montre que ces bactéries ne forment pas un groupe monophylétique. En nombre d'espèces de bactéries isolées, le phylum *Firmicutes*, et plus particulièrement la famille *Ruminococcaceae*, sont les plus représentés. En fait, de nombreuses bactéries cellulolytiques du genre *Clostridium* avaient été classées dans la famille *Clostridiaceae*, telles que *C. cellulolyticum*, mais les analyses phylogénétiques récentes ont permis de reclasser ces organismes dans la famille *Ruminococcaceae* (Ludwig *et al.*, 2007). Au sein du phylum *Firmicutes*, d'autres familles comprenant des bactéries cellulolytiques anaérobies isolées sont représentées, telles que les familles *Clostridiaceae* (*Clostridium. cellulovorans*), *Lachnospiraceae* (*Butyrivibrio fibrosolvans*) et *Thermoanaerobacterales* (*Caldicellulosiruptor obsidiansis*) (Ludwig *et al.*, 2007). Les phyla *Fibrobacteres* (*Fibrobacter succinogenes*) et *Bacteroidetes* (*Bacteroides cellulosilyticus*) sont également représentés, mais seules quelques espèces cellulolytiques ont été isolées à ce jour. Comme c'est le cas chez la plupart des bactéries, ces taxa sont hétérogènes du point de vue des caractéristiques physiologiques. Pendant longtemps, de nombreuses bactéries ont été décrites comme cellulolytiques en raison de leur capacité à dégrader la carboxyméthylcellulose (CMC), dérivé soluble de la cellulose. Cette activité seule ne permet pas la déconstruction de la cellulose native, bien qu'elle y contribue également, et ces organismes ne peuvent donc pas être qualifiés de cellulolytiques. Comme mentionné dans le chapitre précédent, la plupart des bactéries cellulolytiques isolées sont également hémicellulolytiques, mais quelques bactéries uniquement hémicellulolytiques ont également été isolées et appartiennent aux phyla *Bacteroidetes* (*Prevotella paludivivens*) et *Firmicutes* (*Clostridium xylanolyticum*).

Bien que parfois très distantes sur le plan phylogénétique, les bactéries cellulolytiques anaérobies ont acquis et partagent la capacité de dégradation des composés lignocellulosiques, conférée par des systèmes enzymatiques variés et parfois très complexes. Bien que des systèmes de dégradation des lignocelluloses aient été identifiés chez plusieurs bactéries isolées en culture, ils restent encore peu caractérisés chez la plupart des bactéries, en particulier en raison de la diversité des mécanismes. Le développement des approches haut-débit, en particulier la métagénomique, offre néanmoins la possibilité d'explorer et de mieux comprendre ces fonctions au sein d'écosystèmes complexes (Brulc *et al.*, 2009; Maki *et al.*, 2009). Les systèmes les mieux caractérisés seront décrits ci-dessous.

III.2. Systèmes enzymatiques de dégradation anaérobies des lignocelluloses

La déconstruction des lignocelluloses fait intervenir de nombreuses enzymes présentant différentes spécificités de substrat et de modes d'action. Chez les bactéries cellulolytiques anaérobies ces enzymes possèdent des activités carbohydriques estérases (CE), polysaccharides lyases (PL) ou glycosides hydrolases (GH). Le terme cellulases désigne les enzymes actives sur la cellulose et hémicellulases celles actives sur les hémicelluloses. Bien que ces termes apparaissent maintenant comme simplistes au vu des données moléculaires obtenues sur ces enzymes, ils sont encore largement utilisés.

III.2.1. Diversité et mécanismes enzymatiques

Les cellulases et hémicellulases ont une architecture souvent complexe comprenant divers modules fonctionnels et structuraux. Cette complexité se traduit par la répétition d'un même module ou par l'association de plusieurs modules de fonctions différentes au sein d'une même protéine (Figure 4). Parmi les domaines catalytiques, les domaines à activité glycoside hydrolase (GH) sont les plus représentés dans ces protéines. En plus des domaines catalytiques, de nombreux domaines non catalytiques jouent un rôle important dans l'activité de l'enzyme et dans sa spécificité de substrat. Ces domaines peuvent être des domaines de liaisons au substrat, nommés Carbohydrates Binding Module (CBM) (Guillén *et al.*, 2010), ou de liaisons à d'autres protéines (par exemple les dockérines, voir plus loin). Les domaines CBM impliqués dans la reconnaissance et la liaison au substrat, faciliteraient également l'interaction entre l'enzyme et le substrat en orientant le site actif vers le site de clivage et en créant des perturbations locales entre les chaînes de celluloses (Morais *et al.*, 2010; Wang *et al.*, 2008). D'autres domaines, dont les fonctions ne sont pas encore bien établies, peuvent également être présents.

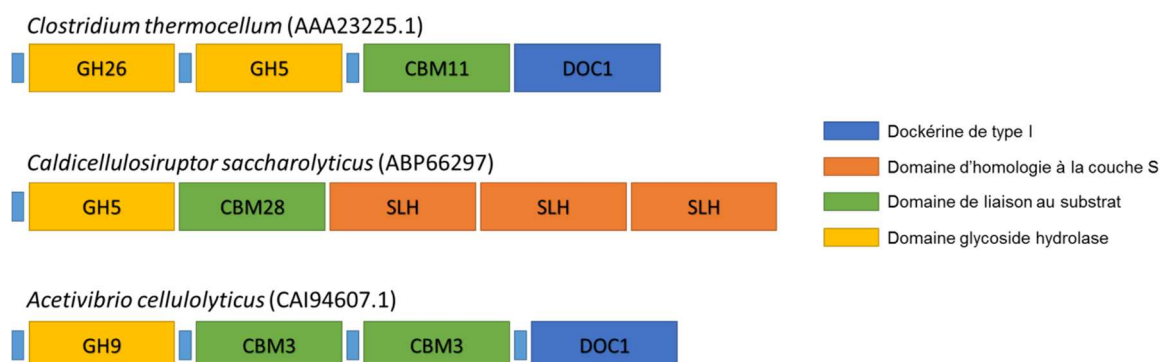
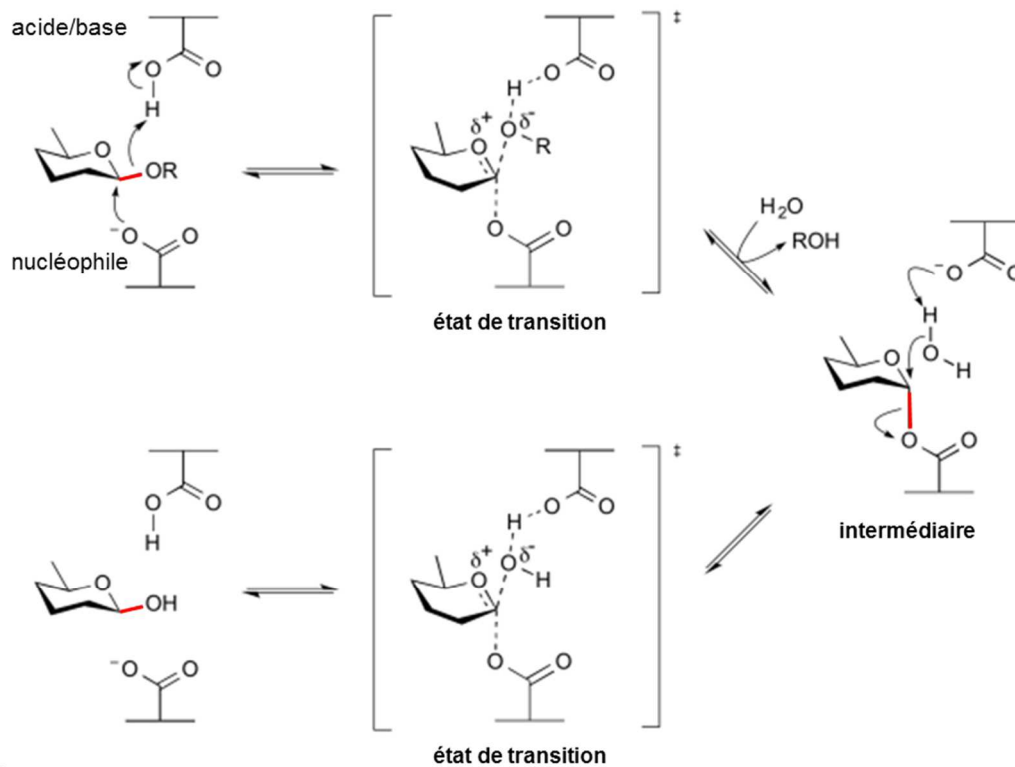


Figure 4. Structure modulaire de quelques cellulases.
Adapté de Sukharnikov *et al.*, (2011) et de Medie and Davies (2012)

Diverses méthodes de classification peuvent être appliquées à ces enzymes. La première consiste à les distinguer, comme toutes enzymes, selon leurs numéros EC (Enzyme Commission number) qui regroupent les enzymes catalysant une même réaction chimique : les cellulases sont par exemple regroupées sous les numéros EC 3.2.1.X. Les cellulases et hémicellulases peuvent également être distinguées selon leur spécificité de sites de clivage. Trois types d'activités sont par exemple associés aux cellulases, qui catalysent toutes l'hydrolyse des liaisons β -1,4 glycosidiques : endoglucanases, exoglucanases et β -glucosidases. Les activités endo/exoglucanases font référence à la capacité de cliver respectivement à l'intérieur (régions amorphes) ou depuis les extrémités (réductrices ou non) des chaînes de cellulose tandis que les β -glucosidases clivent les cellodextrines libérées. Les exoglucanases sont processives, elles libèrent des unités de cellobiose ou de glucose depuis le site de clivage sans libérer le substrat entre les réactions successives. Parmi les endoglucanases, certaines sont processives, à l'instar des exoglucanases (Gilad *et al.*, 2003). La distinction endo/exo s'applique également aux hémicellulases.

Il est également possible de distinguer les glycosides hydrolases selon leur mécanisme d'hydrolyse. Deux principaux mécanismes sont décrits : selon que la stéréochimie de départ du carbone anomérique impliqué dans la liaison glycosidique est conservée ou non il s'agira respectivement d'un mécanisme de rétention ou d'inversion (Figure 5). L'inversion est généralement une réaction acide/base en une étape tandis que la rétention est réalisée en deux étapes par des réactions acide/base et attaque nucléophile (Gloster *et al.*, 2008; McCarter and Withers, 1994). Ces réactions mettent en jeu deux acides aminés, typiquement l'aspartate et le glutamate.

Mécanisme de rétention (β -glycosidase)



Mécanisme d'inversion (β -glycosidase)

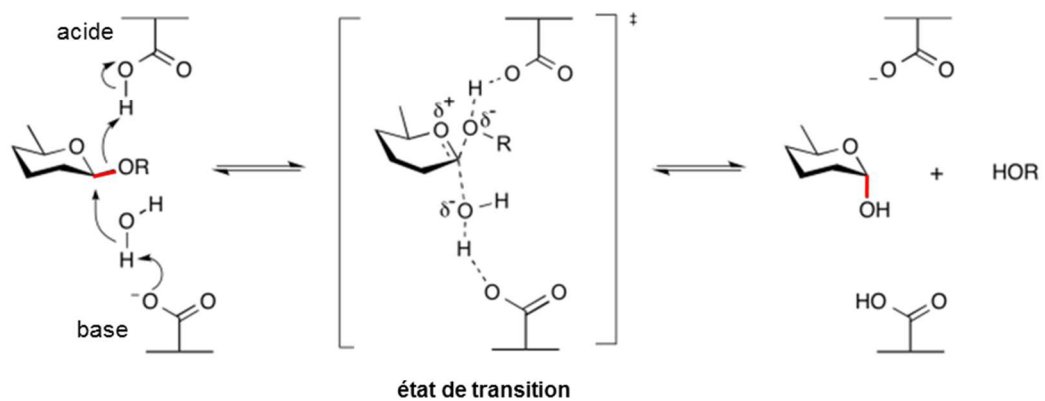


Figure 5. Exemples de mécanismes d'hydrolyse de la liaison β -glycosidique.

Adapté de <http://www.cazypedia.org>.

Enfin, face à la diversité architecturale et à la plurimodularité des enzymes, une autre classification a été adoptée, basée sur les séquences protéiques des enzymes et reflétant plus les caractéristiques structurales (Cantarel *et al.*, 2009). Ces informations sont regroupées dans la base de données CAZy (Carbohydrate Active enZymes) (Cantarel *et al.*, 2009). Le niveau hiérarchique le plus utilisé est celui de la famille, qui regroupe les domaines, catalytiques ou CBM, qui possèdent une forte similarité de séquence. Les familles de domaines catalytiques peuvent être subdivisées en sous-familles, plus homogènes fonctionnellement, ou regroupées en clans, lorsqu'elles possèdent un repliement identique et que la position des acides aminés catalytiques est conservée (Cantarel *et al.*, 2009). À ce

jour 127 familles de domaines glycoside hydrolase et 67 familles de domaines CBM ont été créées. En raison de la modularité, une même protéine peut être répertoriée dans plusieurs familles et les familles sont souvent polyspécifiques, regroupant des enzymes actives sur différents substrats. C'est le cas de la famille 48 (GH48), qui comprend des cellulases et des chitinases (Sukharnikov *et al.*, 2012). La reconstruction phylogénétique des relations entre les séquences en acides aminés de différentes protéines comprenant des domaines GH48 a montré que de nombreuses séquences de protéines fongiques sont proches de celles de protéines de *Firmicutes*, et que la distribution des protéines d'un même organisme est paraphylétique (Sukharnikov *et al.*, 2012). Des résultats similaires ont été obtenus avec des séquences de protéines comprenant des domaines GH9 (Ding *et al.*, 1999), ce qui met en évidence la prépondérance des transferts horizontaux de gènes, qui désignent tout transfert d'information génétique non héritée d'un ancêtre commun, dans l'acquisition de la capacité de dégradation des lignocelluloses. Les glycosides hydrolases sont très représentées dans tous les domaines du vivant et participent à de nombreux mécanismes cellulaires, en particulier le métabolisme des sucres et pour la biosynthèse des glycoprotéines. Dans les génomes des bactéries cellulolytiques, les domaines GH5 et GH9 sont très représentés, et ces séquences sont présentes en plusieurs exemplaires (Blouzard *et al.*, 2010; Dam *et al.*, 2011). Les domaines GH48 sont beaucoup moins représentés et sont généralement présents en un ou deux exemplaires dans ces génomes (Blouzard *et al.*, 2010; Izquierdo *et al.*, 2010). Néanmoins, les enzymes portant des domaines GH48 apparaissent souvent comme des enzymes clés pour la dégradation de la cellulose (Devillard and Goodheart, 2004; Gold and Martin, 2007).

Ces enzymes, cellulases et hémicellulases, sont excrétées dans le milieu par les bactéries et agissent généralement en synergie, ce qui augmente leur activité globale (Beukes *et al.*, 2008; Henrissat *et al.*, 1985). Une autre stratégie de dégradation des lignocelluloses consiste en la formation d'un complexe multienzymatique extracellulaire de très haut poids moléculaire, généralement situé à la surface des cellules et appelé cellulosome. Des bactéries telles que *C. thermocellum* et *C. cellulolyticum* utilisent cette stratégie de dégradation tandis que les bactéries du genre *Caldicellulosiruptor* décrites ne l'utilisent pas (Doi, 2008). De nombreuses études et revues ont été dédiées aux cellulosomes (Bayer *et al.*, 2008, 2004, 1998; Béguin and Lemaire, 1996; Doi and Kosugi, 2004; Doi *et al.*, 2003; Gilbert, 2007; Schwarz, 2001; Shoham *et al.*, 1999) et la section suivante a pour objectifs de mettre en exergue leurs principales caractéristiques.

III.2.2. Les cellulosomes : des complexes multi-enzymatiques particuliers

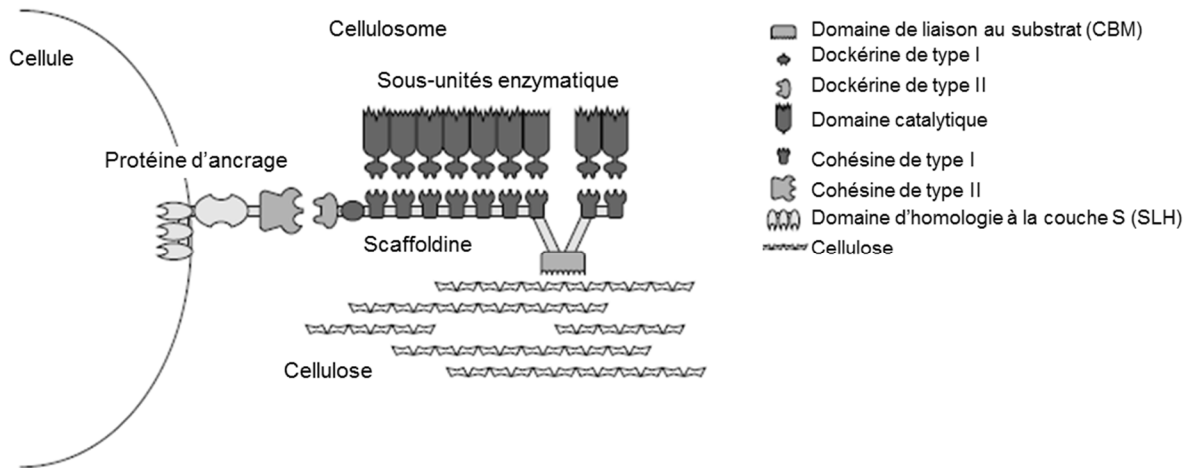


Figure 6. Structure générale des cellulosomes.
Adapté de Shoham *et al.* (1999).

La première caractéristique particulière des cellulosomes est leur structure. D'abord mis en évidence chez *C. thermocellum*, les cellulosomes ont ensuite été identifiés chez de nombreux autres organismes cellulolytiques anaérobies, principalement chez des bactéries appartenant à la famille *Ruminococcaceae*. Dans le système le plus simple, le cellulosome se caractérise par la présence d'une protéine structurale sans activité catalytique, appelée scaffoldine (protéine « échafaudage ») et d'une variété d'enzymes cellulosomales, comprenant principalement des cellulases et des hémicellulases (Figure 6). Dans la plupart des cas, la scaffoldine comprend également un domaine CBM tandis que de nombreuses enzymes cellulosomales en sont dépourvues (Cantarel *et al.*, 2009). La scaffoldine est plurimodulaire et comprend de nombreux domaines de liaisons aux enzymes, appelés cohésines, qui interagissent avec d'autres types de domaines de liaisons, appelés dockérines, présents au sein des enzymes cellulosomales. Les éléments catalytiques sont donc liés au cellulosome grâce à l'interaction entre cohésines et dockérines, dites de type I. De ce fait, les enzymes cellulosomales se distinguent des enzymes libres principalement par la présence de ces modules dockérines. Ce type d'interactions peut également être impliqué dans l'ancrage cellulaire du cellulosome, engageant des domaines cohésines et dockérines dites de type II (Adams *et al.*, 2006). En effet, dans la plupart des cellulosomes un domaine dockérine particulier du cellulosome interagit spécifiquement avec le domaine cohésine de protéines membranaires. Ces protéines sont elles-mêmes liées à la cellule bactérienne *via* des

domaines d'homologies à la couche S – enveloppe cellulaire composée de (glyco)protéines et présente chez de nombreuses bactéries – appelés SLH (Surface Layer Homology) (Desvaux *et al.*, 2006; Sára and Sleytr, 2000). La couche S peut être présente chez les bactéries présentant un Gram positif ou négatif, la différence résidant dans les mécanismes de fixation de cette enveloppe en fonction des caractéristiques de ces deux types de structure membranaire.

Ce schéma général décrit assez bien l'organisation de cellulosomes « simples » tels que ceux présents chez la plupart des *Clostridia* mésophiles incluant *Clostridium cellulovorans* et *C. cellulolyticum* (Bélaich *et al.*, 1997; Tamaru *et al.*, 2010). En revanche, il existe une grande diversité d'organisation des cellulosomes au sein des bactéries cellulolytiques et l'organisation apparaît beaucoup plus complexe chez certaines bactéries telles que *C. thermocellum* et *Acetivibrio cellulolyticus* (Dassa *et al.*, 2012; Xu *et al.*, 2003). Leurs cellulosomes sont en effet composés de scaffoldines d'ancrage qui permettent l'adhésion de multiples scaffoldines primaires porteuses d'unités catalytiques. Ces interactions font à nouveau intervenir une reconnaissance entre domaines cohésines et dockérines, généralement de type II. Par exemple, chez *A. cellulolyticus*, le cellulosome peut comprendre jusqu'à deux scaffoldines d'ancrage liées entre elles (Xu *et al.*, 2003). Chez *Bacteroides cellulosolvens*, une seule scaffoldine d'ancrage a été identifiée mais le grand nombre de domaines cohésines présent dans les scaffoldines d'ancrage et primaire en font le cellulosome le plus grand actuellement décrit, portant à 110 le nombre maximal théorique d'unités catalytiques présentes sur un cellulosome (Ding *et al.*, 2000).

Au sein d'une même espèce, il existe une certaine plasticité et variabilité dans la composition des cellulosomes. Ces aspects ne sont pas encore totalement élucidés et compris, mais il pourrait s'agir d'une combinaison de plusieurs facteurs, notamment l'existence de mécanismes de régulation et des phénomènes d'affinités entre sous-unités induisant des différences liées à la cinétique d'assemblage. En effet, la reconnaissance entre cohésines et dockérines est type spécifique (par définition) mais pas enzyme spécifique, ce qui permet à la bactérie d'exhiber potentiellement une multitude de cellulosomes différents. En particulier, il a été mis en évidence que les activités portées par les cellulosomes varient en fonction du substrat disponible (Blouzard *et al.*, 2010; Cho *et al.*, 2010; Gold and Martin, 2007; Raman *et al.*, 2009). Par exemple, il a été montré qu'en présence de paille de blé, *C. cellulolyticum* produit des cellulosomes riches en hémicellulases qui ne sont pas détectés en présence de cellulose pure (Blouzard *et al.*, 2010). Néanmoins, les cellulosomes produits dans une condition donnée ne sont pas homogènes entre eux et différentes sous-populations de cellulosomes peuvent être identifiées (Han *et al.*, 2005; Pohlschröder *et al.*, 1995). Certaines sous-populations apparaissent cependant majoritaires, ce qui laisse penser que l'assemblage du cellulosome n'est pas aléatoire et qu'il met en jeu des mécanismes de reconnaissance encore peu connus. Il a notamment été montré que l'auto-assemblage extracellulaire du cellulosome est gouverné par l'interaction entre cohésines et

dockérines (Carvalho *et al.*, 2003). La variabilité d'affinité entre cohésines et dockérines serait à l'origine de la formation de sous-populations de cellulosomes (Bomble *et al.*, 2011). Malgré cette possible variabilité, l'interaction entre cohésines et dockérines semble espèce-spécifique (Pagès *et al.*, 1997). Par exemple, les domaines dockérines de *C. cellulolyticum* et *C. thermocellum* ne sont pas interchangeables et cette spécificité serait principalement liée à quelques acides aminés impliqués dans l'interaction cohésines/dockérines (Carvalho *et al.*, 2003; Mechaly *et al.*, 2001). Comme pour les enzymes, l'analyse phylogénétique des séquences des domaines cohésines et dockérines de différentes protéines met en évidence des relations phylogénétiques entre les bactéries qui varient selon le marqueur considéré. L'analyse des séquences de domaines cohésines et dockérines de plusieurs *Clostridia* montre ainsi une nette séparation au sein de la famille *Ruminococcaceae* (Ding *et al.*, 1999; Pagès *et al.*, 1999; Rincon *et al.*, 2003). Les domaines cohésines et dockérines présents chez *C. cellulolyticum* (*Ruminococcaceae*) apparaissent, d'un point de vue phylogénétique, plus proches de ceux de *C. cellulovorans* (*Clostridiaceae*), que de ceux de *A. cellulolyticus* et *C. thermocellum* (*Ruminococcaceae*) qui sont pourtant des espèces plus proches d'un point de vue évolutif.

Les propriétés fonctionnelles des cellulosomes, adaptés à la déconstruction des lignocelluloses sont également remarquables. D'un point de vue enzymatique, le cellulosome permet d'optimiser la synergie entre les différentes fonctions nécessaires à la déconstruction des lignocelluloses, à la fois grâce à la présence simultanée de plusieurs activités enzymatiques de spécificités diverses et grâce à un espacement optimal entre les composants catalytiques. Le cœur de cette machinerie est l'interaction entre cohésines et dockérines qui, à la fois, assure l'assemblage des unités catalytiques et donne un certain degré de liberté à ces enzymes pour l'attaque des fibres de lignocelluloses. De plus, le cellulosome peut jouer un rôle important dans la colonisation cellulaire des composés lignocellulosiques, ce qui offre également d'autres avantages à la cellule. Cet aspect sera détaillé dans la section dédiée à la colonisation bactérienne.

En conclusion, les systèmes enzymatiques lignocellulolytiques apparaissent extrêmement flexibles et permettent aux bactéries de s'adapter au substrat disponible. La synergie d'action, intra- (entre domaines d'une même protéine) et inter-moléculaire, des enzymes libres et des cellulosomes, joue un rôle important dans la dégradation des composés lignocellulosiques. Grâce à l'action de ces enzymes, ces composés deviennent utilisables par les organismes, cellulolytiques, hémicellulolytiques et saccharolytiques, *via* l'assimilation et la métabolisation des produits solubles. Ces mécanismes ont particulièrement été étudiés chez *C. thermocellum* et chez *C. cellulolyticum*, modèle des *Clostridia* cellulolytiques mésophiles et sujet d'étude de ce travail de thèse.

III.3. *C. cellulolyticum* : modèle des *Clostridia* cellulolytiques mésophiles

Modèle des *Clostridia* cellulolytiques mésophiles, *C. cellulolyticum* est une bactérie dont le génome a récemment été séquencé, assemblé et annoté (2009). Ce génome circulaire comporte plus de 4 millions de paires de bases et code 3390 protéines. Quelques caractéristiques ont été décrites dans le chapitre précédent, dont quelques généralités sur l'organisation de son cellulosome. Ce chapitre a pour objectifs de compléter ces informations par des données récentes sur la régulation de l'expression des gènes cellulosomaux et de présenter les principales caractéristiques de son métabolisme carboné. Ces compléments permettront de donner les premiers éléments nécessaires à l'interprétation d'une partie de travail de thèse, centrée sur cette bactérie.

III.3.1. Compléments sur le cellulosome de *C. cellulolyticum*

Le système enzymatique de dégradation des lignocelluloses de *C. cellulolyticum* est composé de 149 CAZymes (Blouzard *et al.*, 2010; Xu *et al.*, 2013). Les domaines GH les plus représentés parmi ces enzymes sont les domaines GH5, GH9 et GH43 qui sont respectivement associés à des activités sur la cellulose ou les hémicelluloses, endoglucanases et xylanases ou β -xylosidases (Cantarel *et al.*, 2009). Au sein du génome de *C. cellulolyticum*, de nombreux gènes codant des protéines cellulosomales sont organisés en clusters, actuellement au nombre de deux. Le premier à avoir été décrit a été nommé *cip-cel*, et comprend 12 gènes : le gène codant la scaffoldine (*cipC*), une protéine portant un domaine cohésine mais dont la fonction est encore inconnue (*OrfX*) et 10 enzymes présentant principalement une activité endoglucanase, certaines présentant également une activité β -mannanase ou rhamnogalacturonane lyase (Desvaux, 2005). Le dernier opéron à avoir été décrit a été nommé *xyl-doc* et comprend 14 gènes codant des enzymes cellulosomales (Blouzard *et al.*, 2010). Ces enzymes possèdent principalement des activités xylanases et galactosidases, mais aucune ne possède d'activités cellulases.

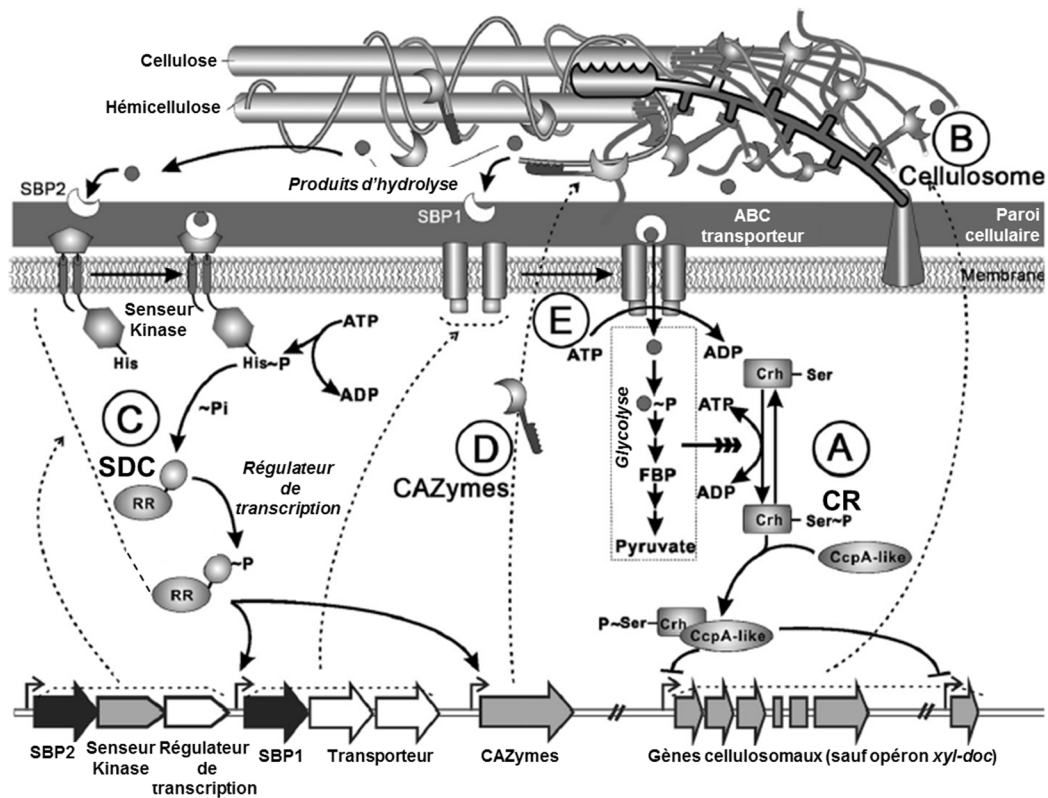


Figure 7. Modèle présentant les mécanismes de régulation des CAZymes chez *C. cellulolyticum*.

SBP : Solute Binding Protein, RC : Répression Catabolique, SDC : Système à Deux Composants. Adapté de Xu *et al.* (2013). A et B : régulation dépendante de la concentration des intermédiaires de la glycolyse ; C, D et E : régulation dépendante de la concentration des produits d'hydrolyse.

Récemment, une étude de transcriptomique de *C. cellulolyticum* en présence de divers substrats, entre autres de la cellulose, du xylose ou des tiges de maïs, a révélé des éléments de la régulation de la transcription de certains gènes cellulosomaux et a permis de proposer un modèle de régulation (Xu *et al.*, 2013). Les principaux résultats ont montré que deux types de mécanismes de régulation sont mis en jeu, la répression catabolique, activée par le niveau d'intermédiaires métaboliques, et le système à deux composants, qui permet la transduction d'un signal extracellulaire à un facteur de transcription. Il est apparu que toutes les CAZymes ne sont pas régulées de la même façon et ces résultats ont été synthétisés dans un modèle (Figure 7).

Ce modèle montre que la répression catabolique intervient dans la régulation de la plupart des gènes cellulosomaux, dont l'opéron *cip-cel* mais pas l'opéron *xyl-doc* et répond aux niveaux de concentration intracellulaire d'intermédiaires de la glycolyse (Figure 7 A et B). Le système à deux composants intervient dans la régulation d'autres CAZymes et des transporteurs de sucres et est activé par les niveaux de concentration des produits d'hydrolyse du substrat (Figure 7 C, D et E). L'activation de l'expression des transporteurs permet l'internalisation des molécules solubles qui seront catabolisées

lors de la glycolyse, ce qui met en évidence une interconnexion entre ces deux systèmes de régulation ainsi que l'interconnexion entre la dégradation du substrat et la fermentation.

III.3.2. Le catabolisme de la cellulose chez *C. cellulolyticum*

Les voies du catabolisme des glucides chez *C. cellulolyticum* ont été déduites par analogie avec celles présentes chez *C. thermocellum* et ont été confirmées par la mesure des activités clés de ces voies (Figure 8). L'internalisation des produits d'hydrolyse du substrat est généralement réalisée par des transporteurs spécifiques de type ABC (ATP Binding Cassette), qui permettent le passage de ces produits grâce à l'énergie fournie par l'hydrolyse de l'ATP (adénosine triphosphate). Les hexoses sont ensuite métabolisés *via* la voie d'Embden-Meyerhof-Parnas (EMP), qui est le type le plus commun de glycolyse, conduisant à la production de pyruvate et de deux molécules d'ATP par hexose catabolisé. Une interaction critique existe entre le flux carboné et le flux d'électrons (Desvaux, 2005), mais ce dernier ne serait pas traité ici. Le pyruvate est un métabolite intermédiaire commun à différentes voies fermentaires et permet d'aboutir à la production d'acétate, d'éthanol et de lactate.

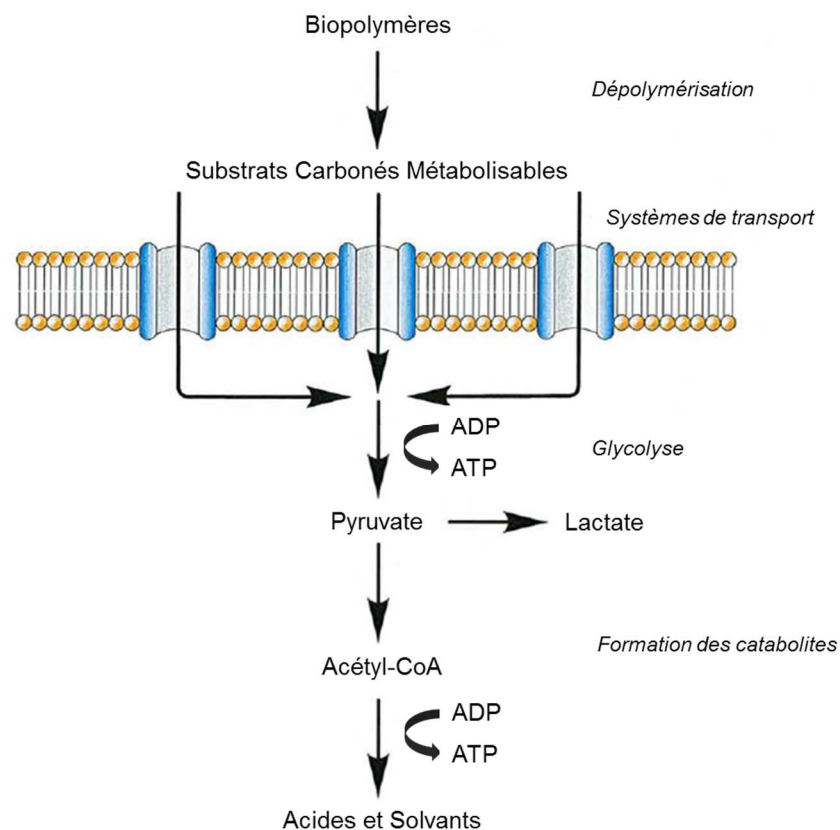


Figure 8. Schéma simplifié du transport et de la fermentation des produits d'hydrolyse de la cellulose chez *C. cellulolyticum*. Adapté de Desvaux (2001).

L'étude de ces flux métaboliques a été réalisée sur cellulose pure et sur cellobiose, en cultures discontinues ou continues (Desvaux *et al.*, 2000; M Desvaux *et al.*, 2001a; M. Desvaux *et al.*, 2001; Giallo *et al.*, 1985, 1983; Guedon *et al.*, 2000, 1999b). Les études métaboliques approfondies nécessitent la mise en place de conditions de cultures adaptées, telle que les chémostats. Les principales conclusions, synthétisées dans une revue (Desvaux, 2005) sont brièvement décrites ci-après.

Il a été noté la sensibilité des régulations métaboliques au flux de carbone entrant, ce qui a été mis en évidence lors de la culture sur cellobiose. Il a en particulier été montré que le profil fermentaire observé pendant la croissance sur cellobiose est différent de celui observé pendant la croissance sur cellulose, les produits finaux majoritaires étant le lactate et éthanol dans un cas, et l'acétate et l'éthanol dans l'autre. Le profil fermentaire observé sur cellobiose est considéré comme une dérégulation du métabolisme puisque le substrat naturel de *C. cellulolyticum* est la cellulose. Le taux d'utilisation du cellobiose est supérieur à celui de la cellulose et ces résultats mettent en évidence le rôle prépondérant de la déconstruction de la cellulose dans la régulation de ce flux de carbone entrant.

A également été mis en évidence la sensibilité de cette bactérie à l'acidification du milieu de culture, *via* des changements importants dans les profils fermentaires, tels qu'une production importante de lactate accompagnée d'une excrétion de pyruvate, et un arrêt brutal de la croissance à un pH de 6,2.

D'autre part, deux nœuds métaboliques ont été identifiés : la phosphoglucomutase, qui catalyse le déplacement d'un groupement phosphate transformant le glucose-1-phosphate et le glucose-6-phosphate, impliquée dans la régulation du flux de carbone entrant vers le métabolisme central et la pyruvate ferredoxine oxydoréductase, impliquée dans la régulation des flux énergétiques (ATP) et électroniques. Plus précisément concernant le premier nœud métabolique cité, lors de l'augmentation du flux de carbone entrant, l'étape catalysée par la phosphoglucomutase devient limitante pour le catabolisme carboné. Une accumulation de glucose-1-phosphate est alors observée et conduit au détournement du flux de carbone vers la production de glycogène et possiblement vers la biosynthèse d'exopolysaccharides. Concernant, le deuxième nœud métabolique cité, il est apparu que l'excrétion de pyruvate observée pour les flux de carbone entrant élevés pourrait être due aux limitations cinétiques de la pyruvate ferredoxine oxydoréductase qui ne serait plus en mesure de prendre en charge le flux de carbone provenant de la glycolyse.

L'ensemble de ces résultats montre que le catabolisme des glucides chez *C. cellulolyticum* est peu adapté à des flux de carbone entrant élevés comme c'est le cas chez d'autres *Clostridia*, tel que *Clostridium acetobutylicum*, une bactérie saccharolytique réalisant une fermentation acétonobutylique (acétone, éthanol et butanol). Des travaux d'ingénierie métabolique sont réalisés

afin de surmonter les limitations métaboliques de *C. cellulolyticum*, en particulier en ciblant les deux nœuds métaboliques identifiés (Guedon *et al.*, 2002). Néanmoins, comme le suggèrent les résultats obtenus sur cellobiose, la dégradation du substrat reste une étape cinétiquement limitante à leur bioconversion. Jusqu'à présent les différents chapitres ont permis de décrire les systèmes enzymatiques de dégradation des composés lignocellulosique ainsi que quelques aspects de la fermentation des produits d'hydrolyse. La colonisation des composés lignocellulosiques par les bactéries cellulolytiques constitue également une phase importante de ce processus et sera décrite dans le chapitre suivant.

IV. Adhésion et dynamiques de colonisation des composés lignocellulosiques

Les microorganismes cellulolytiques adhèrent généralement aux composés lignocellulosiques et peuvent former ou être présents dans des biofilms (Song *et al.*, 2005). Les biofilms sont largement étudiés, principalement en raison de leurs implications majeures dans de nombreux problèmes sanitaires et économiques. Les connaissances acquises sur les biofilms ont également permis d'exploiter ce processus biologique dans de nombreux bioprocédés, tels que la biorémédiation (Singh *et al.*, 2006) ou le traitement des eaux usées (Qureshi *et al.*, 2005). Dans le cadre de la méthanisation de déchets lignocellulosiques, la contribution de l'étape de colonisation à la dégradation n'a jamais été estimée mais pourrait en constituer une étape-clé (Desvaux, 2005; Devillard and Goodheart, 2004). Cette étape sera particulièrement étudiée dans ce travail de thèse. Ce chapitre a pour objectifs de décrire l'étape de colonisation de déchets lignocellulosiques par les microorganismes cellulolytiques et de présenter quelques exemples tirés d'études dédiées à ce processus.

IV.1. Colonisation des substrats lignocellulosiques

Le terme biofilm fait ici référence à une population microbienne englobée ou non dans une matrice de substances exopolymériques (EPS) et adhérente à une surface organique ou inorganique. Selon les définitions, ce terme peut également inclure les agrégats, tels que les floccs et les granules, ou être restreinte à une population englobée dans une matrice d'EPS (Costerton *et al.*, 1995).

Ce mode de vie est très répandu chez les microorganismes (bactéries, archées, protistes) et apparaît comme leur mode de vie préférentiel dans la plupart des écosystèmes. Comparé au mode de vie planctonique, les avantages du mode de vie en biofilm pour ces organismes sont nombreux : persistance au sein des niches écologiques, résistance aux conditions environnementales peu favorables, comme la présence de substances toxiques et de perturbations physico-chimiques, la présence de prédateurs et finalement une vie en communauté (Watnick and Kolter, 2000). La proximité des cellules les unes avec les autres favorise la coopération entre les microorganismes, qu'ils soient de la même espèce ou non. Dans le cadre de la dégradation des composés lignocellulosiques, la colonisation du substrat apporte également d'autres avantages tels qu'une meilleure rétention des enzymes sur le site de dégradation et une consommation plus efficace des produits de l'hydrolyse, parfois inhibiteurs de la cellulolyse.

IV.1.1. Exemples d'études réalisées sur souches pures

La description de la colonisation de la cellulose par les bactéries cellulolytiques anaérobies n'est pas récente (Dinsdale *et al.*, 1978) mais l'étude précise des dynamiques de colonisation n'a fait l'objet de recherches intensives que très récemment, principalement grâce à une utilisation plus large des techniques de microscopie et de traitement d'images, telle que l'imagerie en 3D et en temps réel.

Un premier modèle a été proposé pour décrire la colonisation des composés lignocellulosiques par les *Clostridia*, incluant la famille *Ruminococcaceae*, cellulolytiques mésophiles dont les premières étapes ressemblent à celles du schéma général de la formation des biofilms (Desvaux, 2005; Miron *et al.*, 2001). Il comprend 4 étapes : (i) le transport de la bactérie vers le substrat, (ii) l'adhésion non spécifique, faisant intervenir des forces physico-chimiques, telle que la force de Van der Waals, et les reliefs du substrat (iii), l'adhésion spécifique, impliquant des structures exprimées par les bactéries et iv) la colonisation, qui correspond à la prolifération des cellules adhérentes sur le support. Ce modèle correspond à la formation d'un biofilm mais ne comprend pas la dynamique d'évolution de ce biofilm en fonction de la dégradation de son support.

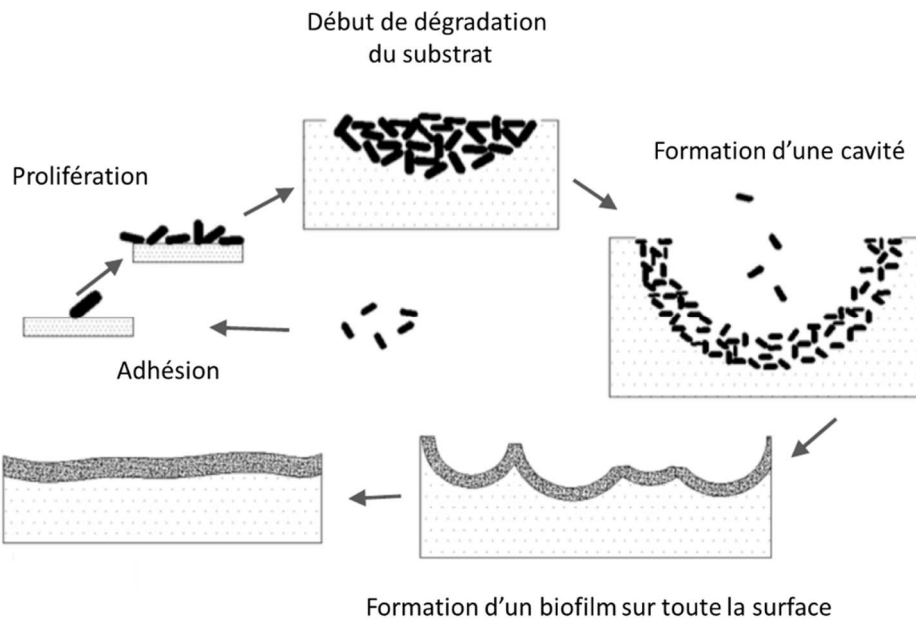


Figure 9. Schéma de la formation d'un biofilm pendant la dégradation de cellulose.
Adapté de Wang *et al.* (2011).

Une étude récente réalisée sur la colonisation de membrane de filtration en cellulose par *Caldicellulosiruptor obsidiansis* a permis de décrire à la fois la dynamique de colonisation de la cellulose par cette bactérie et la dynamique de dégradation du substrat (Wang *et al.*, 2011). Cette étude a été réalisée par microscopie confocale, ce qui permet de réaliser des sections optiques à différentes profondeurs dans l'objet et ainsi de reconstituer une image tridimensionnelle. Les résultats ont montré que le processus commence, comme décrit précédemment, par l'adhésion de bactéries et par leur prolifération à la surface de la cellulose (Figure 9). En se divisant ces cellules forment des colonies qui s'étendent à l'intérieur du substrat. Des cavités sont alors formées, qui peuvent éventuellement se rejoindre à mesure que le substrat est dégradé. Un biofilm très fin (environ 10 µm) peut couvrir la surface de la cellulose. Dans cette même étude, des cavités ont également été observées avec *C. thermocellum*. Dans une autre étude, un biofilm multicouche a également été décrit pour cette dernière bactérie (Hamilton-Brehm *et al.*, 2010). Cependant la plupart des observations de biofilms chez les bactéries cellulolytiques anaérobies indiquent la formation de biofilms monocouches (Tableau 2).

Tableau 2. Caractéristiques de quelques biofilms de bactéries cellulolytiques anaérobies isolées.

Souche	Substrat	Type de culture	Biofilm	Technique
1. <i>Fibrobacter succinogenes</i> – <i>Ruminococcus albus</i> – <i>Ruminococcus flavefaciens</i>	Luzerne	Discontinue	Mono	MEB
2. <i>Clostridium celluloliticum</i>	Cellulose	Continue	Mono	MEB
3. <i>Clostridium thermocellum</i>	Cellulose	Continue	Mono	Confocal
4. <i>Caldicellulosiruptor obsidiansis</i>	Cellulose	Discontinue	Multi	Confocal

Mono : biofilm monocouche, Multi : biofilm multicouche, MEB : Microscopie électronique à balayage. Adapté de Wang *et al.* (2011). Références : 1. (Miron *et al.*, 1989), 2. (Desvaux, 2005), 3. (Dumitrache *et al.*, 2013), 4.(Wang *et al.*, 2011).

Une autre étude récente centrée sur les dynamiques de colonisation et de dégradation de fibres de coton a été réalisée avec *C. thermocellum* (Dumitrache *et al.*, 2013). Contrairement à l'étude précédemment citée et à la plupart des études où des informations sur la colonisation ont été mentionnées, cette étude a été réalisée en cellule à flux (« flow-cell ») et non en réacteur classique. Une « flow-cell » est un dispositif de microfluidique qui permet de réaliser simultanément une culture en continue et des observations microscopiques directement avec ce support (Hansen *et al.*, 2000). Les résultats ont montré que la colonisation se déroule selon le modèle décrit précédemment (Figure

9). Cette étude a de plus permis de mettre en évidence que le recrutement de cellules est également un phénomène important lors de la colonisation de la surface, qui ne dépend donc pas que de la division de cellules déjà adhérentes. D'autre part, des bactéries à différents stades du cycle cellulaire, cellules en division ou en sporulation, peuvent être observées. Le détachement de cellules est un phénomène continu durant toute la croissance du biofilm bien qu'un nombre plus important de cellules se détache à mesure que le substrat est consommé. Bien que l'étude ait été réalisée en « flow-cell », les résultats obtenus ne sont pas issus d'acquisitions en continu mais de ré-échantillonnages de la culture, en raison de l'utilisation de Syto9, un colorant des acides nucléiques, non adapté à une utilisation en continu compte-tenu de sa toxicité. Très récemment, une étude a décrit l'utilisation de « flow-cell » couplée avec de l'imagerie en temps réel grâce à l'utilisation d'un colorant lipophile non cytotoxique pour l'observation de biofilms de *C. thermocellum* en présence de fibres de coton ou de membrane de cellulose (Wang *et al.*, 2013). L'utilisation de telles techniques, adaptées à l'observation de ces biofilms, permettra d'obtenir plus d'informations sur la dynamique de colonisation et de dégradation des composés lignocellulosiques par les bactéries cellulolytiques anaérobies.

IV.1.2. Exemples d'étude en biomasses complexes

Dans le rumen, la colonisation des particules lors de leur digestion apparaît comme un phénomène important (McAllister *et al.*, 1994). Une grande partie de la population microbienne adhère à ces particules et ces consortia sont majoritairement impliqués dans leur dégradation (Koike *et al.*, 2003). Dès les premières études, il a été observé que les bactéries adhérentes étaient présentes dans des cavités du substrat, suggérant que ces cavités étaient créées suite à sa dégradation (Dinsdale *et al.*, 1978).

Dans le contexte de la méthanisation, peu d'études incluent des observations de la colonisation des composés lignocellulosiques. Une étude a toutefois été particulièrement dédiée à cet aspect (Song *et al.*, 2005). Le principal résultat est la mise en évidence de la formation d'un biofilm mixte à la surface des particules de cellulose, comprenant à la fois des bactéries et des archées méthanogènes englobées dans une matrice. Dans ces conditions, le biofilm joue un rôle central puisqu'il réunit les microorganismes capables de réaliser l'ensemble des étapes de la bioconversion totale de la cellulose en méthane.

IV.2. Adhésion bactérienne à la cellulose

Par rapport à l'ensemble des étapes, l'adhésion des premières cellules a lieu rapidement, au cours des dix premières minutes environ, la constitution du biofilm pouvant être observée de quelques heures à plusieurs jours plus tard. De nombreux facteurs physico-chimiques et biologiques peuvent jouer un rôle dans l'adhésion des bactéries au substrat. Ces facteurs peuvent être liés aux propriétés physico-chimiques du substrat ou de l'environnement, telles que la surface disponible, les charges, le pH et la température, ainsi qu'au compartiment biologique, en lien avec la concentration en cellules ou les phénomènes de compétition microbienne (Miron *et al.*, 2001). L'influence de quelques-uns de ces facteurs a été étudiée sur plusieurs souches de bactéries cellulolytiques anaérobies et les résultats sont synthétisés dans le tableau ci-dessous (Tableau 3). Ces résultats montrent que les souches de bactéries cellulolytiques anaérobies les plus étudiées possèdent des sensibilités différentes aux conditions physico-chimiques et biologiques en ce qui concerne leur capacité d'adhésion à la cellulose.

Tableau 3. Effets de quelques facteurs physico-chimiques sur l'adhésion de souches pures cellulolytiques à la cellulose.

	Charges	pH	T (°C)	Ions	Sucres solubles
1. <i>C. thermocellum</i>	+	=	=		
	(molécules chargées +)	(4-9,5)	(4-60)		
2. <i>R. flavefaciens</i>		=	=	-	=
		(3,5-7,5)	(4-52)	(en absence de Ca ²⁺ /Mg ²⁺)	(≤1%)
2. <i>F. succinogenes</i>		-	-	=	=
		(4,5-6)	(4-30 / 43-52)		(≤1%)
		-			-
		(>7,5)			(5%)

T : température, = : pas d'effets, + : effet positif, - : effet négatif. Références : 1. (Bayer *et al.*, 1983), 2. (Roger *et al.*, 1990).

Contrairement à la quantité d'information disponible sur les systèmes enzymatiques des bactéries cellulolytiques anaérobies, les mécanismes impliqués dans leur adhésion au substrat sont beaucoup moins connus. Les protéines potentiellement impliquées sont celles comportant des domaines CBM, telles que les scaffoldines ou d'autres protéines ancrées à la membrane, des pili ou enfin des glycoprotéines (glycocalyx) (Miron *et al.*, 2001; Mosoni and Gaillard-Martinie, 2001; Weimer *et al.*, 2006). Le rôle des cellulosomes dans l'adhésion a d'abord été suggéré chez *C. thermocellum*, principalement en raison de la présence de protubérances membranaires regroupant plusieurs cellulosomes à la surface de la cellule (Lamed *et al.*, 1987). Chez *C. cellulolyticum*, la présence de protubérances n'a jamais été observée (Ferdinand *et al.*, 2013). En revanche, il a récemment été

montré que la scaffoldin *CipC* joue un rôle dans l'adhésion mais que d'autres structures sont probablement impliquées, puisque l'inactivation de son expression n'empêche pas l'adhésion (Ferdinand *et al.*, 2013). Le rôle du cellulosome ne semble cependant pas déterminant dans le profil de colonisation puisque les cavités formées dans la cellulose, mentionnées ci-dessus, sont observées chez *C. thermocellum*, qui possède un cellulosome, et *C. obsidiansis*, qui n'en possède pas (Wang *et al.*, 2011). L'implication des pili dans l'adhésion a été décrite chez *R. albus* probablement également en combinaison avec les cellulosomes (Morrison and Miron, 2000; Rakotoarivonina *et al.*, 2002). Chez *C. cellulolyticum*, la présence de gènes codant des protéines impliquées dans la formation des pili a récemment été décrite mais leur rôle dans l'adhésion n'a encore jamais été démontré (Imam *et al.*, 2011). Enfin, chez *C. thermocellum* la présence d'une structure proche d'un glycocalyx a été suggérée suite à l'observation de l'adhésion de lectines à la surface de la cellule (Dumitrache *et al.*, 2013).

L'ensemble de ces observations montre que les informations disponibles sur la dynamique de colonisation et de dégradation du substrat ainsi que sur les mécanismes impliqués sont peu nombreuses. D'autre part, bien que certains mécanismes aient été décrits chez la plupart des bactéries cellulolytiques anaérobies, d'autres mécanismes encore inconnus semblent également responsables de l'adhésion des bactéries à leur substrat.

Les données bibliographiques présentées dans ce chapitre et dans les deux chapitres précédents mettent en évidence l'adaptation des bactéries cellulolytiques anaérobies à la dégradation d'un substrat récalcitrant, la lignocellulose. Cette adaptation se manifeste par leur présence dans des niches écologiques spécifiques, par l'interaction avec d'autres microorganismes, par la production de systèmes enzymatiques complexes et enfin par la formation de biofilms adaptés à la dégradation d'un substrat insoluble. Malgré cette adaptation, la biodégradation des composés lignocellulosiques en conditions anaérobies est une étape cinétiquement limitante dans la bioconversion de cette ressource. Aussi, l'utilisation de prétraitements des composés lignocellulosiques en amont des bioprocédés peut permettre d'en améliorer les performances pour une utilisation à grande échelle de ces bioprocédés.

V. Prétraitements des composés lignocellulosiques

Les prétraitements des composés lignocellulosiques regroupent l'ensemble des traitements appliqués à ces composés en amont des bioprocédés. Ils visent à diminuer leur récalcitrance à la dégradation biologique et ainsi améliorer les rendements de valorisation pour une exploitation industrielle. Ces traitements agissent sur la structure et/ou sur la composition chimique des lignocelluloses. Le développement de stratégies de prétraitement des composés lignocellulosiques fait l'objet de recherches intensives depuis des décennies et notamment dans le contexte de la production de biocarburants (principalement de bioéthanol) à partir de résidus lignocellulosiques telles que les pailles (biocarburants de 2^{ème} génération). Une littérature conséquente existe sur ces différentes stratégies et de nombreuses revues ont été consacrées à leur comparaison (Hendriks and Zeeman, 2009; Kumar *et al.*, 2009; Mosier *et al.*, 2005; Taherzadeh and Karimi, 2008; Wyman *et al.*, 2005). Ce chapitre n'a pas pour objectif de dresser une liste exhaustive de l'ensemble des stratégies mises en œuvre, qui est régulièrement mise à jour dans des revues, mais de donner un aperçu des stratégies les plus étudiées dans le contexte de la méthanisation de déchets lignocellulosiques.

V.1. Aperçu des stratégies de prétraitements

Les principaux facteurs impliqués dans la récalcitrance des composés lignocellulosiques à leur biodégradation et généralement ciblés par les prétraitements sont la cristallinité de la cellulose, la présence de lignine et la surface accessible ; les prétraitements développés ciblent au moins l'un de ces facteurs (Mosier *et al.*, 2005). La contribution de chacun de ces facteurs à la récalcitrance globale du composé est difficilement évaluable en raison de leur interconnexion. De nombreux autres facteurs, tels que le degré de polymérisation et la porosité, joueraient également un rôle dans la récalcitrance (Mansfield *et al.*, 1999). Les prétraitements de base sont généralement regroupés selon leur nature, mais de nombreux prétraitements combinent plusieurs approches. Ils incluent les traitements biologiques, chimiques et physiques. L'efficacité de certains de ces prétraitements dans un contexte de méthanisation a déjà été démontrée (Baba *et al.*, 2013; Frigon *et al.*, 2012; Jackowiak *et al.*, 2011; Liu *et al.*, 2002; Pommier *et al.*, 2010; Romano *et al.*, 2009; Sharma *et al.*, 1988; Teghammar *et al.*, 2010; Vivekanand *et al.*, 2013; Xiao and Clarkson, 1997; Yuan *et al.*, 2012). Quelques exemples sont décrits ci-dessous.

L'explosion à la vapeur consiste en une injection de vapeur d'eau à hautes température et pression, dans la biomasse lignocellulosique, suivie d'une chute brutale de la température et de la pression. Cette stratégie a pour effet principal d'augmenter la surface accessible en solubilisant les hémicelluloses (Hendriks and Zeeman, 2009). L'effet sur la solubilisation des lignines est plus mineur mais permet néanmoins leur déstructuration. Une étude récente réalisée sur du bouleau a montré qu'un traitement à 220°C pendant 10 minutes réduisait d'un facteur 2,7 le taux de xylanes et permettait d'améliorer le rendement en méthane d'un facteur 1,8 (Vivekanand *et al.*, 2013). L'augmentation de la température ou du temps de traitement avait au contraire un effet délétère sur la production de méthane. Ces mêmes paramètres opérationnels ont également été appliqués au prétraitement de tubes de cartons (Teghammar *et al.*, 2010). Cette étude-ci a également montré que l'effet de ce traitement était amélioré par l'ajout de 2% de soude et 2% de peroxyde d'hydrogène, améliorant d'un facteur 2 le rendement en méthane par rapport au matériau non prétraité. L'explosion à la vapeur a également été appliquée sur une fraction lignocellulosique de déchets de compositions variées et les résultats ont montré que le rendement en méthane était amélioré y compris lorsque le traitement était appliqué en cours de digestion (Liu *et al.*, 2002).

Des consortia microbiens complexes issus de divers écosystèmes cellulolytiques ont récemment été utilisées pour prétraiter des déchets de papier. Parmi ces consortia figure une communauté microbienne enrichie sur cellulose à partir d'un inoculum issu d'un compost (Haruta *et al.*, 2002). Le prétraitement en conditions aérobies et thermophiles de différents types de déchets carton et papier, tels que du papier de bureau, du papier journal, par ce consortium, réalisé avant une digestion anaérobie classique, a permis de fortement augmenter le rendement en méthane (augmentation > 100%, selon les déchets) lorsque le prétraitement était appliqué pendant 7 jours (Yuan *et al.*, 2012). Pendant le prétraitement, plusieurs solvants et acides gras volatils ont été produits à partir de la cellulose et des hémicelluloses contenues dans les déchets. Ces produits, dont la concentration maximale a été atteinte après 7 jours de prétraitement, ont été des substrats directement disponibles pour la digestion anaérobie. Dans une autre étude, du liquide ruminal (pH 7,2) a été utilisé pour prétraiter du papier journal (Baba *et al.*, 2013). Ce prétraitement a permis d'augmenter le rendement en méthane de 2,6 fois par rapport au déchet non prétraité lorsqu'il était appliqué pendant 6h à 37°C, l'augmentation du temps de traitement n'améliorant pas ce résultat. Plusieurs acides gras volatils ont été produits pendant ce prétraitement, reflétant la solubilisation d'une partie importante de la lignine en plus de la dégradation de la cellulose et des hémicelluloses.

Les broyages ou découpages consistent principalement en une réduction par voie mécanique de la taille des déchets. Ce type de prétraitements a été appliqué à des déchets végétaux et des déchets papiers. Une première étude réalisée sur différents types de déchets végétaux a montré que le broyage de pailles (blé, riz et chiendent) permettait d'améliorer le rendement en méthane d'un facteur 1,5 environ, tandis que le broyage de feuilles de diverses espèces de plantes ne conduisait pas à une amélioration importante de cette production (Sharma *et al.*, 1988). Il a également été montré que les augmentations les plus fortes du rendement en méthane étaient obtenues pour des tailles de particules allant de 6 mm à 400 µm, le gain obtenu en-deçà de cette taille n'améliorant que très peu la production. Une autre étude a été réalisée sur des mélanges de papiers et de cartons, l'un reconstitué et l'autre étant un déchet réel entrant en ISDND (Pommier *et al.*, 2010). Les résultats ont montré que pour ces deux mélanges, le découpage fin (< 1 mm) ne conduisait ni à l'amélioration du rendement en méthane, ni à l'amélioration de sa cinétique de production. Dans une étude réalisée à plus large échelle, deux types de broyages ont été mis en œuvre sur des résidus de fourrage (Lindmark *et al.*, 2012). Ces prétraitements, qui ont abouti à des particules de taille fine, dont la moitié faisait moins de 125 µm, ont permis d'améliorer les rendements en méthane d'un facteur 1,5.

L'ensemble des études réalisées sur des prétraitements de fractions lignocellulosiques montrent que leur efficacité dépend de nombreux facteurs, et en particulier du composé à traiter. D'autre part, le type de procédé de digestion anaérobie appliqué aux composés lignocellulosiques prétraités constitue également un facteur important. En effet, de nombreuses interactions entre les propriétés du déchet prétraité et les paramètres opératoires du bioprocédés peuvent exister, comme cela a été montré pour la taille des particules obtenues après broyage et le taux de matière sèche employé (Motte *et al.*, 2013).

V.2. Les prétraitements dans un contexte de méthanisation

Lors de l'élaboration d'un prétraitement adapté aux composés lignocellulosiques dans le contexte de leur méthanisation, certains critères principaux sont recherchés (Hendriks and Zeeman, 2009; Mosier *et al.*, 2005). Le premier critère est l'amélioration de leur biodégradabilité dans le bioprocédé employé tout en préservant les polysaccharides (cellulose et hémicelluloses) dont la bioconversion est au cœur de leur valorisation. Dans certains cas, le prétraitement appliqué est trop intense et un effet délétère est observé sur les rendements en méthane. La limitation de la formation de fortes concentrations d'inhibiteurs biologiques, tel que le furfural (Palmqvist and Hahn-Hägerdal, 2000), est aussi recherchée afin d'assurer le bon fonctionnement du bioprocédé en aval. Le temps de traitement, qui peut varier de quelques minutes à plusieurs semaines selon les cas, est également l'un de ces critères, puisqu'il peut entraîner des coûts et des problématiques de stockage. D'un point de vue environnemental et

sanitaire, la demande en énergie, en ressources, particulièrement en eau, et l'utilisation de produits dangereux sont également pris en compte. Enfin, les coûts liés au prétraitement, tels que les coûts des réactifs, liés au recyclage et énergétiques, figurent également parmi les critères principaux.

En résumé, le choix des prétraitements repose sur un compromis entre l'impact sur les bioprocédés en aval d'une part et les coûts fixes et opératoires générés d'autre part. Dans cette optique, les prétraitements mécaniques par broyages ont souvent été critiqués en raison de leur forte consommation d'énergie (Hendriks and Zeeman, 2009). Généralement les conditions nécessaires pour atteindre une taille de particule inférieure à 1 mm ne sont pas économiquement rentables (Barakat *et al.*, 2013). Néanmoins une étude récente sur l'utilisation d'un broyage fin sur des résidus de fourrages (présentée ci-dessus) a montré que l'un des types de broyages employé avait un bilan énergétique positif et était compatible avec une installation sur un site industriel de méthanisation (Lindmark *et al.*, 2012). Il serait utile de réaliser des analyses rigoureuses de l'impact des prétraitements, de type ACV, afin d'en estimer leur potentiel en prenant en compte de nombreux aspects, mais elles sont rarement effectuées.

VI. Organisation du travail de thèse

La méthanisation des composés lignocellulosiques est en plein développement et suscite un fort intérêt en raison de son fort potentiel de valorisation. Cependant, leur traitement par cette filière est limité en raison de la récalcitrance de la lignocellulose à la dégradation, liée en partie à son hétérogénéité de composition et de structure. Une littérature abondante existe sur la dégradation des lignocelluloses en conditions anaérobies, en particulier dans les écosystèmes animaux, et sur les prétraitements de ces composés. Néanmoins, peu d'informations sont disponibles sur l'établissement de la colonisation des composés lignocellulosiques par les bactéries cellulolytiques anaérobies et son déterminisme, et sur le lien entre colonisation et dégradation. La compréhension des mécanismes biologiques et des facteurs pouvant les influencer, notamment en fonction des prétraitements appliqués, permettrait donc d'améliorer les performances de dégradation des matériaux lignocellulosiques au cours de leur méthanisation.

Dans ce contexte le sujet de thèse traite de la colonisation des déchets lignocellulosiques en conditions anaérobies en lien avec l'efficacité de leur dégradation. En particulier le déterminisme de cette étape sera abordé, notamment concernant les caractéristiques physico-chimiques des substrats avec ou sans prétraitements. L'ensemble du processus sera pris en compte afin de mieux comprendre leurs effets sur l'ensemble de la bioconversion de ces composés. Cette problématique recouvre en partie différents aspects :

- Structures physico-chimiques des substrats
- Performances de méthanisation
- Colonisation du substrat
- Structures et fonctions des communautés microbiennes

Afin d'aborder ces aspects une stratégie transversale a été développée combinant des approches sur systèmes simples et des approches sur systèmes complexes (Figure 10). La distinction faite dans ce manuscrit entre substrats « simples » et « complexes » est majoritairement liée à l'absence ou présence de lignines dans ces matériaux.

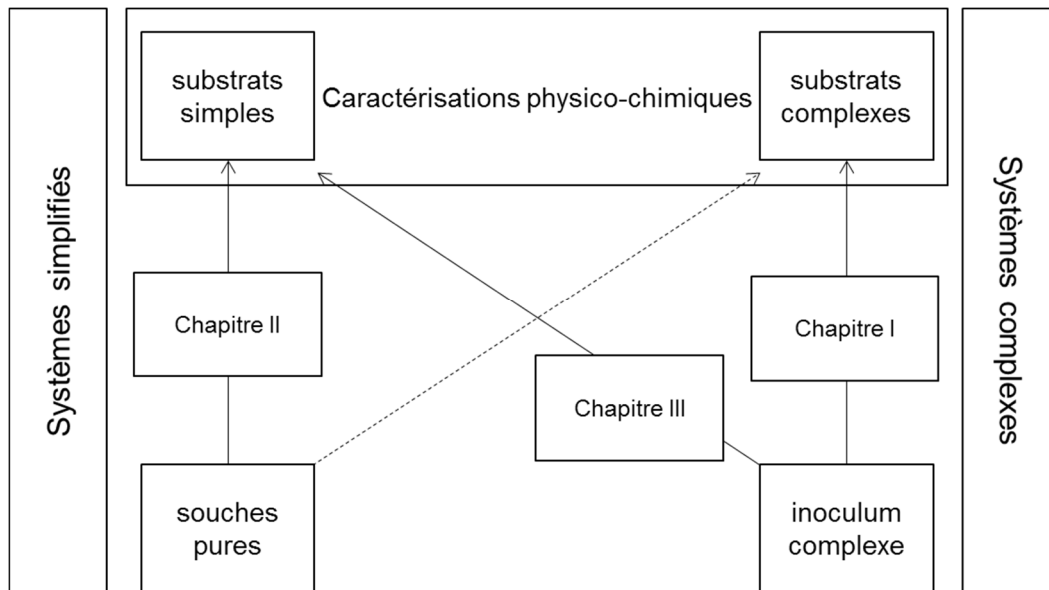


Figure 10. Stratégie expérimentale mise en place pendant la thèse.

La flèche en pointillés fait référence à une étude réalisée dans le cadre d'un stage de Master dont les résultats n'étaient pas d'une qualité suffisante pour l'exploitation.

Systèmes simples

Cette approche utilisant des modèles biologiques permet de réaliser une première étude du déterminisme de la colonisation du substrat en lien avec les performances de dégradation. Elle permet également une étude approfondie des mécanismes cellulaires influencés par les différentes conditions grâce aux informations déjà disponibles sur ces modèles. Les travaux réalisés concernent la colonisation et la dégradation de celluloses pures et de déchets cellulosiques « simples » par des souches cellulolytiques modèles. Les deux souches choisies sont *Clostridium cellulolyticum* et *Acetivibrio cellulolyticus*, qui sont proches des souches présentes dans les méthaniseurs. Les principaux résultats obtenus sont présentés dans le chapitre II, dans l'article intitulé « Integrated quantitative whole-proteome analyses of the biofuel-producing *Clostridium cellulolyticum* ».

D'autre part, afin de mettre en évidence les interactions entre micro-organismes lors de la dégradation de la cellulose, des études ont été réalisées avec plusieurs souches pures, entre les deux souches cellulolytiques étudiées et entre les souches cellulolytiques et une archée méthanogène modèle, *Methanosarcina barkeri*. Ces résultats n'ont pas permis de préparer une publication et sont donc présentés dans un chapitre complémentaire.

Systèmes complexes

Cette deuxième approche permet l'étude du déterminisme de la colonisation du substrat en lien avec les performances de dégradation dans des conditions de méthanisation, plus réalistes. Les travaux

réalisés concernent la dégradation de paille de blé et de carton plat, déchets lignocellulosiques modèles. Les cartons en général peuvent en particulier être très présents dans les ordures ménagères (environ 8% en masse de déchets (ADEME, 2009)). La paille de blé et le carton plat sont ici prétraités par broyages, en présence d'une biomasse complexe issue du digesteur d'ordures ménagères résiduelles et de biodéchets de Varennes-Jarcy (91), fonctionnant en conditions mésophiles et anaérobies. Les résultats obtenus sont présentés dans le chapitre I, dans l'article intitulé « Dry grinding of lignocellulosic substrates shapes microbial communities during methanization ».

Afin de renforcer les liens entre les deux approches, des travaux ont été réalisés avec les déchets celluloseux sélectionnés lors de la première approche dans des conditions de méthanisation en présence de la biomasse complexe. La concordance de résultats entre souches pures et biomasse complexe est peu étudiée dans le cadre d'études de bioprocédés. Cette partie permet de valider en conditions plus réalistes les performances de dégradation et d'émettre des hypothèses plus précises sur les comportements microbiens et sur les structures des communautés microbiennes. Les principaux résultats obtenus seront présentés dans le chapitre III, dans l'article intitulé « Overview of methanization from distinct cellulosic materials based on comparative metaproteogenomics ».

Caractérisations des substrats

En parallèle, un autre axe a été dédié à l'étude des caractéristiques des substrats afin d'établir un lien avec les performances de dégradation et les phénomènes biologiques observés. Les résultats obtenus, issus de nombreuses collaborations, sont intégrés aux chapitres correspondants.

Une partie importante du travail de thèse a consisté à adapter des méthodes de biologies classiques, telles que la culture de souches pures anaérobies et l'observation de la colonisation, et d'écologie moléculaire, telles que la (méta)protéomique et la métagénomique, à l'étude de la dégradation des composés lignocellulosiques dans les conditions utilisées pendant la thèse.

RESULTATS EXPERIMENTAUX

CHAPITRE 1

Avant-propos

L'objectif de ce chapitre est d'étudier l'effet de prétraitements, susceptibles de modifier les structures fines des composés lignocellulosiques et donc de modifier les rendements de dégradation de ces composés. Différents types de broyages par voie sèche ont été appliqués à de la paille de blé et du carton plat. Ces broyages ont été réalisés dans le cadre d'une collaboration avec l'UMR IATE de Montpellier. Dans le cadre de leurs travaux ces broyages ont été très étudiés sur la paille de blé et permettent donc de bénéficier d'un grand nombre d'informations sur leurs effets (Silva and Rouau, 2011; Silva *et al.*, 2012). D'autre part, les broyages par voie sèche présentent l'avantage de ne pas consommer d'eau et de ne pas produire d'effluents, toxiques ou non. Les broyages en général permettent également une bonne homogénéisation du contenu du digesteur. Les différentes techniques qui ont été utilisées sont :

- Le broyage à couteaux : le traitement consiste en des effets de coupe de la matière réalisés à l'aide de couteaux empilés sur un ou plusieurs arbres.
- Le broyage à boulets : le broyage est obtenu par les chocs et les frottements provoqués par la chute de boulets sur la matière à broyer, à l'intérieur d'un tambour rotatif.
- Le broyage à impact : la matière est broyée par les impacts mécaniques à grande vitesse et le cisaillement se produisant entre le matériau et les broches du rotor.
- Le broyage à jet d'air : le broyage est obtenu par les collisions entre les particules fortement accélérées par des jets d'air comprimé.

Différentes intensités ont été appliquées et ont permis d'atteindre des tailles de particules de l'ordre de 170 μm ou 60 μm , sauf pour le broyage à couteaux qui a généré des particules de l'ordre de 4 mm (Figure 11). Les broyages intenses ont en particulier été appliqués pour permettre de détruire les structures cellulaires dans le cas de la paille.

Outre la compréhension des liens entre structures des substrats et performances, les conditions de broyages appliquées permettent aussi d'évaluer les limitations de ces techniques.



Figure 11. Aperçu de quelques fractions de paille de blé obtenues après broyage par voie sèche (photographies).

A : Paille de blé broyage aux couteaux, B : Paille de blé broyage à boulets (65 µm).

L'ensemble de ces fractions ont été testées en méthanisation dans des incubations en culture discontinues (« batch ») non agitées de type BMP (Biochemical Methane Potential), qui est une évaluation normée de la biodégradabilité des composés organiques en conditions anaérobies (Figure 12). Le milieu réactionnel consiste en une solution minérale correspondant à la norme BMP NF EN ISO 11734 supplémentée d'une solution de carbonates (0,25M), du substrat et, d'un inoculum filtré, stabilisé, et concentré, puis conservé à -80°C. Il est issu du digesteur mésophile de Varennes-Jarcy (91). Le ratio inoculum/substrat (basé sur les teneurs en matières volatiles) utilisé dans cette étude a été choisi par rapport à d'autres études réalisées au laboratoire (Boulangier, 2011; Lü *et al.*, 2013).

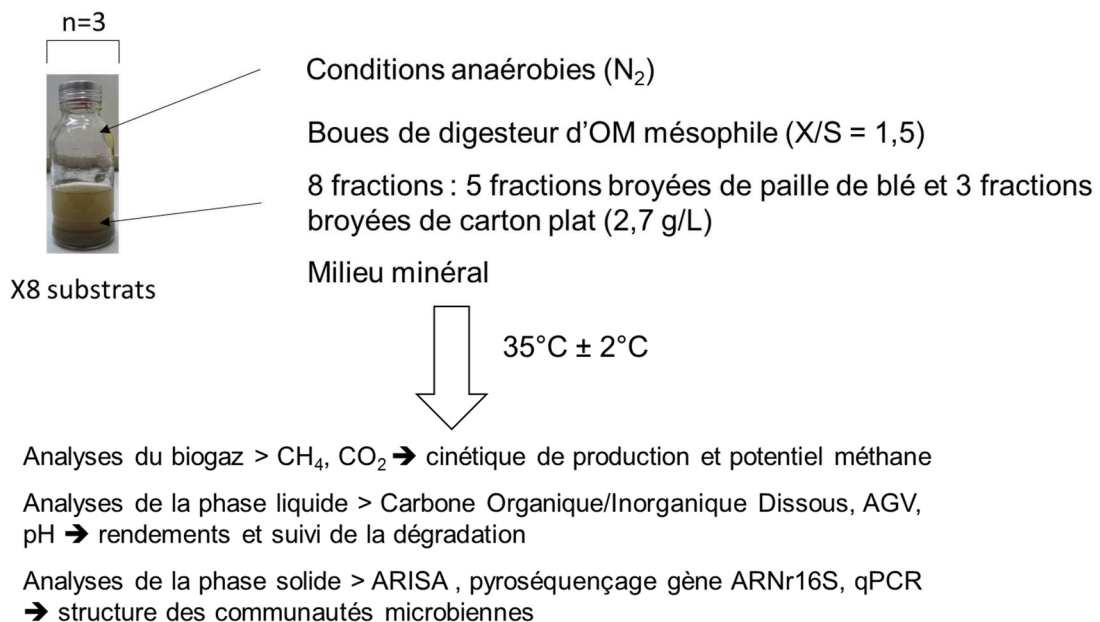


Figure 12. Dispositif expérimental réalisé pour l'étude de la dégradation de paille de blé et de carton plat broyés par une biomasse complexe.

Les analyses réalisées sur ces incubations consistent en un suivi de la production et de la composition du biogaz afin de déterminer principalement la bioconversion des substrats en méthane. Un suivi de la production d'acides gras volatils (AGV) et de carbone organique/inorganique dissous est également réalisé afin d'identifier une possible accumulation d'AGV et de suivre le carbone encore non converti en biogaz. Finalement, une analyse du compartiment microbien est réalisée par ARISA, pyroséquençage des gènes ARNr16S et qPCR (archées, bactéries, *Bacteroidetes* et *Firmicutes*) afin d'obtenir des éléments d'informations sur la structure des communautés microbiennes en lien avec les performances de dégradation obtenues. Plusieurs analyses physico-chimiques des différentes fractions ont également été réalisées afin d'identifier les effets des broyages et d'émettre des hypothèses permettant d'expliquer les résultats observés.

Les résultats obtenus sont présentés ci-après sous forme d'un article en préparation pour soumission à la revue « Bioresource Technology ».

I. Dry grinding of lignocellulosic substrates shapes microbial communities during methanization

Nelly Badalato¹, Nina Pourette¹, Bruno Pontoire², Laurent Mazéas¹, Alain Buléon², Xavier Rouau³, Ariane Bize¹

¹ Irstea, UR HBAN, F-92761 Antony, France

² INRA, UR BIA, F-44316 Nantes cedex 03, France

³ INRA, UMR 1208 IATE, F-34060 Montpellier, France

Abstract

The effects of dry mechanical pretreatments of lignocellulosic materials on their methanization were assessed using cardboard and wheat straw as models. Various methods and settings were used to finely grind these materials and generate eight substrate batches with contrasted properties. Methanization dynamics was analyzed in mesophilic anaerobic batch microcosms. Associated microbial communities were analyzed by fingerprinting and 16S rRNA gene tag pyrosequencing. The grinding methods tested had no influence on cardboard methanization performance. By contrast, a limited positive effect of fine grinding was observed during wheat straw methanization (~ +10% final methane yields for the greater effects). The structure and composition of microbial communities were affected by the nature of the pretreated substrate batch. Recruitment of bacterial species related to *Treponema* was noticeably observed for the cardboard samples. This suggests that subtle lignocellulose structure modifications participate to the selection and shaping of functional microbial communities specific for the material.

Keywords

Anaerobic digestion; Dry grinding; Lignocellulose; Municipal solid waste; Microbial community

Introduction

Municipal solid waste (MSW) methanization is a promising bioprocess as it permits waste-to-energy conversion as well as waste mass reduction. These two features have lead MSW methanization processes to the crossroad between energy demand and ecological concerns. The MSW biodegradable fraction is mainly composed of lignocellulosic materials, representing more than 50% of the MSW mass in developed countries. These materials include papers, cardboards, sanitary textiles and vegetal residues. Similarly to natural lignocellulose, lignocellulosic waste contains cellulose, hemicelluloses and lignin. The precise composition depends on the materials, from the almost pure cellulose materials such as cotton to the mixed materials such as green waste (Foyle *et al.*, 2007). These compounds of natural origin have a high energy potential but they are not readily degradable under anaerobic

conditions. This recalcitrance is due to their highly complex and heterogeneous structure and to the cellulose crystallinity (Mansfield *et al.*, 1999). The presence of lignin especially acts as a barrier to biological attack (Chang and Holtzapple, 2000).

Pretreatments of the lignocellulosic materials can contribute to an increased biodegradability. A lot of strategies have been developed, including physical/mechanical, chemical and biological pretreatments used alone or in combination (Hendriks and Zeeman, 2009; Taherzadeh and Karimi, 2008). All these processes are intended to reduce cellulose crystallinity and lignin content and to increase the surface area accessible for enzymatic hydrolysis (Arantes and Saddler, 2011; Chang and Holtzapple, 2000). Among the mechanical treatments to enhance lignocellulose digestibility, dry grinding processes have many advantages. They do not produce effluents and they are not accompanied with toxic compound release, susceptible to inhibit biological degradation (Palmqvist and Hahn-Hägerdal, 2000). Obviously, the main effect of grinding is the particle size reduction but depending on the employed technology, it can also modify fine structural properties such as cellulose crystallinity (Silva *et al.*, 2012). Therefore, grinding processes can largely improve lignocellulose hydrolysis and many studies have been focused on the effects of lignocellulose grinding on its biodegradation. Most of them established a positive correlation between particle size reduction and biodegradability (Motte *et al.*, 2013; Sharma *et al.*, 1988; Silva *et al.*, 2012; Yuan *et al.*, 2011), while no significant effect was observed in other studies (Pommier *et al.*, 2010; Rivers and Emert, 1988). These contrasted results are mainly linked to the starting material but also to the employed grinding methods (Vidal *et al.*, 2011).

During waste methanization, the bioconversion of lignocellulose to biogas relies on complex anaerobic microbial communities and involves several interdependent degradation steps (Li *et al.*, 2009). Most of the studies cited above primarily focused on the total bioconversion potential or sometimes employed solely pure enzyme cocktails, without specific examination of microbial populations. Consequently, there is limited knowledge on the microbial response to physical alterations of lignocellulosic substrates, and as lignocellulose hydrolysis mostly occurs at the substrate surface in anaerobic environments (Lynd *et al.*, 2002; McAllister *et al.*, 1994), invaluable insight into hydrolysis and bioconversion processes can be obtained from microbiological studies.

The objective of the present work is to assess the effects of wheat straw - a commonly studied material - and cardboard dry grinding on their subsequent methanization and on the associated microbial populations. Various grinding methods were employed and the materials were reduced down to micrometric particles.

Materials and methods

Substrate characterizations and grinding

Two lignocellulosic materials were used, wheat straw (WS) (*Triticum aestivum*) and flat cardboard (C). The raw wheat straw and cardboard batches were the same as those used in Motte *et al.* (2013) and in Abbassi-Guendouz *et al.* (2013) respectively. Prior to fine grinding, the materials were comminuted by cutting milling with a 4 mm screen (WS-KM) and this treatment was studied for the wheat straw only. For the wheat straw, an intermediate step of sieve-based grinding was added. Fine powders were obtained by subsequent ball milling (WS-BM1, WS-BM2, corresponding to distinct median particle sizes, and C-BM), impact grinding (WS-IG and C-IG) and air-jet milling (WS-JM). Different operational conditions were applied to obtain various median particle sizes. All grinding methods were performed as described in Silva *et al.* (2012). A total of 5 wheat straw material batches were subjected to anaerobic digestion: WS-KM, WS-BM1, WS-BM2, WS-IG and WS-JM, and a total of 3 cardboard material batches were similarly studied: C-KM, C-BM and C-IG. Particle size distributions and specific surface were obtained by laser diffraction granulometry as described in Silva *et al.* (2012). Crystallinity indices, estimated by x-ray diffractometry, carbon and nitrogen content and volatile solid (VS) and total solid (TS) were measured as described in Badalato *et al.* (in preparation) and are presented in Table I-1. Scanning electron microscopy was realized on six batches (WS-KM, WS-BM1, WS-BM2, WS-IG, C-KM and C-BM) as described in Badalato *et al.* (in preparation) except that no chemical fixation was used during samples preparation. For wheat straw, the detailed results of Van Soest fractionation were previously available in Motte *et al.* (2013). For cardboard, Van Soest fractionation was previously performed and the detailed results were kindly provided by the authors (Abbassi-Guendouz *et al.*, 2012).

Methanization microcosms

Bioconversion of the substrates batches during anaerobic digestion was studied in lab-scale batch incubation reactors. They consisted of 330 mL glass bottles closed with a screw cap and a septum. Reactors were filled with 0.6 g dry matter (DM) of substrate mixed with a mineral solution (EN-ISO-11734, 1998 and Lü *et al.*, 2013), with a carbonate buffer (0.25 M) and an inoculum. The inoculum originated from an anaerobic solid-waste and biowaste digester operated in mesophilic conditions: the digestion sludge was sieved, stabilized at 35°C, centrifuged (speed, temperature time) and finally aliquoted and stored at -80°C to serve as inoculum. Inoculum had a 35.49/14.26 for TS (%)/VS (%) and a 23.6% total carbon content. It was added to each reactor to a final inoculum-to-substrate ratio of 1.5 volatile solids bases (VS). The mineral solution was added to reach a final volume of 220 mL and the headspaces were flushed with N₂ gas to obtain anaerobic conditions. The reactors were then incubated

at $35\pm 2^\circ\text{C}$. Each reactor condition was operated in triplicate, except for the reactor condition WS-KM that was operated in duplicate. A control reactor without substrate was also conducted.

Analytical methods

Biogas production was monitored over time by pressure measurements and biogas composition was analyzed with a gas chromatograph as described in Qu *et al.* (2009b). Liquid and solid phase analyses were performed on 1.5 mL samples collected through the septum using a syringe equipped with a needle. The fresh samples were centrifuged at 10 000 g for 10 min. Both the pellets (for microbiological analyses) and the supernatants (for chemical analyses) were stored at -80°C . Dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) concentrations were measured in the supernatant with a BIORITECH 700 analyzer. Volatile fatty acids (VFA) concentrations were measured using a DX 120 Ion Chromatograph (Dionex) with an IonPac ICE-AS1 column. Final degradation yields were calculated by estimating the amount of carbon released from the substrate compared to the initial carbon content, based on the final values obtained for DOC, DIC and CH_4 and CO_2 gas cumulated productions. For DOC and DIC, the final concentrations were subtracted from the concentrations measured at day 0. The biogas isotopic composition was analyzed by determining $\delta^{13}\text{C}_{\text{CH}_4}$ and $\delta^{13}\text{C}_{\text{CO}_2}$ values and calculating the apparent fractionation factor α_c (Qu *et al.*, 2009a).

Microbial community analyses

Total DNA was extracted from the pellets sampled at days 0, 18 or 22 and 31 using the PowerSoil™ DNA isolation kit (MoBio Laboratories) according to the manufacturer's instructions. Before amplification, DNA was quantified with the Qubit® dsDNA HS assay kit (Invitrogen). For bacterial and archaeal fingerprinting analyses by Automated Ribosomal Intergenic Spacer Analysis (ARISA), the 16S-23S rRNA intergenic transcribed spacers (ITS) were amplified by polymerase chain reaction (PCR) using the ITSf/ITSr bacterial primer set (Cardinale *et al.*, 2004) and the 1389F, modified from Loy *et al.* (2002) (5'-CTT GCA CAC ACC GCC CGT C-3')/71R (García-Martínez and Rodríguez-Valera, 2000) archaeal primer set. PCR conditions were as follows : 1X PCR buffer, 0.75U of Taq DNA polymerase (ThermoScientific), 0.2mM (each) deoxynucleoside triphosphate, 1.5mM MgCl_2 , 0.2 μM (each) primer and 2 ng of DNA template in a final volume of 25 μL . The mixture was held at 95°C for 15 min, subjected to 35 amplification cycles of 95°C for 30 s, 49°C for 30 s (54.8°C for the archaeal primer set), 72°C for 1 min and to a final extension step at 72°C for 10 min. Amplified fragment separation by capillary electrophoresis and profile analyses were performed as described in Fechner *et al.* (2010), using the R package StatFingerprints (Michelland *et al.*, 2009). For tag pyrosequencing analyses, the primer set 340F - 806R was chosen to target archaeal populations and the primer set 27F - 519R to target bacterial populations and amplify variable regions, V1-V2-V3 and V3-V4 respectively of the 16S rRNA gene.

Multiplexed bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) was performed using the Titanium platform (Roche Applied Science) in a commercial facility (Research and Testing Laboratories, Lubbock, TX) and the results were analyzed as previously described in Lü *et al.* (2013). Most abundant OTUs sequences were identified using the EzTaxon-e server (<http://eztaxon-e.ezbiocloud.net/>; Kim *et al.*, 2012) on the basis of 16S rRNA sequence data. Rarefaction curves were calculated using the R package Vegan (Oksanen *et al.*, 2011). Quantitative PCR targeting *Bacteria*, *Archaea*, *Bacteroidetes* and *Firmicutes* populations were realized as described in Carlei *et al.* (in preparation) using primer sets described in Yu *et al.* (2005) for bacteria and archaea, and in Bacchetti De Gregoris *et al.* (2011) for *Bacteroidetes* and *Firmicutes*. Briefly, iQ Supermix (Bio-Rad) and iQ Sybr Green Supermix (Bio-Rad) were employed for bacteria and archaea and for *Bacteroidetes* and *Firmicutes* respectively. The amplifications were performed in a CFX96 real-time thermal cycler (Bio-Rad).

Results and discussion

Waste characteristics

The wheat straw and the cardboard were characterized prior to dry grinding, particularly to determine their chemical composition (Table I-1). Based on the chemical composition, the wheat straw appeared as more biodegradable than the cardboard as shown by the higher volatile solid content and the lower lignin content. Besides, the cellulose content was higher in the cardboard while the hemicellulose content was higher in the wheat straw. After dry grindings, particle size distributions were analyzed (Table I-2). For all the materials, two ranges of particle size were generated (~50-60 µm and ~170 µm), expectedly corresponding to two different ranges of specific surface area. However, specific surface area measurements are just indicatives since the granulometry method used can not measure internal porosity. Estimation of the crystallinity index, used to assess the proportion of crystalline region in cellulose, showed no difference among the grinding methods and intensities. Concerning the wheat straw batches it was previously shown that the employed dry grinding methods could have an effect on the cellulose crystallinity (Silva *et al.*, 2012) however the strongest effect on cellulose crystallinity was observed for ultra-fine particles, lower than 50 µm in size. Besides diffractograms showed an unexpected peak at 29.33° in cardboard batches probably reflecting the presence of additives, such as glue (Teghammar *et al.*, 2010).

Table I-1. Raw materials characteristics

Substrate	TS/VS (%/%)	C/N (%/%)	Lignin (%)	Hemicellulose (%)	Cellulose (%)
Wheat straw	93.2/86.6	44.5/22.4	11 ^a	35 ^a	50 ^a
Cardboard	94.1/74.4	37.7/18.9	15.6 ^b	14.1 ^b	65.8 ^b

^{a, b} Values obtained from Van Soest fractionation from Motte *et al.* (2013)^a and Abbassi-Guendouz *et al.* (2012)^b

Four wheat straw and two cardboard batches were analyzed by electron scanning microscopy (Fig. I-1). For wheat straw, in the 4mm cutting milled batches a smooth surface, with pores, was observed (Fig. I-1 A) and the cell wall structures were preserved (Fig. I-1 B). After the ball milling treatment, the wheat straw surface appeared as more irregular and cell walls were less distinct, consistent with the increase in treatment intensity and the particle size decrease (Fig. I-1 C to F). Therefore it could be suggested that the lignin-hemicellulose matrix is preserved during cutting milling compared to ball milling which still act as a barrier and lower cellulose accessibility. For cardboard, the cutting milling produced a more fibrous material than the ball milling. For both batches no specific structural pattern were observed.

Table I-2. Batch reactors and particle characteristics.

Substrate	Treatment	Reactors	Median particle size (μm)	Particle size span ^a	Specific surface (m^2/g)	Crystallinity index (%)
Wheat straw	Cutting milling	WS-KM	nd	nd	nd	nd
	Ball milling	WS-BM1	170	3.57	0.105	nd
		WS-BM2	65	5.64	0.323	35
	Impact grinding	WS-IG	163	2.90	0.1	35
	Air-jet milling	WS-JM	61	2.91	0.317	35
Cardboard	Cutting milling	C-KM	171	4.79	0.156	nd
	Ball milling	C-BM	54	5.14	0.339	30
	Impact grinding	C-IG	61	4.83	0.331	30

nd: not determined

^a the particle size span characterizing the particle size distribution was calculated according to Silva *et al.*, (2012), as follows: $(d_{90}-d_{10})/d_{50}$, where the d_{10} , d_{50} or d_{90} values represent the 10th, 50th and 90th percentile of total volume.

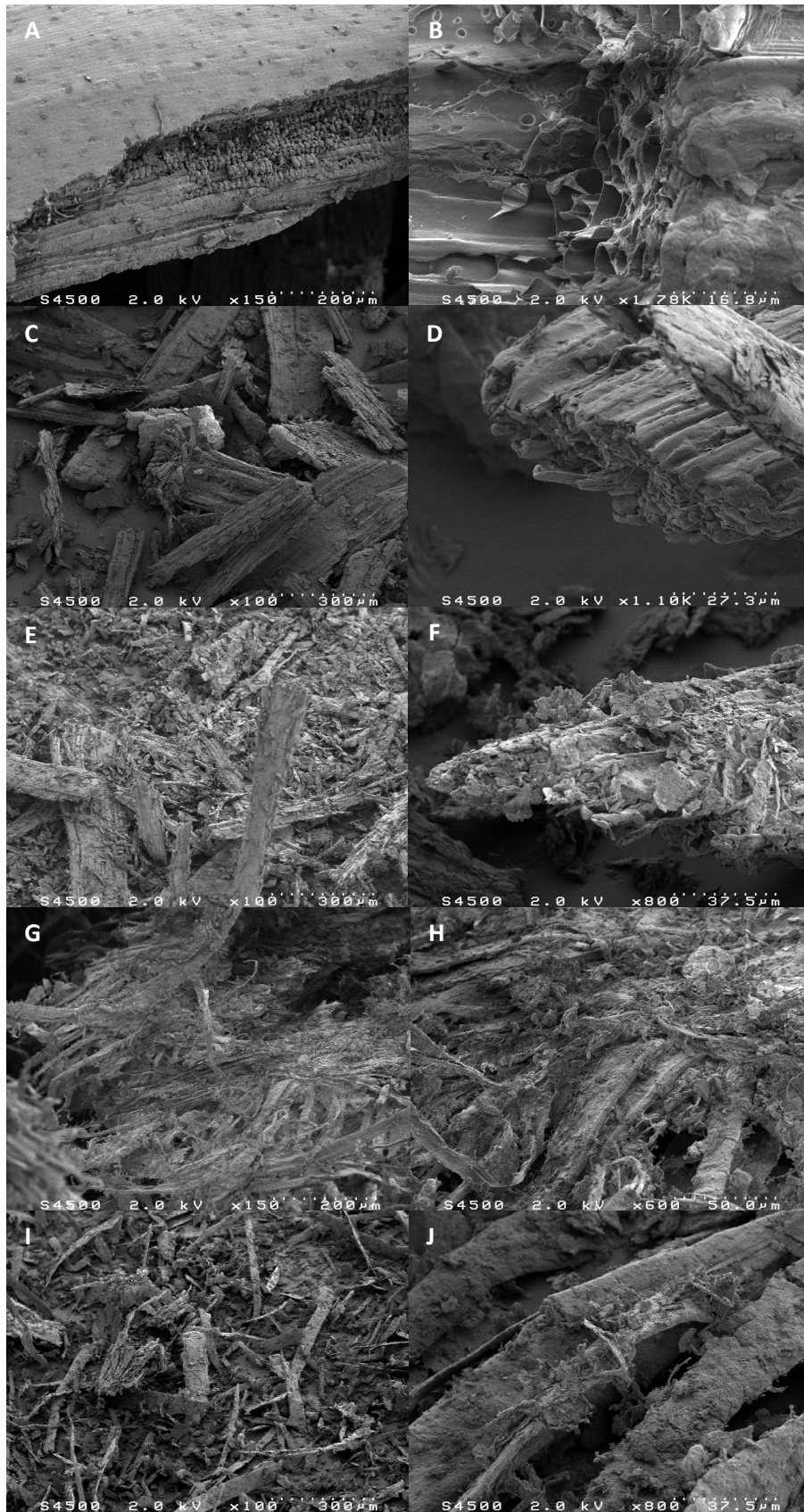


Fig. I-1. Scanning electron micrographs of wheat straw (A-F) and cardboard (G-J) grinded batches. (A-B) 4mm cutting milled wheat straw (WS-KM), (C-D) 170 μm ball milled wheat straw (WS-BM1), (E-F) 65 μm ball milled wheat straw (WS-BM2), (G-H) 171 μm cutting milled cardboard (C-KM) and (I-J) 54 μm ball milled cardboard (C-BM).

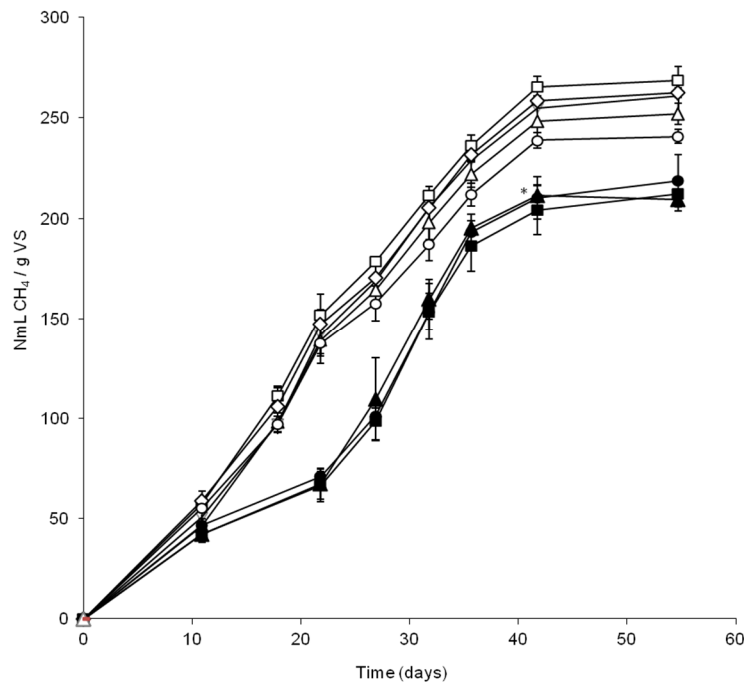


Fig. I-2. Cumulative methane production during anaerobic batch experiments.

Wheat straw fractions (empty symbols) and cardboard fractions (filled symbols): WS-BM2 (□), WS-BM1 (△), WS-IG (▽), WS-JM (◇), WS-KM (○), C-IG (■), C-KM (▲) and C-BM (●). Error bars are the standard deviations calculated for the triplicate samples analyzed at each time point. *Biogas composition was not analyzed at this time point, values were estimated based on the cumulative biogas production at the same time point and the biogas composition at the previous time point.

Anaerobic digestion performance

Cumulated methane productions obtained for the different batches of wheat straw and cardboard are shown in Fig. I-2. For all the batches, a good reproducibility was observed among replicates, and no methane production was detected in the blank test (data not shown). The observed methane production kinetics were similar to those commonly described for the anaerobic degradation of a lignocellulosic substrate (Pommier *et al.*, 2010; Qu *et al.*, 2009b). Based on the curves, 3 distinct production phases occurred, a first increase with contrasted production rates according to the material, followed by a rapid production, and finally stabilization as degradation neared completion. Similar trends were observed for total biogas production and a plateau was reached after 50 and 41 days for wheat straw and cardboard incubations respectively (data not shown). Methane and carbon dioxide were the major products detected in the gas phase and hydrogen was detected only at very low concentrations (less than 1%). The kinetic behavior shown by the cardboard batches was overall different from that observed for the wheat straw ones, the methane production appearing as less regular over time for the cardboard. The methane production kinetics obtained for all wheat straw fractions were similar (Fig. I-2, empty symbols). Average maximum methane yields reached during

incubation for each treatment are reported in Table I-3. The WS-BM2 reactor showed the highest methane yield with 268.7 NmL CH₄/g VS. This value corresponded to an 11.6% increase compared to the lowest methane yield, obtained for the 4 mm cut fraction (WS-KM). Values obtained for WS-IG and WS-JM reactors were close to the one obtained for WS-BM2 reactor. By contrast, WS-BM1 methane production was slightly higher than the value obtained for the WS-KM reactor. Chemical analyses in the supernatant revealed that acetate and propionate were produced at low levels. No VFA nor DOC accumulation were observed, but the sampling time point intervals probably did not permit to measure the maximal value for VFA and DOC concentrations, which were probably reached during the first week of incubation. Besides, final degradation yield values showed that more substrate was utilized in WS-BM2 reactor than in WS-KM reactor reaching a maximum of around 42% (Table I-3). Hence, it appears that ball milling with the most reduced particle sizes had the strongest effect on wheat straw methanization. Moreover, the batch treated with impact grinding, with the lowest specific surface, showed similar performances as those obtained with air-jet milling. These results are in good agreement with those obtained by Silva *et al.*, where ball milled wheat straw showed the higher release of reducing sugars after hydrolysis with pure enzymes, compared to air-jet milling and impact grinding for equivalent particle size (~50 μm) other grinding treatments (Silva *et al.*, 2012).

Table I-3. Final methane and degradation yield obtained in batch experiments with the grinded substrates.

Reactors	Average maximum methane yield (NmL CH ₄ /g VS)	Average final degradation yield (%)
WS-KM	240.8 ± 3.3	30.5 ± 1.4
WS-BM1	252.0 ± 5.2	35.8 ± 2.0
WS-BM2	268.7 ± 6.9	41.7 ± 3.5
WS-IG	260.9 ± 6.5	35.1 ± 1.8
WS-JM	262.6 ± 2.4	36.4 ± 1.7
C-KM	209.1 ± 5.5	31.7 ± 4.1
C-BM	218.3 ± 12.8	28.4 ± 0.8
C-IG	212.2 ± 2.6	29.2 ± 2.2

Since no decrease in cellulose crystallinity was observed with these treatment conditions it appears that this parameter is not a determinant factor in these conditions for wheat straw anaerobic degradation. In addition since the performances obtained with impact grinding was similar to those obtained with ball milling, with lower average particle size, suggests that decreasing particle could not

be the only important factor to enhance substrate biodegradability. Hence, other properties may be involved such as degree of polymerization or porosity and could be explored to better understand their role in substrate recalcitrance (Mansfield *et al.*, 1999). Among the cardboard pretreated batches, methane production kinetics were all highly similar (Fig. I-2). By contrast to what was observed for the wheat straw batches, no differences were observed among the 3 tested cardboard batches regarding maximum methane yields (Table I-3). These values are similar to those obtained in Pommier *et al.* (2010). However, in contrast to the results obtained for the wheat straw fractions, no significant differences were observed between the 3 studied batches concerning the maximum methane yields (Table I-3). These results also showed that the cardboard is more recalcitrant to degradation than the wheat straw in these conditions, since the average final degradation yields were around 30% for the cardboard batches compared to 30-40% for the wheat straw batches. The absence of grinding effects on cardboard methanization may be due to the more homogeneous nature of this material at the micrometer scale. Therefore, it appears that the sole surface area increase was not a major determinant for cardboard methanization with the employed operating conditions. However it could not be completely excluded since no higher particle size were tested.

Microbial community structures

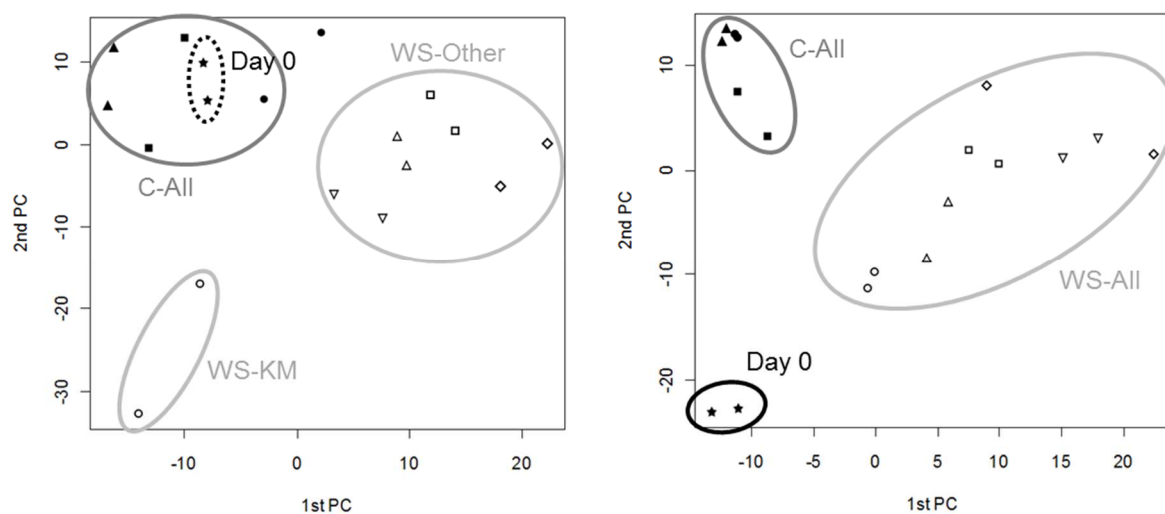


Fig. I-3. PCA biplots of bacterial (left) and archaeal (right) ARISA profiles from day 0 and 31. Representative plots obtained for samples, at day 0 (★) and day 31 for WS-BM2 (□), WS-BM1 (△), WS-IG (▽), WS-JM (◇), WS-KM (○), C-IG (■), C-KM (▲) and C-BM (●). The explained variances for principal components 1 and 2 are respectively 40.50% and 23.38% (left), 40.12% and 30.01% (right).

Microbial community structures were studied by ARISA for samples collected at day 0 and day 31 during the phase of stable methane production (see Materials and methods). Based on the obtained electrophoregrams, the profiles were similar at day 0 for all the reactors, for bacteria and archaea respectively (data not shown). The profiles were analyzed by principal component analysis (PCA) and representative biplots are shown in Fig. I-3. These two analyses primarily showed that the cardboard degrading community profiles were well separated from the wheat straw degrading community profiles. Fig. I-3 (left) revealed that the bacterial community profile from the WS-KM reactor was very distant from the other wheat straw degrading community profiles. In addition, a more important profile shift occurred over time in wheat straw containing reactors compared to cardboard containing reactors. Finally, cardboard community profiles appeared less complex than wheat straw community profiles based on the electrophoregrams (data not shown). This difference is possibly due to the presence of a cardboard adapted community in the inoculum, but other factors cannot be excluded, such as a more advanced degradation stage for wheat straw reactors compared to cardboard reactors. Within archaeal communities, the profile shift between day 0 and day 31 was more strongly marked than for bacteria (Fig. I-3). Similar to the results obtained within the bacterial community, the cardboard archaeal community profiles was quite distant from the wheat straw archaeal community. All together, these results suggest that the substrate grindings led to broad microbial population changes. This is consistent with the general knowledge on lignocellulose degradation, since lignocellulose hydrolysis mostly occurs at the substrate surface in anaerobic environments (Lynd *et al.*, 2002; McAllister *et al.*, 1994). It seems that microbial functional groups other than lignocellulolytic microorganisms are also impacted by the treatments, as shown by the differences within archaeal community profiles. Indeed, archaea in anaerobic digesters are generally involved in methanogenesis and not in lignocellulolysis. These changes in functional groups distant from lignocellulose deconstruction could result from the presence of different bacterial communities leading to different microbial associations and also to differences in the fermentation products profiles and/or in their production kinetics. It is noteworthy that the cutting milling on straw seems to have an impact on the substrate sufficient to lead to the selection of distinct microbial communities.

To study in more details the microbial community structure, 16S rRNA gene pyrosequencing data were generated and analyzed for one replicate of WS-KM, WS-BM2 and C-IG and for two time points during the methane production phase (day 18 and 27 for straw and 22 and 27 for cardboard batches) and the inoculum. In total 16 516 and 17 186 sequences were retrieved for *Bacteria* and *Archaea* domains respectively. Rarefaction curves, estimating the number of observed OTUs in function of the number of sequences included in random dataset subsamples, were calculated for each sample and microbial

domain. They did not reach a plateau, suggesting that only abundant OTUs were efficiently retrieved (data not shown).

For *Bacteria*, sequences assigned to the *Bacteroidetes* phylum were dominant in all samples (data not shown). Sequences assigned to the *Firmicutes* phylum, on the other hand were less abundant (data not shown). This observation contrasts with the results from a similar previous study of the bacterial diversity during pure cellulose digestion in batch mesophilic microcosms, where the *Firmicutes* phylum dominated (Li *et al.*, 2009). The main difference in OTU affiliation at the phylum level between the wheat straw and cardboard samples was the higher abundance of *Spirochaetes* in the cardboard samples. In-depth analyses of the dataset showed that for all the wheat straw and cardboard samples, the most represented genera (except *Petrimonas* and *Treponema*) showed quite similar proportions and evolutions between the two time points, suggesting that differences observed on ARISA profiles may be linked to different compositions at lower taxonomic levels. The genus *Alkaliflexus* was the most represented and was particularly dominant in finely grinded batches, WS-BM2 and C-IG (Fig. I-4). Cultured species the closest from the dominant OTUs affiliated to the *Alkaliflexus* genus were *Natronoflexus pectinivorans* and *Alkalitalea saponilacus* (>93% sequence identity) according to the EzTaxon-e server (Kim *et al.*, 2012) (data not shown). These two species were isolated from soda lakes and can grow on pectin and/or xylan (Sorokin *et al.*, 2011; Zhao and Chen, 2012). It suggests that the organisms from the retrieved OTUs could participate in hemicellulose degradation and be selected by fine grindings. Regarding *Spirochaetes*, OTUs related to the genus *Treponema* were abundant in all cardboard samples, representing around 20% of all the retrieved sequences (Fig. I-4). The closest cultured match was *Treponema lecithinolyticum* (~90%) a saccharolytic species (Wyss *et al.*, 1999). Despite the low sequence similarity it suggests that these microorganisms may participate, at least partially, in carbohydrates fermentation, since known *Treponema* species have already been described in lignocellulose degrading environments (Stanton and Canale-Parola, 1980; Warnecke *et al.*, 2007). Another interesting result is that more sequences were affiliated to the genus *Acetivibrio* in the wheat straw ball milled samples (~3% of total sequences) compared to other samples (data not shown). The closest cultured match for these OTUs were *Clostridium straminisolvens* (>95%). This is not a surprising result since *Acetivibrio sp.* and *C. straminisolvens* are cellulolytic species and are both members of the *Ruminococcaceae* family (Kato *et al.*, 2004; Ludwig *et al.*, 2007; Patel *et al.*, 1980). Since the OTUs related to known cultured cellulolytic or hemicellulolytic microorganisms represented a low proportion of the assigned sequences it seems that the most abundant microorganisms involved in lignocellulose degradation in the reactors could be distant from the known cellulolytic or hemicellulolytic cultured microorganisms.

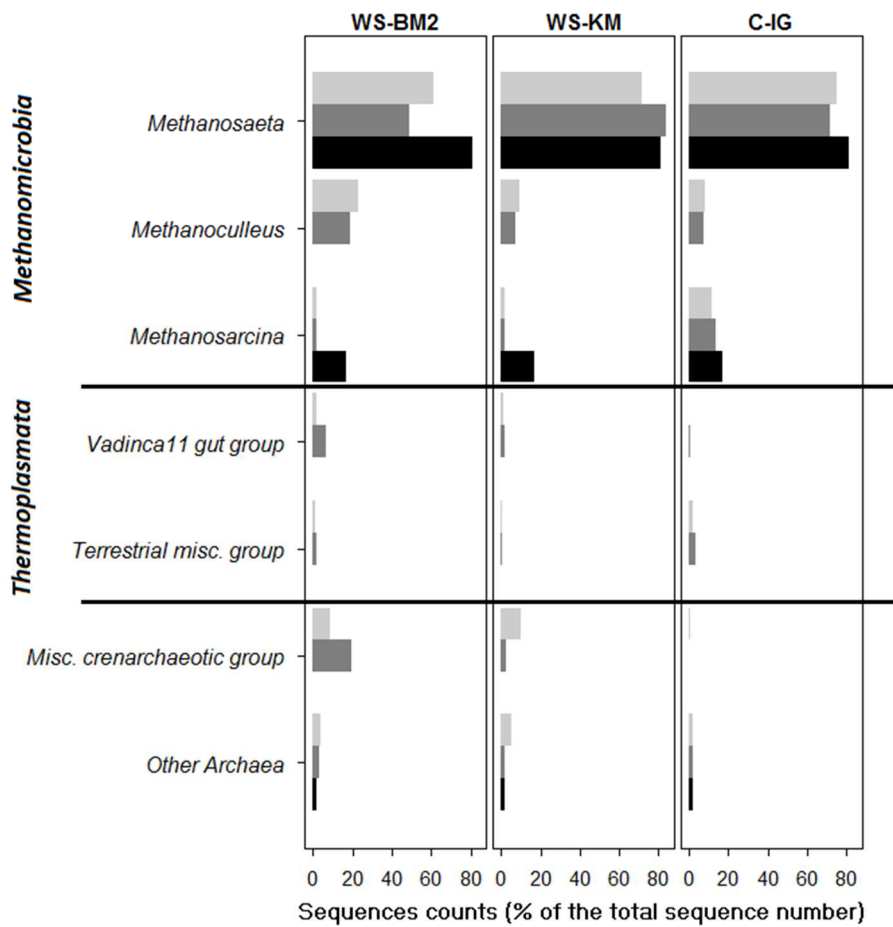
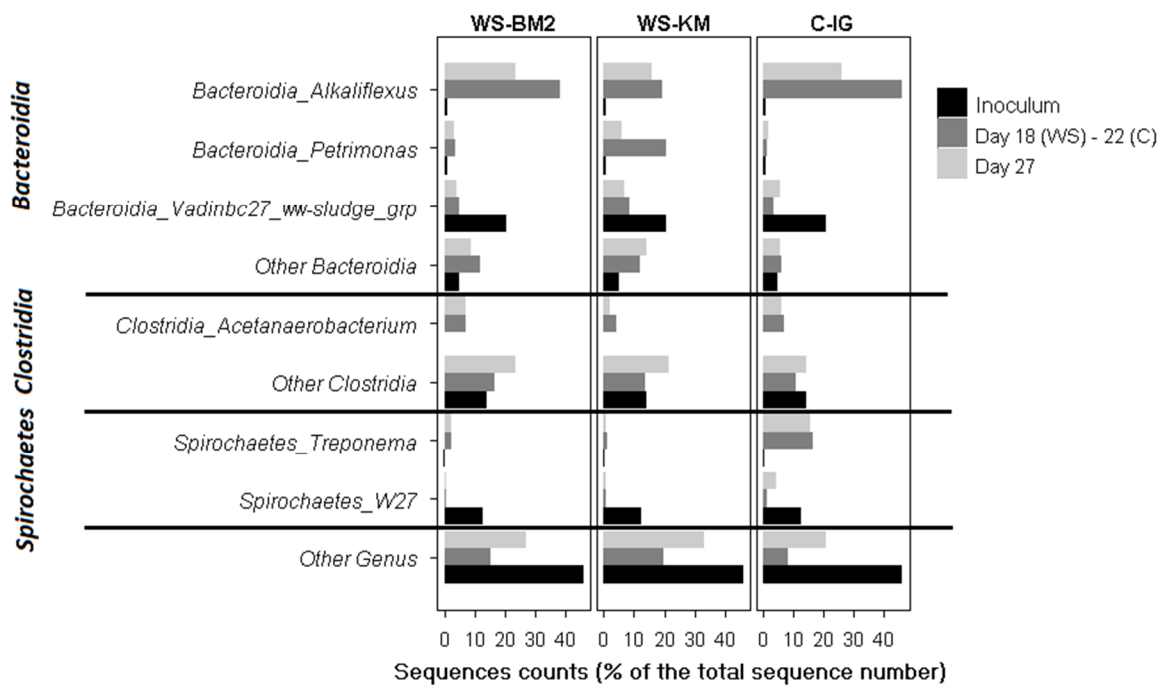


Fig. I-4. Overview of bacterial (top) and archaeal (bottom) diversity from 16S rDNA pyrosequencing. Results are shown for the most abundant genus (> 4% of total sequences count) for both domains. Results were retrieved from one reactor for each condition.

Quantitative real-time PCR was performed to target four microbial populations, *Bacteria*, *Archaea*, *Bacteroidetes* and *Firmicutes* (Fig. I-5). No significant effects on the bacterial and archaeal population growth were observed among the three conditions, since similar levels were detected. Surprisingly, these results were not consistent with the 16S rRNA pyrosequencing analysis since *Firmicutes* appeared as the dominant populations in 16S copy numbers compared to *Bacteroidetes*. Also levels measured for these two phyla were very low compared to those measured for *Bacteria*, probably due to a lack in universality and specificity of primers sets used to target these populations.

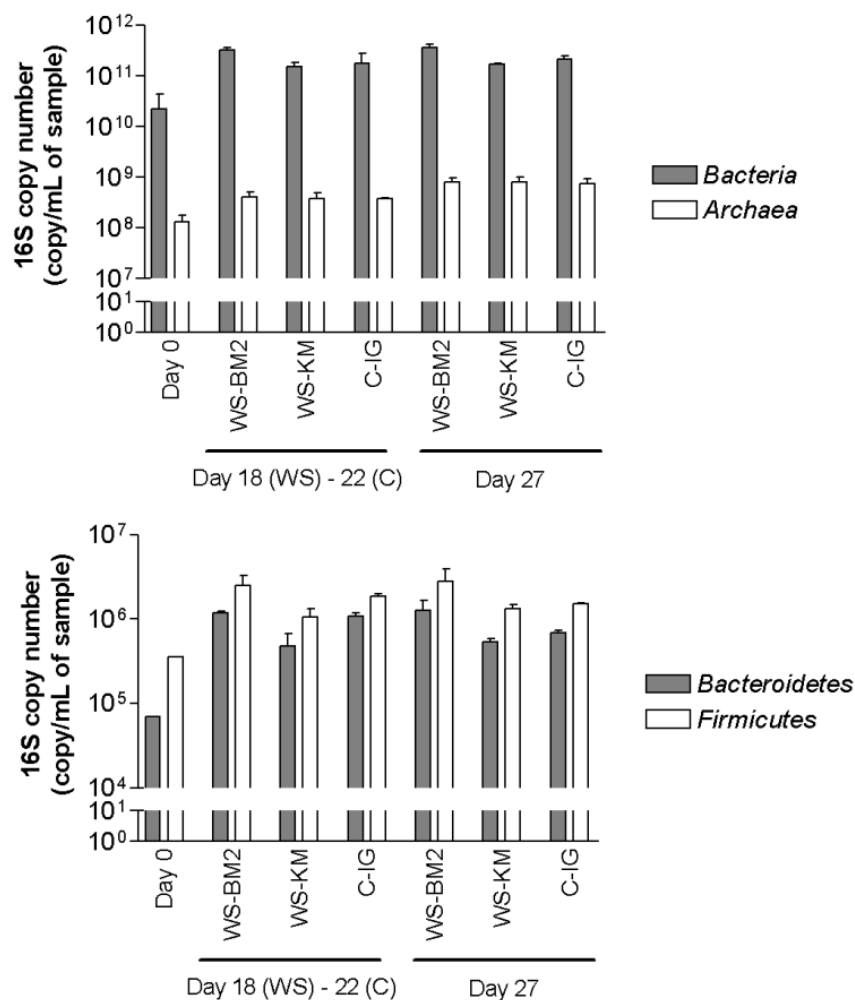


Fig. I-5. Quantitative real-time PCR of bacterial and archaeal (top) and *Bacteroidetes* and *Firmicutes* populations (bottom).

By contrast to the highly diverse bacterial community, the archaeal community was largely dominated by OTUs related to the phylum *Methanomicrobia* in all samples and more precisely to the *Methanosaeta* genus (Fig. I-4, bottom). The closest cultured match for these OTUs were *Methanosaeta concili* (>95%). Other OTUs related to the genus *Methanoculleus* were also abundant but they were

not detected in the inoculum. The closest cultured match was *Methanoculleus bourgensis* (96%). Finally, OTUs related to the genus *Methanosarcina*, that were closely related to *Methanosarcina mazei* (>99%), were abundant in the inoculum and cardboard samples but not in the wheat straw samples. From the great abundance of the genus *Methanosaeta*, it seems that the acetoclastic methanogenesis could be the dominant pathway for methane production, but there was a relatively high proportion of OTUs belonging to the genus *Methanoculleus*, which suggests that the hydrogenotrophic methanogenesis pathway could also have been activated. The activation and dominance of the acetoclastic methanogenesis pathway in all the reactors was confirmed by isotopic composition analyses of the biogas (data not shown).

Conclusions

The effects of wheat straw and cardboard dry grinding on their methanization and on associated microbial populations were assessed. No influence was observed on cardboard methanization performance. By contrast, a limited positive effect of high particle size reduction by ball milling was observed during wheat straw methanization (~ +10% final methane yields). The structure and composition of microbial communities responsible for lignocellulose-to-methane bioconversion was affected by the nature of the pretreated substrate batch. Recruitment of bacteria from the *Treponema* genus was noticeably observed for the cardboard samples. Subtle lignocellulose structure modifications could thus participate to the selection and shaping of functional microbial communities specific for the material.

Acknowledgments

This work was supported by Région Ile-de-France (R2DS 2010-08 project). The authors acknowledge Claire Dumas and Renaud Escudié (INRA UR LBE, France) for graciously providing the native wheat straw and cardboard used in this study. Finally, the authors acknowledge Franck Seara and Said El Adami for providing digestion sludge from a French mesophilic waste methanization plant (SYVOM, Urbasys SAS).

Points-clés

Confirmations :

- L'effet positif mais limité du broyage fin de la paille (Sharma *et al.*, 1988)
- Pas d'effet du broyage du carton, même à une échelle plus fine (Pommier *et al.*, 2010)

Nouveaux éléments :

- Communautés sélectionnées distinctes entre la paille et le carton, principalement *via* la sélection de bactéries proches du genre *Treponema* lors de la dégradation du carton
- Sensibilité des communautés microbiennes aux changements de structures fines des matériaux lignocellulosiques
- Le taux de cristallinité de la cellulose ne semble pas être un facteur prépondérant, cependant les conditions de l'étude ne permettent pas de conclure sur l'absence d'effet de ce facteur

CHAPITRE 2

Avant-propos

Comme mentionné précédemment, les travaux réalisés sur des systèmes modèles permettent une étude plus approfondie des mécanismes cellulaires influencés par les différentes conditions grâce aux informations déjà disponibles. Les deux souches choisies sont *Clostridium cellulolyticum* et *Acetivibrio cellulolyticus*, qui sont proches des souches présentes dans les méthaniseurs.

La première phase de l'étude de systèmes modèles consistait en la mise au point d'une technique de mesure de l'adhésion par cytométrie en flux. Cette technique a déjà été appliquée au suivi de l'adhésion de bactéries sur des particules synthétiques (Beloin *et al.*, 2008). Le principe ici était d'utiliser des billes de cellulose comme seule source de carbone pour la culture des souches pures cellulolytiques, *C. cellulolyticum* et *A. cellulolyticus*, et de suivre la quantité de bactéries adhérees au cours du temps et dans différentes conditions. Le comptage des bactéries aurait été rendu possible grâce à un colorant cellulaire, non adsorbable par les billes, qui émet un signal lorsque les billes passent devant le laser du cytomètre. Par cette technique il aurait été également possible de suivre la taille des billes et donc leur vitesse de dégradation. Les résultats obtenus ont montré que les billes de cellulose sont dégradées par les deux souches étudiées, leur taille, évaluée par des observations en microscopie, diminuait et des produits de fermentation apparaissaient au cours du temps. Cependant, aucune colonisation du substrat n'a pu être observée. Une analyse des billes de cellulose a montré qu'une substance était également présente en plus de la cellulose et pourrait correspondre à une résine entourant les billes et empêchant la colonisation. Suite à ces conclusions, l'utilisation de la méthode de mesure de l'adhésion de bactéries par cytométrie en flux a été abandonnée.

Afin de poursuivre l'étude des modèles biologiques, d'autres substrats celluloses ont été testés : cellulose microcristalline en poudre, papier filtre, coton en disques, mouchoir en papier et membrane de cellulose. A l'exception du mouchoir en papier, ces substrats sont des modèles déjà cités dans la littérature pour l'étude de la dégradation de la cellulose. Des travaux préliminaires ont été réalisés afin de suivre la colonisation et la dégradation de ces différents substrats par les deux souches cellulolytiques. Les résultats ont montré que tous ces substrats étaient colonisés et dégradés par les deux souches, à l'exception du coton en présence d'*A. cellulolyticus*. Cependant les vitesses de dégradation étaient différentes, le mouchoir en papier étant le plus rapidement dégradé et le coton le moins rapidement, comme cela a pu être mis en évidence par l'analyse des concentrations en COD et AGV présents dans la phase liquide. Les observations en microscopie ont par ailleurs révélé des profils

de colonisation différents selon les substrats. Enfin, un profil fermentaire différent a été observé chez *C. cellulolyticum* en présence de mouchoir en papier (Kleenex[®]) en comparaison avec les résultats obtenus avec les autres substrats.

Dans le but d'approfondir les résultats, une nouvelle expérience a été réalisée (Figure 13). Les substrats choisis pour cette deuxième expérience étaient le mouchoir en papier, le papier filtre et le coton en disques, afin d'obtenir des résultats contrastés en terme de profils de colonisation et de dégradation, et la souche choisie était *C. cellulolyticum*, en raison de la différence observée dans les profils fermentaires. Les cultures ont été réalisées en bouteilles, en conditions discontinues (« batch ») et non agitées. D'autre part, afin de suivre plus finement la colonisation, des microplaques, placées dans des sachets et en anaérobiose, ont été utilisées en parallèle pour réaliser des cultures dans des conditions similaires avec un nombre de réplicats suffisant avec des observations microscopiques fréquentes. Les microplaques permettent également de s'affranchir de biais de prélèvement qui favoriserait l'échantillonnage des plus petites particules. Le milieu utilisé pour toutes les incubations était un milieu minéral classique pour la culture de souches pures anaérobies, il a été testé et validé pour la culture de toutes les souches microbiennes étudiées pendant la thèse.

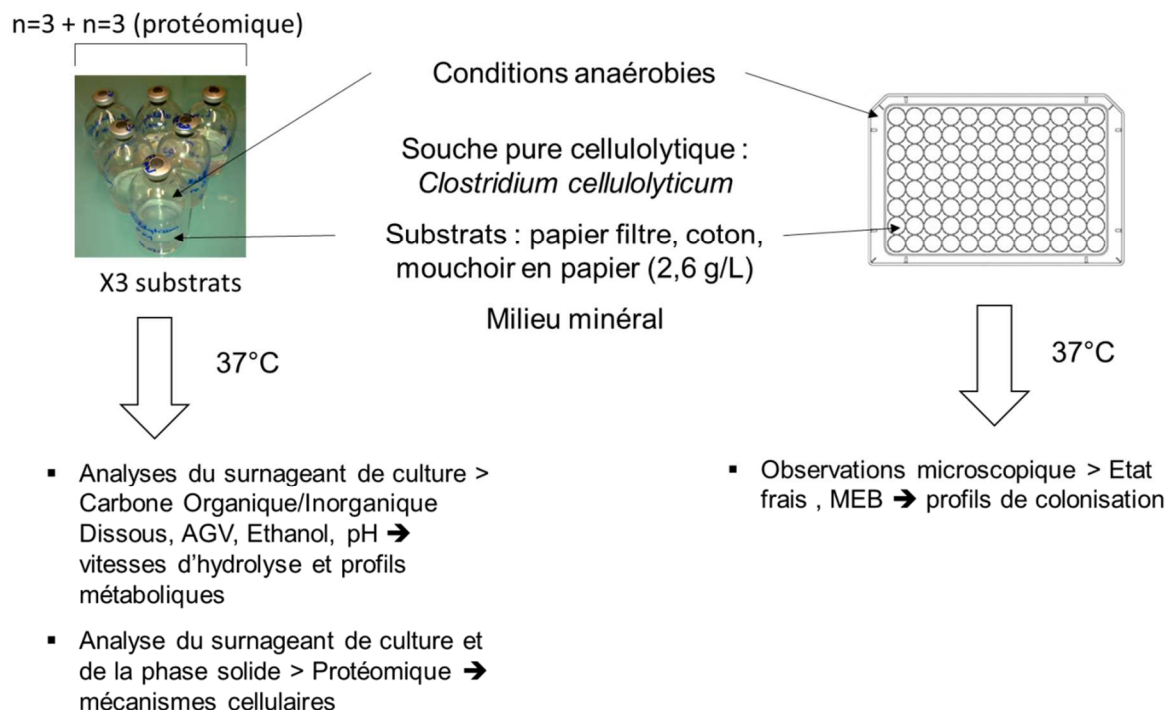


Figure 13. Dispositif expérimental réalisé pour l'étude de la dégradation de trois substrats cellulosiques par *C. cellulolyticum*.

Les analyses réalisées sur les incubations en bouteilles consistaient en un suivi de la production de carbone organique dissous (COD) afin d'évaluer la dégradation des substrats en molécules solubles. La production d'acétate, de lactate et d'éthanol, produits majoritaires de la fermentation, a aussi été suivie pour identifier les profils fermentaires dans les différentes conditions. Le pH est également un paramètre important et a été suivi tout au long de l'expérience. La composition du biogaz produit n'a pas pu être suivie au cours de l'expérience, en raison du dispositif utilisé qui ne permet pas d'assurer le maintien de conditions stériles et qui n'est pas adapté à l'analyse de petits volumes de gaz. Concernant le suivi de la colonisation, des microplaques ont été sacrifiées à chaque point de temps, pour permettre la coloration des cellules et l'observation en microscopie confocale en état frais. D'autre part, une analyse de protéomique quantitative a été réalisée sur deux points de temps pendant la phase de dégradation rapide, lorsque la production de COD augmente fortement, sur des incubations avec du mouchoir en papier et du papier filtre. Le mouchoir en papier a été choisi en raison des profils de dégradation et métaboliques observés lors de sa dégradation. Le papier filtre a été également choisi car il s'agit d'un substrat très utilisé pour l'étude de la dégradation de la cellulose et permet de faire un lien avec la littérature. Les bouteilles utilisées pour cette analyse ont été sacrifiées puisque l'utilisation d'une grande quantité d'échantillon est préférable pour préserver la représentativité de l'échantillon. La démarche de l'analyse de protéomique est détaillée ci-après. Afin de compléter cette étude et de mieux identifier la phase limitante de la dégradation des substrats, plusieurs analyses physico-chimiques des différents substrats ont également été réalisées. Un dosage quantitatif des polysaccharides dans la phase soluble est en cours de développement au laboratoire ce qui permettrait de mesurer leur accumulation. Une étude de la composition des substrats simples par fractionnement Van Soest sera également réalisée. Cette méthode est couramment utilisée pour déterminer la composition des substrats lignocellulosiques et permettra la comparaison avec d'autres substrats, en particulier les substrats complexes utilisés pendant la thèse pour lesquels ces données ont déjà été obtenues.

Les résultats obtenus sont présentés ci-après sous forme d'un article en préparation pour soumission à la revue « Molecular Systems Biology ».

Compléments sur les aspects méthodologiques des analyses protéomiques

Les analyses protéomiques permettent d'étudier les fonctions exprimées par la biomasse. Le schéma général présentant le déroulement de cette analyse est présenté ci-dessous (Figure 14). Les principales étapes sont : l'extraction et la purification des protéines de l'échantillon, la fragmentation de ces protéines en peptides par digestion enzymatique, l'analyse de chaque mélange de peptides par chromatographie en phase liquide couplée à un spectromètre de masse en tandem, et finalement l'identification des protéines grâce aux spectres de masse des peptides entiers et fragmentés. La digestion trypsique et l'analyse en spectrométrie de masse ont été réalisées grâce à une collaboration avec la plateforme PAPPSO (INRA). Deux suites d'instruments de chromatographie et spectrométrie de masse, aux sensibilités et contraintes différentes, ont été utilisées pour les deux études protéomiques présentées dans ce travail de thèse, et seul le principe général de cette étape sera présenté ici.

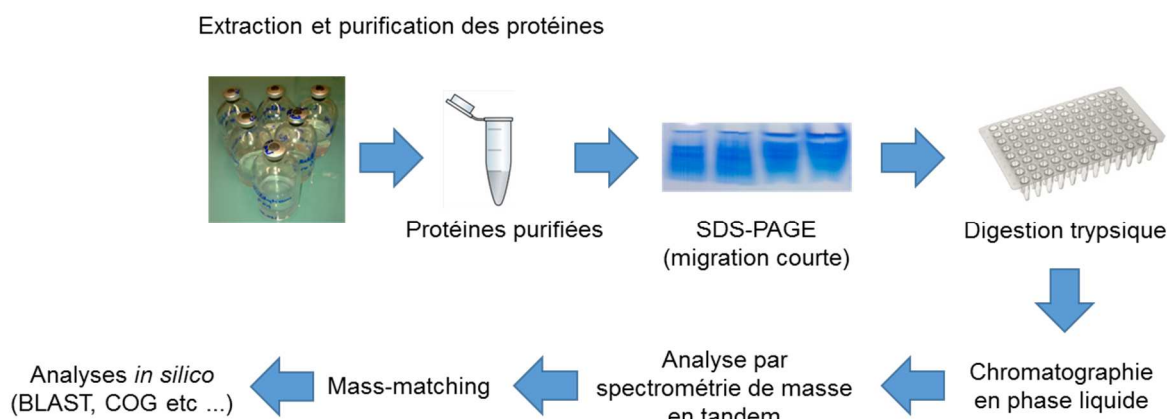


Figure 14. Aperçu du schéma expérimental pour l'analyse protéomique.

La première phase consiste en l'extraction et la purification des protéines. Dans le cadre de la thèse, deux méthodes ont été testées, l'une développée par Lü *et al.* (2013) basée sur une extraction à l'acide trichloroacétique et acétone, et l'autre développée par Hanreich *et al.* (2012) basée sur une extraction au phénol. En souches pures, les rendements d'extraction obtenus étaient légèrement supérieurs avec la méthode au phénol. Avec des biomasses complexes les rendements obtenus étaient aussi légèrement supérieurs pour les échantillons issus du travail de thèse, mais pas pour tous les échantillons issus d'expériences d'autres projets de recherche, suggérant un effet important du type de matrice analysée. Suite à ces expériences préliminaires, la méthode au phénol a été retenue pour l'analyse de l'ensemble des échantillons du travail de thèse. Cette méthode a de plus l'avantage de

comporter peu d'étapes et de permettre de traiter plus d'échantillons simultanément, limitant ainsi les biais liés à des séries d'extractions différentes.

La première étape de l'extraction des protéines consiste à lyser les cellules. Concernant la méthode retenue, il s'agit d'une lyse par sonication dans un bain de glace afin de limiter l'échauffement de l'échantillon. Le lysat, contenant les protéines solubles et quelques fragments de membranes cellulaires, est collecté et du phénol concentré y est ajouté. La phase organique contenant la majorité des protéines est ensuite collectée et la précipitation des protéines est réalisée par l'ajout d'acétate d'ammonium dans du méthanol froid. Après cette étape, les protéines sont contenues dans le culot qui subit plusieurs lavages à l'acétate d'ammonium dans du méthanol froid puis à l'acétone. Les protéines sont finalement solubilisées dans un tampon compatible (Tris-Urée) avec la plupart des méthodes de dosage et d'analyse. La méthode de dosage utilisée ici est un kit de quantification des protéines par puces basé sur le marquage fluorescent des protéines et un système de microfluidique pour la migration des protéines (Bioanalyseur, Agilent High Sensitivity Protein 250 Kit). La purification des protéines est complétée par une migration courte en gel SDS-PAGE (Sodium Dodecyl Sulfate-Polyacrilamide Gel Electrophoresis) d'une quantité égale de protéines pour chaque échantillon, puis un découpage des fragments de gels contenant les protéines

La deuxième phase est la digestion enzymatique, par la trypsine, des protéines contenues dans les fragments de gel. Les peptides contenus dans chaque mélange sont ensuite séparés par chromatographie en phase liquide (nanoHPLC ou nanoUPLC), ce qui permet d'éluer progressivement les peptides avant leur entrée dans le spectromètre de masse. La spectrométrie de masse ne permettant de détecter que des molécules chargées, une étape d'ionisation des peptides est effectuée en entrée du spectromètre de masse (Electron Spray Ionisation, Ionisation par électronébuliseur). Enfin, la masse de chaque peptide est mesurée dans un premier spectromètre de masse. A chaque cycle de spectrométrie de masse, les peptides majoritaires sont sélectionnés, fragmentés et les ions issus de la fragmentation sont analysés dans un deuxième spectromètre de masse en tandem. Il s'agit donc d'une acquisition donnée-dépendante. A l'issue de ces analyses, les spectres de masse correspondant aux peptides entiers et fragmentés, issus des protéines contenues dans l'échantillon, sont obtenus.

La dernière phase est l'analyse des spectres de masses obtenus. La principale méthode utilisée pendant la thèse consiste à comparer les données obtenues expérimentalement avec les données calculées à partir des protéines présentes dans une banque de données, choisie selon le type d'échantillon de départ. La première étape consiste en la création d'une liste de peptides théoriques dont les masses correspondent précisément à celles des peptides de l'expérience. La deuxième étape

consiste en la comparaison des spectres de fragmentation obtenus expérimentalement avec des spectres de fragmentation théoriques. La validité de cette association dépend d'un score calculé à partir du taux de similitude entre les spectres comparés. Le peptide de la base de données associé au score le plus élevé est retenu. Une « e-value » est également calculée pour chaque peptide, qui prend en compte de la validité du score par rapport aux scores des autres peptides dont la masse théorique peut être associée à la masse expérimentale. Une protéine est validée si deux peptides uniques au moins sont validés pour cette protéine et si les valeurs de « e-value » ne dépassent pas des seuils prédéfinis. Suite à l'ensemble de ces étapes une première liste de protéines est obtenue. Un nouveau cycle d'analyse peut-être effectué afin de prendre en compte les modifications post-traductionnelles supposées les plus courantes, dont la présence modifie les spectres de masse. Cette dernière étape permet d'augmenter le nombre de spectres et de peptides identifiés pour chaque protéine. Par définition, elle ne permet pas d'identifier de nouvelles protéines. Prendre en compte les modifications post-traductionnelles dès les premières étapes de l'analyse n'est généralement pas appliqué car cela pourrait aboutir à un fort taux de résultats faux-positifs, sans possibilité de les identifier au sein des résultats. Des analyses quantitatives peuvent également être réalisées afin de comparer les niveaux de protéines dans différentes conditions d'étude. Il existe plusieurs méthodes d'analyses et la principale méthode utilisée lors de la thèse a été réalisée dans le cadre d'une collaboration avec Omics Service et est détaillée dans l'article ci-dessous.

II. Integrated quantitative whole-proteome analyses of the biofuel-producing *Clostridium cellulolyticum*

Nelly Badalato¹, Alain Guillot², Victor Sabarly³, Marc Dubois³, Nina Pourette¹, Bruno Pontoire⁴, Sylvie Durand⁴, Arnaud, Bridier¹, Fabienne Guillon⁴, Véronique Monnet², Théodore Bouchez¹, Alain Buléon⁴, Gérard Mortha⁵, Ariane Bize¹

¹ Irstea, UR HBAN, F-92761 Antony, France

² INRA, UMR 1319 MICALIS, PAPPSO, 78352 Jouy-en-Josas cedex, France

³ Omics Services , 13bis avenue de la Motte Picquet, 75007 Paris, France

⁴ INRA, UR BIA, F-44316 Nantes cedex 03, France

⁵ Pagora, Grenoble Institute of Technology, 38402 Saint-Martin d'Hères, France

Abstract

Biofuel production from lignocellulosic waste and residues is an interesting option in the perspective of mitigating the environmental costs associated to energy production. However, their widespread utilization is hampered by the difficulty to overcome lignocellulos recalcitrance. In the frame of an integrated approach, colonization and metabolic patterns of the model mesophilic cellulolytic bacterium *Clostridium cellulolyticum* were analyzed during growth on 3 distinct cellulosic manufactured materials (facial tissue, filter paper, cotton). A faster and more widespread colonization was observed for facial tissue, followed by filter paper and cotton. The biofilm appearances were also qualitatively different. This behavior was associated to a faster degradation and a higher ethanol-to-acetate ratio for facial tissue compared to the other two materials. Label-free quantitative shotgun proteomics analyses were realized for facial tissue, the more reactive substrate, and filter paper, a previously studied model. Advanced comparative statistical analyses of the proteins identified in the pellets and supernatants enabled the identification of 187 proteins showing a different behavior according to the substrate. Noticeably, the overexpression of the many enzymes from the carbohydrate central metabolic pathway upstream the pyruvate node was strongly supported by the results. Moreover, the cellulosomes (important cellulolysis machineries present in *C. cellulolyticum* and other cellulolytic bacterial species) were shown to exhibit distinct average compositions according to the substrate.

Keywords

Clostridium cellulolyticum / colonization / label-free quantitative proteomics / lignocellulose / waste

Introduction

Cellulose bioconversion is the core of numerous bioprocesses for biofuel production since this biopolymer is carbon-rich and naturally abundant (first generation biofuel). Despite this widespread utilization, cellulose recalcitrance still limits the energy recovery yields and the process cost-efficiency, which is an obstacle to a broader implementation of biofuel production (second and third generation biofuel, currently in an exploratory research phase). This particularly holds for bioprocesses operated in anaerobic conditions (Jeihanipour *et al.*, 2011). Many factors account for cellulose recalcitrance such as accessible surface area, presence of other recalcitrant compounds such as lignin, degree of polymerization, porosity and cellulose crystallinity (Mansfield *et al.*, 1999). However, the precise contribution of each of them is difficult to assess and is only partially known at present.

Consolidated bioprocesses (CBP) have been described as a promising approach to overcome the recalcitrance key-issue. CBP rely on the utilization of pure microbial strains performing both the cellulolysis and the fermentation to biofuel, typically ethanol, in a single-step continuous process. Currently in their exploratory research phase, they are the subject of intensive investigation. In natural environments, cellulose bioconversion is mainly realized by widely distributed cellulolytic microorganisms using complex enzymatic machineries (Lynd *et al.*, 2002). In particular, *Clostridia* members, including *Clostridium cellulolyticum*, seem to play a major role for anaerobic cellulose digestion and are the most studied cellulolytic anaerobes so far (Lynd *et al.*, 2002). These bacteria are generally found adhered to cellulose, thus gaining a first access to cellulose hydrolysis sites and hydrolysis products (Lynd *et al.*, 2002; McAllister *et al.*, 1994). Some strains isolated so far are promising candidates for CBP for ethanol production (Higashide *et al.*, 2011; Jin *et al.*, 2011) and closely related members have been detected in anaerobic digesters treating high cellulose content waste (Li *et al.*, 2009).

In this context, cellulose recalcitrance and bioconversion has been extensively studied for biomass residues and agricultural waste and products (Hansen *et al.*, 2011; Himmel and Bayer, 2009; Ogeda *et al.*, 2012), the latter being currently used or studied as input for bioethanol production bioprocesses (first and second biofuel generation respectively). However, new sources of lignocellulosic materials are currently emerging as attractive options for biofuel production based on fermentation and they received a more limited attention until now in this perspective. Municipal solid waste (MSW) is an important source of cellulosic materials (e.g. sanitary textiles, papers, cardboards) since the latter account for up to 50% of the MSW mass in developed countries (ADEME, 2009; EPA, 2011). Using the cellulosic fraction of MSW for biofuel production potentially emits less greenhouse gases than using cultivated crops and does not compete for land with food production.

For the present study, we developed an integrative approach to investigate the colonization and metabolic patterns exhibited by the cellulolytic bacterium *Clostridium cellulolyticum* grown on three manufactured cellulosic substrates – facial tissue, filter paper and cotton pads – related to those found in MSW. *Clostridium cellulolyticum* is a model of mesophilic anaerobic cellulolytic bacterium (Desvaux, 2005). Its cellulose degradation system, a cellulosome associated to 143 Carbohydrate-Active Enzymes (CAZymes, Cantarel *et al.*, 2009) has been well studied upon different carbon sources. Information about cellulosome expression and modulation is already available (Blouzard *et al.*, 2010; Xu *et al.*, 2013). Metabolic characteristics have also been examined, especially upon cellobiose *versus* cellulose, showing metabolic nodes in the central metabolism pathways (Desvaux, 2006, 2004).

Here, a label-free quantitative shotgun proteomic approach was conducted, revealing distinct patterns regarding CAZymes, central metabolism and proteins related to substrate colonization. More than 1200 proteins were identified in total. Another strong point is the integration of detailed chemical and structural characterizations of the substrates, including cellulose crystallinity and degree of polymerization. Indeed, manufactured cellulosic materials are expected to harbor distinct features compared to the well-characterized native vegetal lignocellulose models, such as wheat straw. The possible role of the substrate features in the distinct colonization patterns is addressed: many studies were previously dedicated to the impact of cellulose characteristics on enzyme adhesion (Chang and Holtzapple, 2000; Hansen *et al.*, 2011; Rivers and Emert, 1988) but only few studies tackled their impact on bacterial colonization (Weimer *et al.*, 1990). Concerning the adhesion of cellulolytic bacteria, the effect of the substrate characteristics has been relatively less studied than environmental factors such as pH, temperature and soluble sugars, cations and cell concentrations (Roger *et al.*, 1990).

This work is the first global shotgun proteomic study of *C. cellulolyticum*. It provides detailed insights into the degradation strategy of this bacterium, an invaluable knowledge to design more efficient cellulose bioconversion processes.

Results

Substrate characteristics

Important structural and chemical characteristics of the cellulosic substrates used in this study – facial tissue, filter paper and cotton pads – were determined to assess their role in substrate recalcitrance. The main differences concerned mainly the crystallinity index (CI) and the molecular weight distributions. CI is commonly measured to estimate the amount of amorphous regions, more easily degradable and of crystalline regions, more recalcitrant, in cellulose. The molecular weight may also

have an impact on the substrate degradability. For convenience, these substrates will be referred to hereafter as tissue, paper and cotton, respectively.

Table II-1. Summary of the substrate's main characteristics

Substrates	TS/VS (%)	C/N (%)	COD (g/L)	CI (%)	DP
Tissue	96.0/95.8	43.1/0.03	1.14	94	~970
Paper	96.5/96.4	41.1/0.05	1.11	74	~1,950
Cotton	94.9/94.2	41.9/0.07	1.06	50	~7,700

TS/VS: Total Solid/Volatile Solid – C/N: Carbon/Nitrogen – COD: Chemical Oxygen Demand – CI: Crystallinity Index – DP: Degree of Polymerization corresponding to the highest peak of the molecular weight density distribution (Fig. II-1).

Measurement values of water, ash, carbon and nitrogen contents were highly similar for the three substrates (Table II-1). The low (<1%) ash contents indicated that they did not contain high amounts of minerals. Mineral substances such as calcium carbonate are frequently used to manufacture specific cellulosic materials and could thus have persisted at various levels in the final products. The carbon and nitrogen contents, close to 42% and <0.1% respectively, were compatible with high carbohydrate contents. In addition, the dominant chemical functional groups were the same in all substrates and typical of cellulosic materials (Fig. SII-1): based on the spectra obtained from FTIR analyses, primary and secondary alcohols and glycosidic bonds were identified. By contrast, great differences between the substrates were detected concerning the CI and the molecular weight distributions (Table II-1, Fig. II-1). For the three substrates, CI ranged from 50% to 90%, from the more amorphous cellulose in tissue to the more crystalline cellulose in paper. However, the precise values of this calculated index are greatly linked to the measurement technique used (Park *et al.*, 2010). Molecular weight composition analyses provide information about the degree of polymerization (DP) of carbohydrate molecules. The obtained molecular weight distributions showed that the three substrates had very distinct profiles (Fig. II-1). For paper, the curve was smooth with a single wide major peak and it was close to a log-normal distribution: it indicated an overall well-solvated, unaggregated and quasi pure cellulosic substrate. By contrast, the distribution profile obtained from tissue was wider and irregular, and for cotton, narrower. Peaks observed around $10^{6.6}$ (Fig. II-1, area 2) probably corresponded to molecules with a high molecular weight and/or to aggregates. Typically, cotton pads are made from natural cotton linters in which cellulose is highly polymerized; the production method used or the presence of additives can also generate abundant molecular aggregates. A peak around $10^{5.2}$ (Fig. II-1, area 1) was only observed for tissue and could be the signature of hemicellulosic molecules, which have lower molecular weights than cellulose (Xiao *et al.*, 2001). The DP corresponding to the highest peak of each distribution is reported in Table II-1.

Based on these characterizations, tissue was predicted to be the most accessible substrate due to a low CI, a low DP and the possible presence of hemicelluloses. Due to opposite trends regarding the CI and the DP, it was difficult to predict which of the cotton or the paper could be more biodegradable.

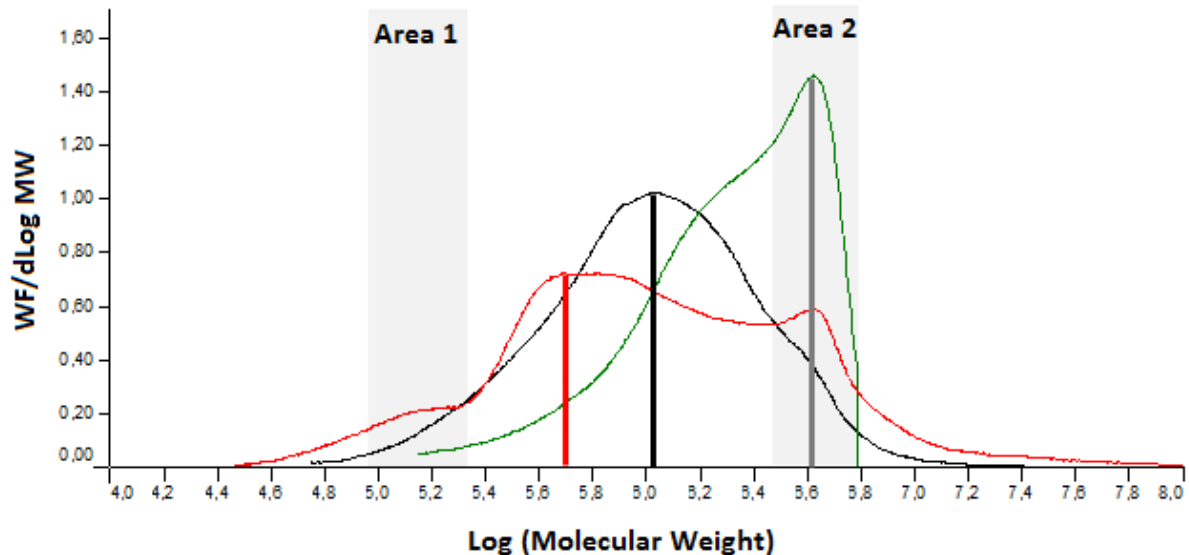


Fig. II-1. Molecular weight density distributions of the cellulosic materials.

X-axis: log(Molecular Weight). Y-axis: normalized weight-fraction. WF: Weight Fraction – MW: Molecular Weight. Red: tissue – Black: paper – Green: cotton. Vertical lines: Log(molecular weight) of the highest peak for each substrate (see Materials and Methods). Area 1: hemicellulose distribution peak (for tissue). Area 2: peaks corresponding to high molecular weight polymers and/or aggregates.

Colonization and metabolic patterns

To compare the behavior of - and the metabolic patterns associated to the three distinct substrates during their bioconversion by the mesophilic cellulolytic bacterium *C. cellulolyticum*, each substrate was used separately as the sole carbon source in triplicate anaerobic pure strain cell cultures (see Materials and Methods). Specific patterns were observed, especially for the tissue incubations.

The colonization dynamics of the cellulosic substrates by the bacterial cells was qualitatively followed over the course of the degradation by wet mount microscopy observations; scanning electron microscopic (SEM) observations on fixed samples were also performed after 46 hours of incubation (see Materials and Methods, Fig. II-2). For none of the tested substrate were protuberances observed on *C. cellulolyticum* cell surface (Fig. II-2 A-C). This is in good agreement with previous studies of the same strain (Ferdinand *et al.*, 2013) and differs from observations of the thermophilic cellulolytic strain *C. thermocellum* grown on filter paper (Bayer and Lamed, 1986). Contrasted colonization dynamics and patterns emerged over time (Fig. II-2). The colonization was the fastest for tissue. It seemed to first initiate at localized spots (Fig. II-2 D, after 46 hours of incubation); at a later stage, the bacterial colonies were abundant and regularly distributed over the cellulose fibers (Fig. II-2 G, after 70 hours of

incubation). By comparison, the colonization of cotton surface appeared as slower and less widespread, even after 70 hours of incubation (Fig. II-2 C, F, I). Indeed, a lower number of unevenly distributed cell heaps were observed. Paper was associated to a somewhat intermediate colonization speed and extent (Fig. II-2 B, E, H): if the colonies seemed more abundant for paper than for cotton, a large portion of the substrate surface still appeared as cell-free.

Based on SEM observations (Fig. II-2 A-C), the biomass was not organized in thick biofilms. However, the observations suggested the presence of an extracellular polysaccharide (EPS) matrix for tissue and paper (Fig. II-2 C), with bacterial cells possibly embedded into this matrix (Fig. II-2 B, arrow). For cotton, no EPS matrix was detected and all colonization heaps were composed of a few isolated cells (Fig. II-2 C). Consistent with all the above-presented elements concerning microbial colonization, the cellulose fibers appeared as the most destructured for tissue, followed by paper and then cotton (Fig. II-2 A-C, after 46 hours of incubation). For tissue, some bacterial cells may have been completely covered by cellulose fibers (Fig. II-2 A, arrow), suggesting that the destructuration could have been more advanced than what was suggested by SEM surface visualization; however this remains quite speculative.

All the data were thus consistent with tissue being colonized the fastest and the most efficiently, followed by paper and then cotton.

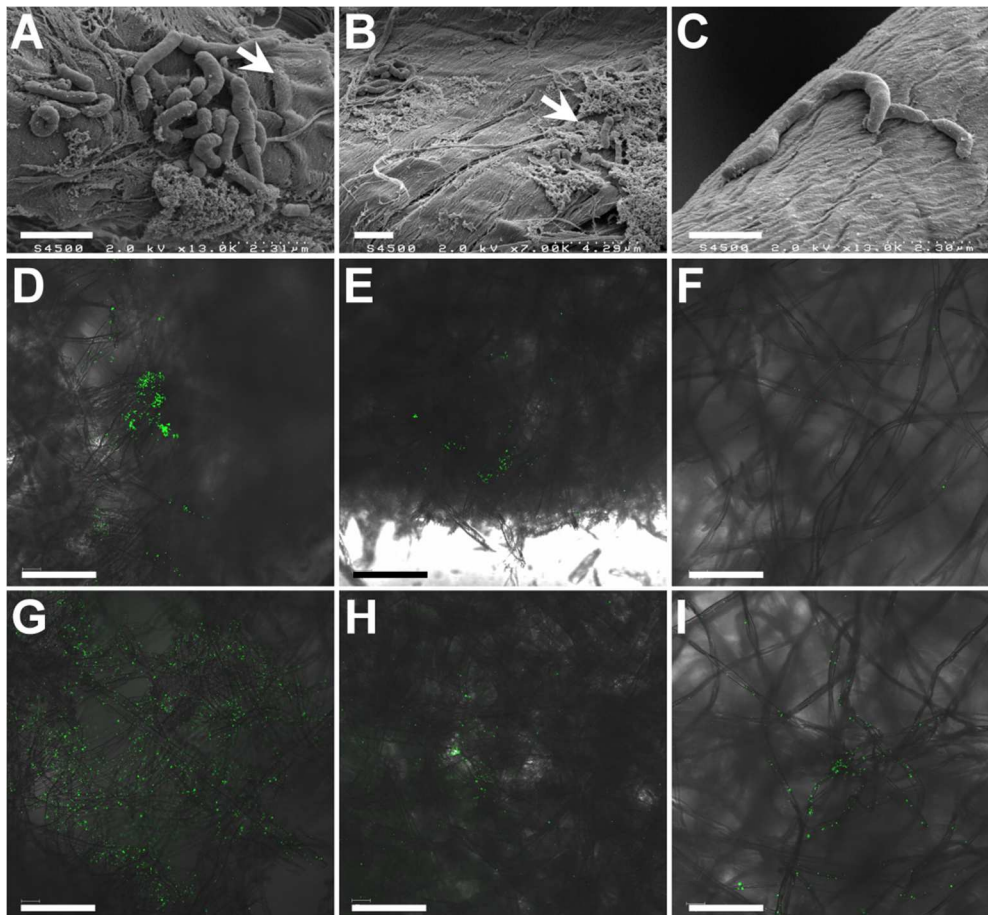


Fig. II-2. Illustrative microscopy images of substrate colonization.

A-C: Scanning Electron Microscopy micrographs obtained after 46 hours of incubation – bars = 2 μ m. D-I: Confocal Laser Scanning Microscopy images acquired on wet mount samples stained with a cellular esterase activity marker (green) – bars = 200 μ m. Left column (A, D, G): tissue incubations – Middle column (B, E, H): paper incubations – Right column (C, F, I): cotton incubations. Second line (D-F): images obtained after 46 hours of incubation. Third row (G-I): images obtained after 70 hours of incubation. Arrows: areas possibly corresponding to bacterial cells embedded in EPS matrix or in cellulose fibers.

To analyze the metabolic patterns, the concentrations of major end-products of cellulose fermentation by *C. cellulolyticum* – acetate, ethanol and lactate – were monitored in the liquid phase, as well as the dissolved organic (DOC) and inorganic (DIC) carbon concentrations and the pH (Fig. II-3 A-F). The cumulated gas production was determined by analyzing the head space at the final time point (Fig. II-3 G). The results primarily showed a distinct metabolic profile for the tissue incubations.

Based on the DOC analyses (Fig. II-3 A), bioconversion occurred at a faster rate for tissue than for the two other substrates during the 100 first hours of incubation. As previously reported, acetate was the major fermentation end-product (Desvaux *et al.*, 2000). At the final time point, acetate levels were similar and ethanol levels were at close levels for all substrates. By contrast, more lactate was produced on facial tissue than on the two other substrates. In more details, at early stages, (less than 24 hours

of incubation), the concentrations of acetate, ethanol and lactate accounted for less than or about 50% of the measured DOC. The presence of soluble sugar may have explained these differences. Between 22 and 142 hours of incubation (from day 1 to day 6), the major fermentation end products accounted for 80 to 90% of the measured DOC, and for less than 80% during the latest incubation stages. The ratios of acetate-to-lactate concentrations were around 5 for the tissue and 10 for the paper and the cotton, all over the incubation time. The ratios of ethanol-to-acetate increased for all the substrates during the early stages of the incubation (around 0.15) and were stable around 0.3 for cotton and paper after 22 hours of incubation (day 1), stable around 0.45 for the tissue after 70 hours of incubation (day 3).

For *C. thermocellum* incubations, it has previously been shown that a significant part of soluble carbon is in the form of soluble sugars resulting from hydrolysis, extracellular pyruvate and amino acids (Ellis *et al.*, 2011). For *C. cellulolyticum*, extracellular pyruvate excretion has already been measured, in particular during growth with cellulose concentrations over 6.7g/L in batch culture with a regulated pH (Desvaux *et al.*, 2000); free amino acids in continuous culture have also been previously detected for *C. cellulolyticum* (Guedon *et al.*, 1999a).

To conclude, the measured end-product concentrations were in good agreement with previous knowledge on *C. cellulolyticum*. A distinct metabolic profile was observed for tissue compared to paper and cotton, the latter being associated with highly similar metabolic profiles, especially during the first 100 hours of incubation. For tissue incubations, higher ethanol-to-acetate and lactate-to-acetate ratios were obtained, suggesting a more important carbon-flux entering the bacterial cells.

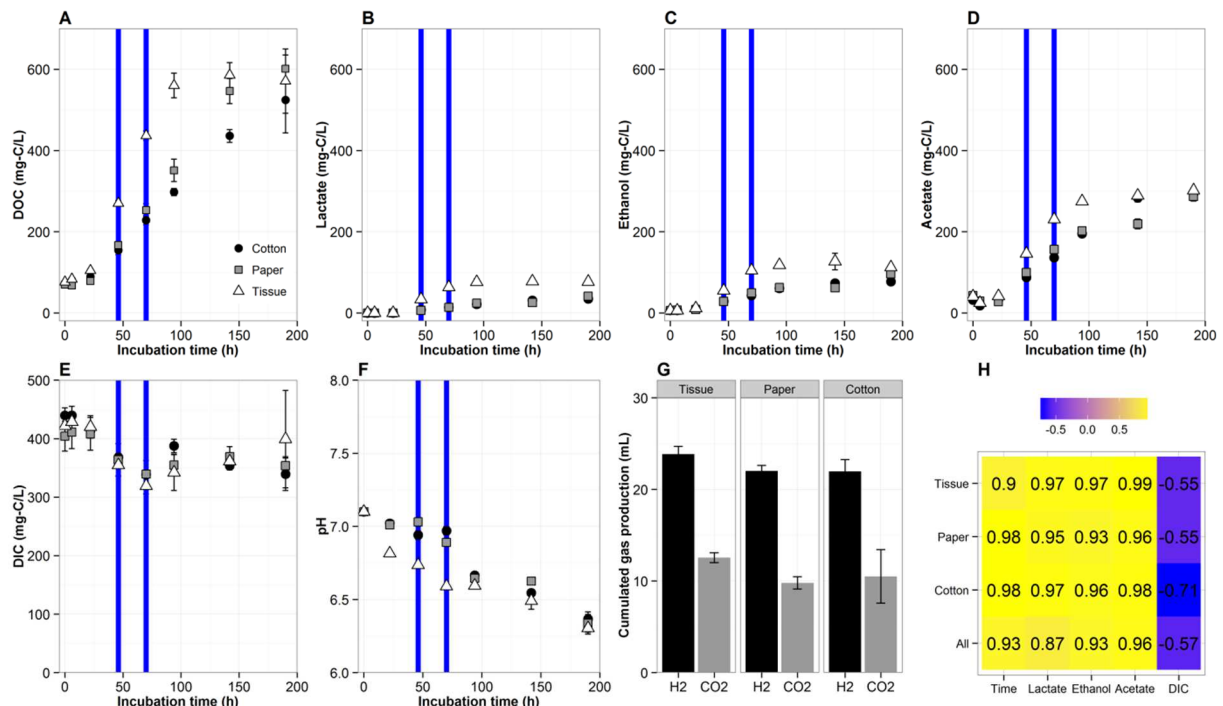


Fig. II-3. Gas and liquid phase analyses of the incubations related to the fermentation metabolic patterns.

A: DOC, Dissolved Organic Carbon – E: DIC, Dissolved Organic Carbon – G: cumulated gas production at the final incubation time point – H: correlation scores of incubation time, lactate, ethanol, acetate and DIC values with the DOC values, calculated for each substrate separately and for all the substrates together (All). A-G: mean \pm SD from the triplicate experimental values are shown. A-F: the legend for all these panels is shown in panel A; vertical lines indicate the time points corresponding to the proteomic datasets (see sections below).

Overview of the proteome datasets

To identify biological functions associated to the distinct patterns observed for colonization and metabolite production, label-free quantitative shotgun proteomic analyses were conducted (see SC approach in Materials and Methods). Analyses were realized during the DOC production phase, after 46 and 70 hours of incubation, on additional triplicate cell cultures incubated simultaneously to those used for colonization and fermentation analyses. Tissue and paper were selected for the proteomic analyses in order to compare the substrate identified as the most reactive, tissue, to a more studied model substrate, paper. The cellular pellets as well as the supernatants, containing excreted proteins, were analyzed separately (gel-like images are shown in Fig. SII-2). The amounts and yields of purified proteins were highly similar among the samples from a same fraction type (pellets *versus* supernatants).

In total, around 25 000 spectra were generated by LC-MS/MS for each sample (Table II-2). Spectra were assigned to peptides by a mass-matching approach using X!Tandem search engine, the X!TandemPipeline (3.3.0) (Craig and Beavis, 2004; Fenyö and Beavis, 2003) and the *C. cellulolyticum*

Uniprot database (see Materials and Methods). Among these spectra around 29% were assigned to peptides for the pellets, and 24% for the supernatants. This statistically significant difference ($p < 0.01$, Mann-Whitney test) concerning spectra assignment could result from distinct levels of post-translational modifications (PTMs) of intracellular and extracellular proteins. In *C. thermocellum*, it was recently shown that almost half of the detected cellulosomal proteins are modified by PTMs such as signal peptide cleavage and amino acid oxidation or methylation (Dykstra *et al.*, 2013). Extracellular proteins could possibly be modified to a greater extent than intracellular proteins, in average, causing lower assignment levels in supernatants: indeed, most of the possible PTMs were not taken into account for protein validation, as it is standardly done to avoid high levels of false positive identifications. The analyses led to an average of 594 and 361 non-redundant identified proteins validated by at least two unique peptides, for the pellets and supernatants respectively. To assess the analysis depth, rarefaction curves were estimated based on the number of unique proteins identified. They showed that almost all proteins identifiable with the present method were successfully retrieved for each sample (Fig. II-4 A), although the number of validated peptides differed according to the fraction type (pellets *versus* supernatants). For each fraction and substrate, proteins identified from all the replicates and from the two time points were pooled and compared (Fig. II-4 B). Most of the proteins (353) were found in all samples, showing that the degradation of both substrates relied on the activation of core functions required for cellulose degradation.

Table II-2. Statistical indicators related to peptide and protein identification.

	Pellets	Supernatants
Used spectra	25,225 ± 355	25,299 ± 253
Assigned spectra	7,332 ± 421	6,047 ± 439
Spectrum assignment rate	29.07% ± 1.55	23.92% ± 1.75
Identified proteins	594 ± 39	361 ± 24
Identified CAZymes	79 ± 3	55 ± 5

For pellets and supernatants data were pooled from all samples for the two conditions, the two time-points and the three replicates. Values ± standard deviation.

Identified unique proteins were classified using the redundant Clusters of Orthologous Groups (COGs) database, which enables a general functional annotation (Fig. II-4 C). Results showed that most of the proteins identified in the pellets (~40%) and the supernatants (~45%) were related to the metabolism category, i.e. metabolic pathways for substrate degradation, energy production and cell maintenance. The large amount of identified proteins not assigned to a COG category (~15%) was expected since for most bacteria, the COG coverage, reflecting protein sequence conservation among prokaryotic lineages, is approximately 80% of the encoded proteins (Tatusov *et al.*, 2003). A similar percentage

level (~10%) corresponded to proteins closely related to database proteins of unknown or poorly predicted function, which is the largest category in the COG database (Tatusov *et al.*, 2000). Among the 63 identified proteins assigned to the COG category “Poorly characterized”, only 17 are actually proteins of unknown function in *C. cellulolyticum*. According to the CAZy database (Cantarel *et al.*, 2009), CAZymes were found in the pellets and in the supernatants (Table II-2), as was expected since cellulosomes are generally found attached to the cell surface and cellulosomal components are also excreted (Cantarel *et al.*, 2009). A detailed list of the identified CAZymes is presented in Table SII-1.

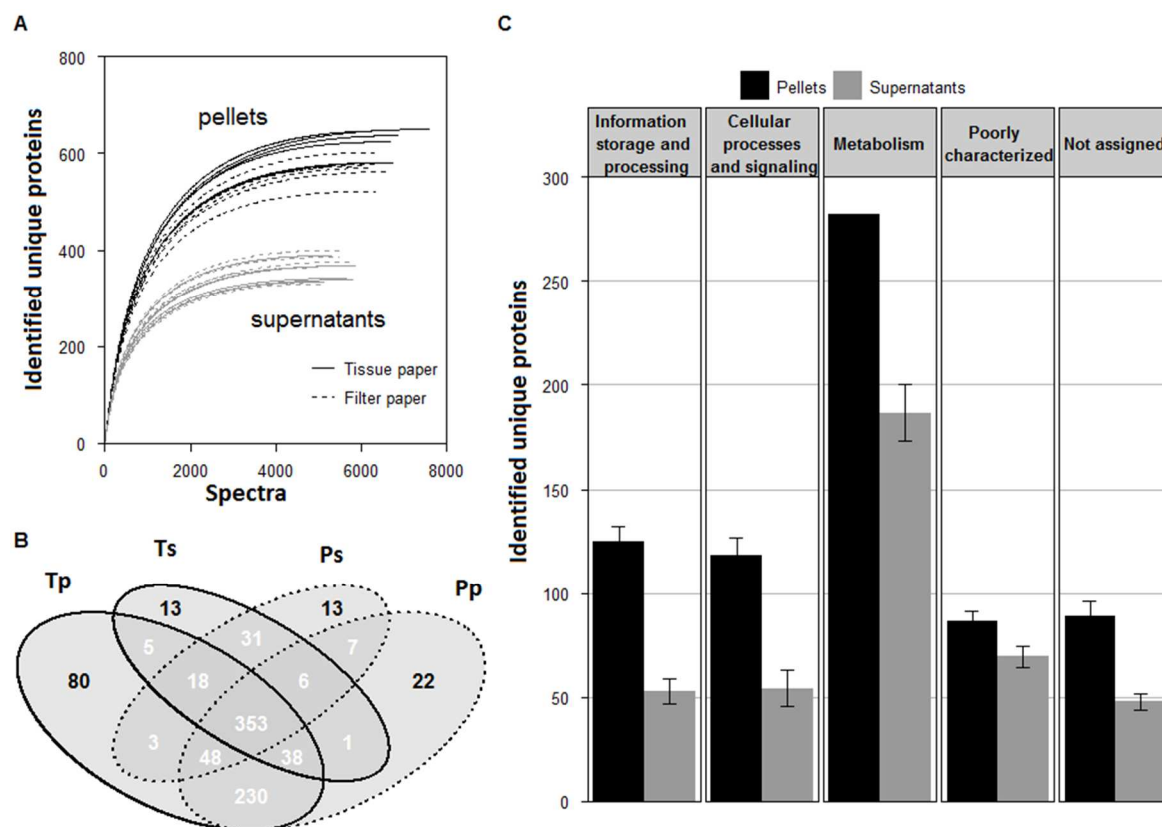


Fig. II-4. Overview of the proteomic results.

A: rarefaction curves based on the number of identified unique proteins (Y axis) in function of the number of used spectra (X axis). B: Venn diagram of the identified unique proteins for each type of substrate and fraction (Tp: Tissue, pellets – Ts: Tissue, supernatants – Pp: Paper, pellets – Ps: Paper, supernatants). C: classification of the identified unique proteins according to COG categories.

Label-free protein quantification

To compare protein levels for the tissue and paper incubations, two label-free quantitative data analyses were implemented independently. These approaches were either based on signal integration (area under curve, AUC) from precursor ion spectra or on spectral counting (SC) of the peptide fragmentation spectra, using PepC software (Heinecke *et al.*, 2010) (see Materials and Methods). These two types of methods are complementary and their benefits and limitations have recently been

reviewed (Neilson *et al.*, 2011). In the present work, the AUC data served as a reference and most statistical analyses were performed on the corresponding datasets (see Materials and Methods). The SC approach, on the other hand, was mainly used to consolidate the results by verifying the overall consistency (Fig. SII-3). A table of all the identified proteins and associated values for both approaches (quantification, statistical tests, etc) is available in the Supplementary Materials. For the analyses presented below and in the next Result sections, only the AUC datasets were considered.

To identify proteins with significantly different levels based on the Protein Abundance Indexes (PAI), the datasets were analyzed with Tobit models after their \log_{10} -transformation and quantile normalization. The regression models were fitted for each identified protein in each type of fraction (pellets *versus* supernatants) and for each substrate (see Materials and Methods). Tobit models are compatible with censored data and are thus suitable for mass spectrometry dataset analyses. The models included the effects of the substrate, the incubation time and the interaction of substrate and time. In total, 187 proteins showed at least one significant effect for one of the fractions, corresponding to 94 and 58 proteins for the pellets and supernatants respectively.

Gene set enrichment analyses of functional categories

To identify global molecular processes affected by substrate or interaction effects, gene set enrichment analyses (GSEA) of various functional categories were performed (Subramanian *et al.*, 2005, see Materials and Methods). GSEA is a procedure to test whether the elements of some predefined biological groups are enriched towards the top or bottom of a ranked list. COG categories, BioCyc metabolic pathways and CAZymes were considered as functional categories. Different scores were used to generate the ranked lists of identified proteins. First, statistical values associated to the substrate effect and to the interaction between substrate and time effect were used (Tobit regression models, see above and Materials and Methods). In addition, since substrate degradation and metabolite production were correlated to DOC concentrations (Fig. II-3 H), a correlation coefficient between protein levels and DOC values obtained for each replicate was also used for ranking (see Materials and Methods). Results are presented in Table II-3 for the main functional ontologies and pathways, comprising at least nine identified proteins. The signs associated to significant enrichments are conventional, as detailed in the table footnotes.

The results showed that statistically significant enrichments were more numerous for the pellets than for the supernatants. In more details, the category “energy production and conversion” was enriched with proteins positively correlated with DOC concentrations and with positive substrate and interaction effects for the pellets. Besides, the glycolysis and gluconeogenesis pathways were enriched with proteins positively correlated with DOC concentrations and with a positive interaction effect.

However the heterolactic fermentation pathway was only enriched with protein with a positive substrate effect. Numerous significant enrichments were detected for the CAZy category, and, surprisingly, this category was enriched with proteins negatively correlated with the DOC concentrations for the pellets only (no significant DOC-related enrichment for the supernatants). Regarding substrate and interaction effects for the CAZy category, contrasted trends were observed: negative-sign enrichments were obtained for the pellets (tissue protein levels < paper protein levels, and tissue protein levels increase less or decrease more than paper protein levels over time) whereas positive-sign enrichments were obtained for the supernatants (tissue protein levels > paper proteins levels, and tissue protein levels increase less or decrease more than paper protein levels over time).

The tRNA charging pathway appeared as enriched only in supernatants, with proteins positively correlated to the DOC concentration and with a positive substrate effect. These proteins, such as tRNA ligases, are not expected to be excreted in the extracellular medium and their presence in supernatants could probably results from cell lysis. Besides, since p-values were not adjusted to gain maximal information about trends for each category, false positives could have emerged due to multiple tests and the tRNA charging trends are likely to be false positives.

Table II-3. Gene Set Enrichment Analyses (GSEA) for major functional categories.

Functional categories (gene sets)	Correlation TOC*		Substrate**		Interaction***	
	Pellet	Supernatant	Pellet	Supernatant	Pellet	Supernatant
Amino acid transport and metabolism (COG)					+	
Energy production and conversion (COG)	+		+		+	-
tRNA charging pathway		(+)		(+)		
gluconeogenesis I pathway	+				+	
heterolactic fermentation pathway			+			
glycolysis III (glucokinase) pathway	+				+	
CAZy	-		-	+	-	+

* +: correlated positively with DOC values, -: correlated negatively with DOC values.

** +: positive substrate effect (tissue > paper), -: negative substrate effect (tissue < paper).

*** +: positive interaction effect (tissue increases more or decreases less than paper), -: negative interaction effect (tissue increases less or decreases more than paper).

(+) probable false positive effect (see result section Gene set enrichment analyses of functional categories)

The activity of fermentation pathways was higher during growth on tissue, as shown by the dynamics of the chemical parameters (see Result section Colonization and metabolic patterns). Based on the synthetic view gained from the enrichment analyses, this increased activity was most probably partly

linked to overall higher protein levels in the corresponding pathways. Besides, the opposite trends observed for the CAZy category could reflect different enzymatic strategies to degrade tissue and paper. To achieve a more detailed understanding of these aspects, examinations of individual proteins belonging to functional groups of interest were performed.

Carbohydrate-active enzymes (CAZymes)

The CAZymes showing significant effects, except for glycosyltransferases proteins (which are not involved in cellulose degradation) are presented in Fig. II-5 with details on their nature. For clarity, enzymes were named after their gene locus name in the following text. It should first be noted that several major cellulosomal proteins did not show any significant effects and are thus not shown in the table (see Table SII-1): the scaffolding protein CipC (Ccel_0728), another cohesin module bearing protein OrfX (Ccel_0733) as well as all proteins encoded by the xyl-doc cluster (encoding proteins more specifically oriented towards hemicellulolysis). Among the 27 CAZymes with a significant effect (Fig. II-5), 20 are cellulosomal components, since they harbor a dockerin module, and they include 4 proteins encoded by the cip-cel cluster (Ccel_0732, Ccel_0734, Ccel_0735, Ccel_0740). Results showed an overall clear pattern since the significant proteins were not the same for pellets and for supernatants, except for Ccel_1648.

Concerning pellets, a large fraction of the significant CAZymes (12) showed a negative substrate effect and they were thus detected at higher levels for the pellets from paper incubations than for those from tissue incubations. These 12 proteins are cellulosomal components and include all the identified significant *cip-cel* cluster proteins. Six of them harbor a GH9 catalytic module often associated with CBM3 modules (Ccel_0734, Ccel_0735, Ccel_0753, Ccel_1249, Ccel_1648); four other enzymes bears other catalytic modules – GH5 (Ccel_0740), GH10 (Ccel_0931), CE2 (Ccel_2162) and PL1 (Ccel_2243); the remaining two are cellulosomal proteins with a putative esterase activity (Ccel_1060) and with unknown function (Ccel_1655). Endoglucanase F (Ccel_0729) was only slightly overrepresented (q -value = 0.015) for the pellets of paper incubations. Only two proteins were more concentrated for the pellets of tissue incubations, a GH3 (Ccel_1139) and a GH5 containing proteins (Ccel_2337).

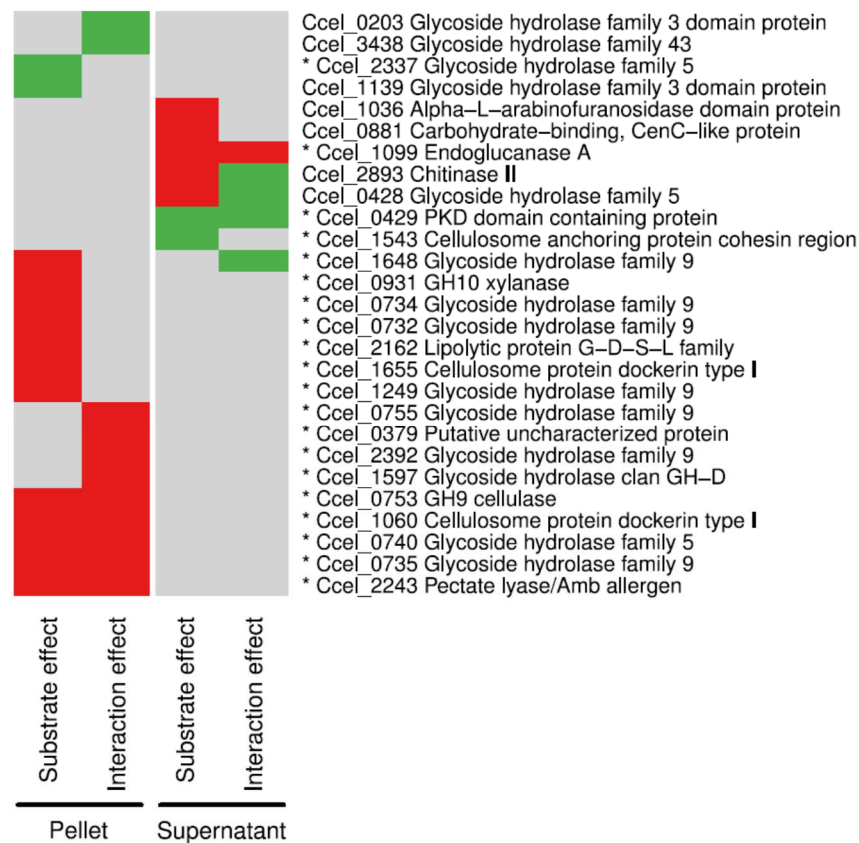


Fig. II-5. Heat map of CAZymes showing significant substrate or substrate by time interaction effects in the pellet or supernatant fraction.

Green: positive effect (tissue > paper for substrate effect, tissue increases more or decreases less than paper for substrate by time interaction effect). Red: negative effect (tissue < paper for substrate effect, tissue increases less or decreases more than paper for substrate by time interaction effects). Grey: no significant effect. * cellulosomal components.

Concerning supernatants, the trend was less clear since the identified proteins showed various behaviors. In more details, proteins more concentrated for the paper incubations were two GH5 (Ccel_0428, Ccel_1099) including the well-known Endoglucanase A, a GH18 (Ccel_2893), a GH51 (Ccel_1036) and a CBM16 (Ccel_0881) containing proteins. Proteins more concentrated for the tissue incubation were a GH44 (Ccel_0429) containing protein and the cohesin bearing protein Ccel_1543. Interestingly, only three cellulosomal proteins were found with significant substrate effect in supernatants and most of the CAZymes showing significant substrate effect also showed interaction effects. Finally, still for the supernatants, significant interaction effects were mainly positive, corresponding to protein levels that increase more or decrease less in tissue than in paper incubations between time points: it indicates that these protein levels were not modulated in a similar way over time in the two types of incubations (tissue *versus* paper).

The CAZyme analysis suggests that the cellulosomes had distinct average compositions in tissue and paper incubations. It could reflect different adaptations of the cellulolysis machinery to the precise substrate (Blouzard *et al.*, 2010). The observed differences are subtle and concern a variety of enzymes, highlighting the complexity of the cellulolysis machinery and of its regulation.

Central carbon metabolism

Based on the enrichment analyses (see Result section Gene set enrichment analyses of functional categories), several metabolic pathways of specific interest regarding cellulose degradation were enriched with proteins showing significant effects among DOC correlation, substrate or substrate by time interaction. Glycolysis/gluconeogenesis and heterolactic fermentation pathways were in particular retrieved. These pathways, previously described for *C. cellulolyticum* (Desvaux, 2005) and including the Embden-Meyerhof-Parnas (EMP) pathway (which generates ATP, NADH, and several biosynthetic precursors such as pyruvate from glucose), were used to map the proteins with significant effects (Fig. II-6).

All 23 enzymatic activities involved in these catabolic reactions were successfully retrieved. Among them, significant effects concerned almost exclusively enzymes overrepresented in the tissue incubations (positive effects) and located upstream the pyruvate node, when adjusting the statistical test for multiple comparisons. Several proteins involved in the uptake and possibly in the sensing of hydrolysis products were also differentially expressed under tissue and paper conditions. When considering the proteins significant based on the p-value only (not significant based on multiple comparisons), 8 additional proteins were retrieved, including 4 proteins located downstream of the pyruvate node and distributed among the lactate, acetate and ethanol fermentation pathways: L-lactate dehydrogenase (LDH) (EC 1.1.1.27), acetate kinase (EC 2.7.2.1), acetaldehyde dehydrogenase (EC 1.2.1.10) and alcohol dehydrogenase (EC 1.1.1.1).

More elements concerning the significant proteins identified with the most stringent test (multiple comparisons) are detailed below. At least one significant effect, substrate or interaction, was observed for 9 enzymes. These effects were present in the pellets only, as expected, except for alcohol dehydrogenase (EC 1.1.1.1) and phosphopyruvate hydratase (EC 4.2.1.11) that were also associated to significant effects in the supernatants. Among enzymes showing a significant substrate effects all, except for phosphopyruvate hydratase (EC 4.2.1.11), had a positive effect indicating higher protein levels in the tissue incubations compared to the paper ones. These enzymes were all distributed in the EMP, such as glucose-6-phosphate isomerase (EC 5.3.1.9), 6-phosphofructokinase (EC 2.7.1.11), glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) and phosphoglycerate mutase (EC 5.4.2.1). Interestingly, the above-mentioned phosphopyruvate hydratase (EC 4.2.1.11), associated to a negative

substrate effect for both pellets and supernatants, also was associated to a positive interaction effect when considering only the p-value. Besides, two proteins involved in xylose fermentation, xylose isomerase (EC 5.3.1.5) and xylulokinase (EC 2.7.1.17) were retrieved and the xylose isomerase showed a positive interaction effect. Xylose is a classic constituent of hemicellulose.

Upstream of the fermentation pathways shown in Fig. II-6, proteins involved in the binding and transport of cellulose hydrolysis products were also retrieved (see Supplementary Materials). Among them, only one ABC transporter showed significant substrate effect in pellets and was more concentrated under growth with paper. This protein (Ccel_0967) is a type 2 transporter with unknown substrate specificity. However many proteins related to solute-binding proteins (SBP) showed a significant substrate effect in the pellets or in the supernatants. All these SBP belong to the family 1 with the specificity for malto-oligosaccharides, multiple sugars, glycerol phosphate or iron, except for Ccel_1768 which belongs to the family 5 whose specificity is for peptides or nickel (Tam and Saier, 1993). Most of these SBP showed a negative substrate effect and were thus more concentrated in the paper incubations. None of these SBP are those regulated by the two component system, as shown by Xu *et al.* (2013).

This body of data suggests that many proteins involved in sugar uptake and fermentation pathways were present at higher levels in the tissue incubations and this, among other factors, could have contributed to the increased activation of the central metabolic pathways. Other additional factors could typically have been direct regulatory effects of increased carbon intracellular influx and metabolite concentrations on protein activity, through allosteric regulation (Crow and Pritchard, 1977; Veith *et al.*, 2013).

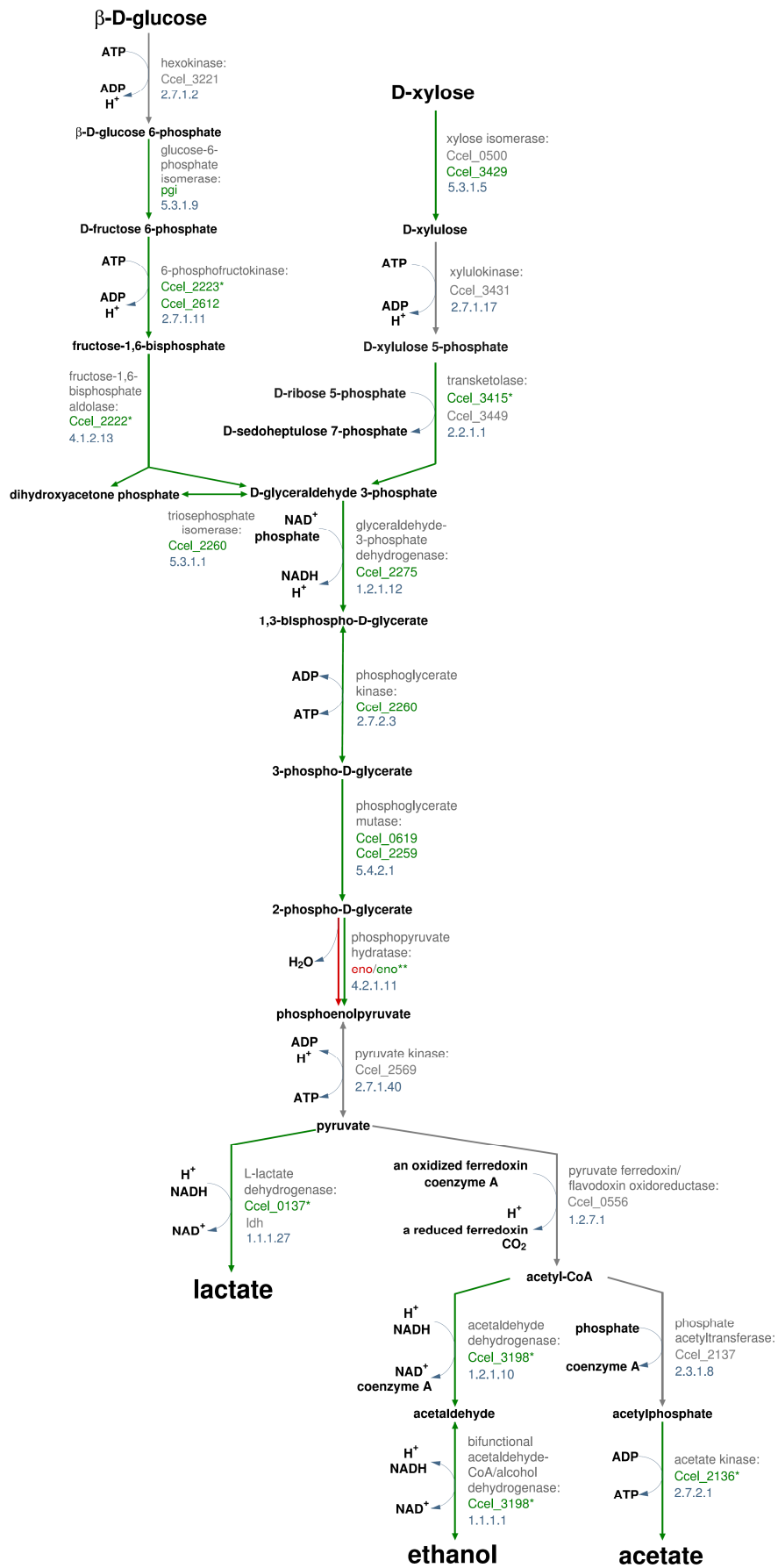


Fig. II-6. Individual proteins with significant substrate or substrate-time interaction effects mapped over *C. cellulolyticum* central carbon metabolism and fermentation pathways adapted from the Biocyc website.

Green: positive effect (tissue > paper for substrate effect, tissue increases more or decreases less than paper for substrate by time interaction effect). Red: negative effect (tissue < paper for substrate effect, tissue increases less or decreases more than paper for substrate by time interaction effects). Grey: protein names and properties, including the EC number * Effect not significant when adjusting for multiple comparisons but still showing the overall trend as p-values ≤ 0.05 . ** Significant negative substrate effect (q-value ≤ 0.01) and positive interaction effect (q-value > 0.01 but p-value ≤ 0.05).

Bacterial adhesion and chemotaxis

C. cellulolyticum tightly adheres to its substrate but mechanisms participating in cell adhesion are not clearly identified to date. It has been shown that the scaffoldin protein *CipC* (Ccel_0728) is only partly responsible for adhesion (Ferdinand *et al.*, 2013). Recently, Imam *et al.*, showed that *C. cellulolyticum* genome encodes for multiple type IV pili, including one *pil* and three *tad* operons (Imam *et al.*, 2011). However, in the present study only three ATPases related to *tadA* (Ccel_0114 and Ccel_0503) and PilT (Ccel_2413), corresponding to distinct type IV pili systems, were retrieved and none of these proteins showed significant effects (see Supplementary Materials). Among proteins participating in chemotaxis, including those involved in flagellar assembly, only one methyl-accepting protein (MCP) Ccel_3326 had a significant substrate effect and was more concentrated in the paper than in the tissue incubations (see Supplementary Materials).

Among other proteins having a significant substrate effect, three are possible candidates for participation in substrate sensing. These proteins of unknown function contain S-layer homology domains, such as Ccel_0274 and Ccel_2393, indicating possible cell-bound proteins, or fibronectin type III domains, such as Ccel_0648, which are sometimes found in CAZymes and can also be linked to cell adhesion (Kataeva *et al.*, 2002).

Discussion

Over the course of the experiment, growth on the three cellulosic substrates led to the same bioconversion yield since final carbon levels in the liquid and gas phases were similar for all incubations. From this result, it seems that all three substrates have the same biodegradability in these conditions. However, it has already been shown that *C. cellulolyticum* has a limited growth in batch culture and that inhibition occurs by cell self-intoxication owing to the accumulation of currently unknown intracellular intermediates (Guedon *et al.*, 1999a). Consequently, this result does not appear as definitive since it could have been different in conditions favoring growth, typically in continuous culture with pH regulation (M Desvaux *et al.*, 2001b).

However, the bioconversion rates were different among the three substrates since the DOC production rate was 2.2 and 1.3 fold higher for tissue and paper incubations respectively, compared to cotton incubations. This difference is likely linked to different cellular activity levels rather than to different cell concentrations since biomass yields were probably similar between tissue and paper incubations, as shown by cellular protein concentration measurements. In relation with this observation, it appeared that more final fermentation products were released during the growth on facial tissue. Interestingly, different ratios between the fermentation product levels were observed. These ratios highlight the catabolism partition between ATP production and co-enzymes reoxydation. The ethanol-to-acetate ratio was higher during growth on facial tissue but always lower than 1, suggesting an optimized growth (Desvaux *et al.*, 2000). The lactate-to-acetate ratio was much lower during growth on facial tissue, suggesting that more carbon flowing from glycolysis is routed through lactate formation. Regarding enzyme concentrations, it appeared that only enzymes upstream the pyruvate node, i.e. the EMP pathway, were overrepresented during growth on tissue when using a false discovery rate of 1 %. However, several enzymes downstream this node, involved in lactate, acetate and ethanol production, also showed a trend towards over-abundance when not adjusting for multiple comparisons (p -values ≤ 0.05). In all cases, protein levels cannot fully account for enhanced enzyme activities since most of the enzymes involved in these pathways are subjected to allosteric regulation. Typically, lactate-dehydrogenase activity is increased by the presence of fructose-1,6-bisphosphate (Crow and Pritchard, 1977). For *C. cellulolyticum*, it has been shown that partition of the carbon flux is greatly influenced by the entering carbon flow and the pH (Desvaux *et al.*, 2000; M Desvaux *et al.*, 2001b). Particularly, pyruvate ferredoxin oxidoreductase is known to be very sensitive: this enzyme cannot handle high carbon flow from glycolysis and its activity greatly decreases with pH (M Desvaux *et al.*, 2001b). In our conditions, the pH was not controlled such that the faster pH decrease (pH 6.75 and 6.6 for 46 and 70h respectively) during the growth on tissue could have contributed to the different metabolic patterns observed. However, ratios and final fermentation products concentrations were still in the range of those reached during growth on cellulose (Desvaux *et al.*, 2000) compared to growth on cellobiose (Guedon *et al.*, 1999a). Altogether, these results suggest that growth on tissue led to higher carbon flux entering the cell following substrate deconstruction by CAZymes and it suggests a higher biodegradability for the tissue compared to the paper and cotton.

Interestingly, colonization patterns also appeared as different since a more widespread colonization was observed on tissue fibers compared to the small-size heaps observed on cotton fibers. This result suggests that sites for adhesion and degradation could be more accessible and/or numerous on tissue than on paper and cotton. This higher accessibility and site number could therefore explain the enhanced biodegradability. This hypothesis is also consistent with the observed metabolic patterns

since it has already been shown that increasing cellulose concentration up to 6.7 g/L leads to an increased ethanol-to-acetate ratio (Desvaux *et al.*, 2000). As cellulose is an insoluble and recalcitrant substrate, the concentration of accessible sites is likely to be a more significant factor than the cellulose mass concentration. Still concerning colonization, proteins involved in cell motility were not differentially expressed suggesting similar ratios between adhered and free cells between facial tissue and filter paper. In contrast to *Ruminococcus albus*, implication of type IV pili in cell adhesion was not shown in our conditions (Rakotoarivonina *et al.*, 2002). Differences in structure and composition among the three substrates principally consist in the cellulose crystallinity and DP. However, it appeared that cellulose crystallinity, generally thought to be the principal factor underlying cellulose recalcitrance, does not fully explain the differences in biodegradability, suggesting an important role for the DP (Mansfield *et al.*, 1999). The presence of hemicellulose may also have an impact on the biodegradability properties since it is hydrophilic and may enhance cellulose accessibility through swelling and hydration.

In cellulolytic bacteria, CAZymes have a central role encompassing cellulose deconstruction and metabolism. CAZymes are known to be highly modulated by the substrate nature, from single monosaccharide to complex lignocellulosic materials (Blouzard *et al.*, 2010; Xu *et al.*, 2013) and here it has appeared that subtle differences in substrate structure and composition also lead to CAZymes modulation. First, higher levels of some cellosomal components during growth on paper compared to tissue were observed here. In the model proposed by Xu *et al.*, cellosomal genes are activated, or not repressed, when intracellular levels of glycolytic intermediates are low (Xu *et al.*, 2013). Since less carbon flows through glycolysis during growth on paper compared to tissue, results obtained here are consistent with this model. Secondly, paper degradation involved more proteins of the GH9 family, whose members are endoglucanases, which is consistent with the presence of longer cellulose chains in this material. By contrast, tissue degradation involved the higher level of a beta-glucosidase, suggesting, although this result is not as straightforward, more readily available cellodextrines.

This study is the first proteome-wide analysis of the cellulolytic bacterium *C. cellulolyticum*, an important model in the field of bioethanol production from cellulosic materials. The integrated approach showed that the use of cellulosic manufactured materials with subtle composition differences resulted similar bioconversion yields but was sufficient to generate distinct colonization and metabolic patterns as well as degradation speeds. Facial tissue appeared as the more reactive substrate and yielded a high ethanol-to-acetate ratio, compared to filter paper and cotton. Advanced comparative statistical analyses of label-free shotgun proteome datasets enabled to identify 187 proteins associated to a significantly distinct behavior in tissue compared to paper, among more than 1200 proteins identified in total. These results suggest the higher level of proteins from the

carbohydrate central metabolic pathway upstream the pyruvate node, during tissue degradation. They also showed distinct average compositions of the cellulosomes according to the substrate material. These results constitute an important step towards the understanding of cellulolysis at the system level and nicely complement the data available on *C. cellulolyticum* metabolism (Desvaux, 2005). At a more applied level, they suggest the interesting potential of the lignocellulosic fraction of MSW for biorefinery purposes.

Materials and methods

Substrate materials and characterizations

Three cellulosic materials, cotton pads (cotton), Whatman No 1 filter paper (paper) and facial tissue (tissue) were used as sole carbon sources. The moisture and ash content of these substrates were determined by oven-drying at 105°C for 24h and at 550°C for 1h following instructions described in EN ISO 12879 and EN ISO 12880. Analyzes were performed using an Elementar Vario EL III analyzer. The samples were placed in tin capsules and injected in a dry combustion chamber at 950°C. Nitrogen oxide was then reduced to N₂. Resulting CO₂ and N₂ were finally separately quantified with a catharometer.

To determine the molecular weight density distributions, cellulosic substrates were dissolved and derivated by tri-carbanilation using N,N-dimethylacetamide (DMAc) (Henniges *et al.*, 2007). The obtained solutions were analyzed in a size-exclusion chromatograph (Viscotek TDA-302 apparatus) equipped with 3 Varian PLGel Mixed B columns (7.8×300) with a guard column in series. The eluent was THF at 1 mL/min, 35°C. Samples were injected at ~1 mg CTC/mL. The injection volume was 100 µL. The coupled detection was UV at 260 nm, RALS/LALS/RI at 670 nm (laser 3 mW) and viscometer detector, all at 35°C in the same block. Refractometry was used as concentration detector and UV as control. The dn/dC was taken at a fixed value of 0.165 in good accordance with literature data and measurements of sample recovery through the column. Data were treated by the OmniSec™ (4.5.6. version) program (Viscotek Co.). The most influencing parameters are (1) choice for peak baselines and integration limits (based on RI and UV signals), and (2) choice for column curve fit (log MW versus elution volume). Both significantly influence the shape of molecular weight distributions, average molecular weight and polydispersity calculations

Diffraction diagrams were monitored by recording X-ray diffraction diagrams every 10 minutes on a Bruker D8 Discover diffractometer. Cu Kα1 radiation (Cu Kα1 = 1.5405 Å), produced in a sealed tube at 40 kV and 40 mA was selected and parallelized using a Göbel mirror parallel optics system and collimated to produce a 500 µm beam diameter. Crystallinity index was calculated using the method developed by Wakelin *et al.* (1959).

For mid-infrared spectroscopy, 2 mg of each material was mixed with 120 mg of spectroscopically pure potassium bromide and measurements were realized as described in Devaux *et al.* (Devaux *et al.*, 2005). The spectral region between 2000 and 700 cm^{-1} was treated using the OPUS software (v 2.06) as indicated before analysis.

Strain and culture conditions

Clostridium cellulolyticum DSM 5812 was grown anaerobically at 37°C as described in Stams *et al.* (1993) (initial pH 7.1) in minimal medium containing, per liter: Na_2HPO_4 , 0.4 g; KH_2PO_4 , 0.4 g; NH_4Cl , 0.3 g; NaCl , 0.3 g; MgCl_2 , 0.1 g; CaCl_2 , 0.1 g; NaHCO_3 , 4.0 g; $\text{Na}_2\text{S}\cdot 9\text{H}_2\text{O}$, 0.2 g. The medium was supplemented with 0.2 mL/L of vitamin solution and 1 mL/L of acid and alkaline trace element solutions (each) (Stams *et al.*, 1993). 5 mg/L resazurin were added to the medium as a redox indicator. Inoculation was realized with 10% (v/v) of a pre-adapted culture grown on 1g/L Sigmacell microcrystalline cellulose. For proteomic and chemical analyses, reactors consisted in 125 mL flasks with 50 mL working volume and the substrates were added to a final concentration of 2.5 g/L. Triplicate flasks containing either tissue, or paper or cotton were dedicated to physico-chemical monitoring of the degradation dynamics. For tissue and paper, six additional replicate flasks were operated and sacrificed for proteomic analyses during the incubation. For imaging, reactors consisted in 96-well microplates incubated in sealed anaerobic bags with 300 μL working volume in each well, using the same minimal medium. For each different substrate, chads were punched and added to 5 wells, resulting in a final substrate concentration of approximately 5 g/L.

Imaging

At each sampling time point, one microplate was sacrificed and the samples were prepared for imaging. The liquid phase was removed from each well and cellulose fibers were washed with 1X phosphate buffer saline (PBS). For live cell imaging, 4 wells were prepared for each substrate as follows: cells were stained with 0.1% ChemChrome V8 (Chemunex), a cellular esterase activity marker, in 1X PBS for 10 min at 37°C. Observations were realized directly after staining using a confocal laser scanning microscope (LSM 510 Zeiss). For scanning electron microscopy, 1 well was prepared for each substrate as follows: samples were fixed in a solution containing 2.5% glutaraldehyde and 0.1 M sodium cacodylate (pH 7.4) for 16h and stored at 4°C until further analyses. Before their observation, samples were prepared as described in (Bridier *et al.*, 2013). The observations were performed by the INRA platform MIMA2.

Physico-chemical analyses of the incubation samples

At each sampling time, fresh samples (2.2 mL) were recovered from the 125 mL flasks: 0.2 mL were used to measure the pH immediately after sampling and the remaining 2 mL were centrifuged at 10 000g for 10 min at 4°C. The obtained pellets and supernatants were stored separately at -80°C. Substrate degradation was estimated by measuring the DOC and DIC concentrations in the supernatants using a BIORITECH 700 analyzer. Volatile fatty acid concentrations (including lactate and acetate) were measured using a DX 120 Ion Chromatograph (Dionex) with an IonPac ICE-AS1 column. Ethanol concentrations were measured by gas chromatography/mass spectrometry (GC/MS) as described in Richard *et al.* (in preparation). Biogas composition was analyzed at the final time point with a gas chromatograph as described in Qu *et al.* (2009b).

Total proteome and exoproteome extraction and MS/MS analyses

During the DOC production phase, after 46 hours and 70 hours of incubation, 3 flasks for each condition were sacrificed and sampled for proteomic analyses. The total volume, except 2 ml used for chemical analyses, was supplemented with 1 mM PMSF (Phenylmethylsulfonyl Fluoride) to inhibit protease activity and centrifuged at 6 000g for 20 min. The collected supernatant was further centrifuged at 13 000g for 15 min to remove remaining substrate debris and cells. Pellets from the first centrifugation round and supernatants from the second one were frozen in liquid nitrogen immediately after the phase separation, and stored at -80°C. Proteins from the pellets were extracted as described in Hanreich *et al.* (Hanreich *et al.*, 2012) with minor modifications. Briefly, cell disruption was performed on the pellets resuspended in 1X PBS, by a bead-beating step using 0.1 mm silica and glass beads, followed by an ultrasonication step. Proteins were extracted using liquid phenol and precipitated with ammonium acetate in methanol. After several washing steps, the protein pellets were resuspended in buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS) as described in Lü *et al.* (2013) and then stored at -80°C. Culture supernatants were filtered through 0.22 µm PVDF membrane filters and proteins were precipitated and resuspended as described above. A total of 24 protein samples were obtained (2 different substrates, 3 replicates, 2 different time points, pellets and supernatants). Protein concentrations were measured using an Agilent 2100 Bioanalyzer with a High Sensitivity Protein 250 kit following the manufacturer's instructions. For each sample, 5 µg of purified proteins were fixed in SDS-PAGE gel (NuPAGE Bis-Tris Mini Gels, Invitrogen). After a short-duration migration, gel fragments (~5 mm) containing the proteins were excised. In-gel trypsin digestion was performed on each gel fraction and the obtained liquid fractions with peptide mixtures were analyzed by shotgun liquid chromatography-tandem mass spectrometry (LC-MS/MS). LC-MS/MS analyses were performed with an Ultimate 3000 LC system (Dionex) connected by a nanoelectrospray interface to a linear ion trap

mass spectrometer coupled to an Orbitrap detector (LTQ-Orbitrap, Thermo Fisher; PAPPISO proteomic platform, INRA, Jouy-en-Josas). Four μl of tryptic peptide mixtures from each sample were loaded at a flow rate of 20 $\mu\text{L}/\text{min}$ onto a precolumn (Pepmap C18, 0.3 \times 5 mm, 100 \AA , 5 μm ; Dionex). After 4 minutes, the precolumn was connected to a separating Pepmap C18 nanocolumn (0.075 \times 15 cm, 100 \AA , 3 μm ; Dionex) and the gradient was set to 300 nL/min. All peptides were separated in the nanocolumn using modified buffer elut-B with a linear gradient of acetonitrile from 2% to 36%, for 18 minutes. The eluting buffers were buffer elut-A: 0.1% formic acid, 2% acetonitrile and buffer elut-B: 0.1% formic acid, 80% acetonitrile. The total run duration was 50 minutes, including the regeneration step. Ionization was performed at the liquid junction, with a spray voltage of 1.3 kV applied to a noncoated capillary probe (PicoTip Emitter, 10 μm ID; New Objective). The peptide ions were analyzed using the Nth-dependent method as follows: (1) full mass spectrometry scan (m/z 300–2000); (2) ZoomScan (scan of the 3 major ions); and (3) MS/MS on these 3 ions using standard peptide fragmentation parameters ($Q_z=0.25$, activation time=30 ms, collision energy=40%).

Protein database search and data analysis

For the AUC approach, the data processing pipeline was designed using the TOPPAS software, part of the OpenMS project. X!Tandem, for which executable is directly provided in the OpenMS archive, was used to perform database searches in batch mode within all the files. Peak lists were created by an OpenMS dedicated tool with an additional processing step. Indeed, a precursor recalculation was computed for each tandem spectrum in order to improve the number of matches between a spectrum and a peptide sequence. Proteins were digested by trypsin and two missed trypsin cleavage sites were allowed. Cysteine carbamidomethylation was set as a fixed modification; methionine oxidation and protein N-terminal acetylation were set as variable modifications for all X!Tandem searches. The mass tolerances in MS and MS/MS were set to 10 ppm and 0.6 Da respectively. Data were searched against a target/decoy concatenated database in order to obtain an FDR (false discovery rate *i.e.* number of decoy hits/number of target hits \times 100) value at the peptide level. All identifications were validated with a final peptide FDR of 5 % and a calculated protein FDR of 1 %.

For the SC approach, the raw data produced on LTQ-Orbitrap mass spectrometer were first converted into mzXML format files with ReADW (<http://sashimi.sourceforge.net>) and, in a second step, protein identification was performed with X!Tandem software (X!Tandem CYCLONE 2010.12.01.1 [<http://www.thegpm.org>]) (Fenyő and Beavis, 2003) and the X!TandemPipeline (3.3.0) against the *C. cellulolyticum* (strain ATCC 35319 / DSM 5812 / JCM 6584 / H10) Uniprot database (version January 2012, <http://uniprot.org>). The X!Tandem search parameters were as follows: trypsin specificity with one missed cleavage, fixed alkylation of cysteine, and variable oxidation of methionine. The mass

tolerance was fixed to 10 ppm for precursor ions and 0.5 Da for fragment ions. For all proteins identified with a protein E-value of <0.01 in the first step, additional peptides were searched to reinforce identification, using similar parameters except that semitryptic peptides and protein N-terminal acetylations were accepted. All peptides identified with an E-value lower than <0.1 were conserved. Rarefaction curves were estimated using the R package Vegan (Oksanen *et al.*, 2011).

Label-free quantification

For the AUC approach, the data processing pipeline was designed using the TOPPAS software, part of the OpenMS project. Peaks were detected in each sample using the « PeakPickerCentroid » algorithm provided by the OpenMS software. Validated identification data were matched with detected peaks. Peaks which were assigned to the same peptide sequence in different samples were used as anchors for retention time alignment between those samples. To perform protein relative quantification in different samples, a Protein Abundance Index (PAI) was calculated and defined as the average of XIC (eXtracted Ion Chromatogram) area values from the three most intense peptides identified for a given protein.

For the SC approach, spectral counts of proteins that were identified in each replicate were used and combined statistical tests were applied to assess the significance (Heinecke *et al.*, 2010). Proteins with a false discovery rate <10% for error estimation, a p-value <0.01 for the T-test and a G-value >2.5 for the G-test of each of the two conditions were assigned as differentially expressed proteins.

Statistical analyses of the label-free quantification proteome datasets

For the AUC approach, all statistical analyses were performed using the programming language R (R Core Team, 2013) on \log_{10} -transformed and quantile normalized Protein Abundance Indexes (PAI). The substrate, incubation time and substrate by incubation time interaction effects were assessed by Tobit models with a censoring threshold equal to the minimum non-zero value for a given protein. With y_{ijkl}^* the normalized \log_{10} PAI of protein l , for the substrate S_i , incubation time T_j and replicate k , the model used for each protein is the following:

$$y_{ijkl} = \begin{cases} \min_{i'j'k': y_{i'j'k'l}^* > 0} (y_{i'j'k'l}^*) & \text{if } y_{ijkl}^* = 0 \text{ and } \exists k': y_{ijk'l}^* \neq 0 \\ y_{ijkl}^* & \text{otherwise} \end{cases}$$

$$\text{with } y_{ijkl}^* = \mu_l + \alpha_{il}S_i + \beta_{jl}T_j + \gamma_{ijl}(S, T)_{ij} + \varepsilon_{ijkl}$$

$$\text{and } \varepsilon_{ijkl} \sim \mathcal{N}(0, \sigma_l^2)$$

Tobit regressions were computed using the Markov chain Monte Carlo algorithm from the R package MCMCpack (Martin *et al.*, 2011). Runs that did not pass the Heidelberger and Welch's convergence diagnostic (Heidelberger and Welch, 1983; Plummer *et al.*, 2006) were discarded. Substrate (α_{ij}) and interaction (γ_{jij}) effect coefficients as well as p-values, relative to the tests whether these coefficients are different from zero, were determined from the posterior distributions. P-values were adjusted for multiple comparisons using Benjamini and Hochberg's false discovery rate (Benjamini and Hochberg, 1995). The significance threshold used corresponds to a false discovery rate of 1 %. Kendall tau rank correlation coefficients and corresponding p-values were determined using the R package Kendall (Mcleod, 2011).

GSEA was performed based on clusters of orthologous groups or COGs (Tatusov *et al.*, 2000) recovered from the *Clostridium cellulolyticum* H10 annotation of the Microscope website (Vallenet *et al.*, 2013), on the metabolic pathways recovered from the *Clostridium cellulolyticum* H10 Pathway Genome Database of the Biocyc website (Caspi *et al.*, 2010) and on the Carbohydrate-Active enZYmes or CAZy (Cantarel *et al.*, 2009). COGs and pathways for which less than nine proteins were found in the dataset were removed. In total, six GSEA were performed using three different scores on the two fractions, supernatant and pellet. The first score, linked to the PAI – DOC correlation, was the Kendall tau correlation coefficients. The second score, linked to the substrate effect, was the signs of the facial tissue substrate coefficients times one minus the substrate effect p-values:

$$\text{sgn}(\hat{\alpha}_{tissue,l}) \times (1 - \mathbb{P}(\hat{\alpha}_{tissue,l} = 0)) = \text{sgn}(\hat{\alpha}_{tissue,l}) \times \mathbb{P}(\hat{\alpha}_{tissue,l} \neq 0)$$

This score, derived from the p-values of the substrate effect, was therefore close to 1 (respectively -1) for proteins significantly more (respectively less) concentrated in the tissue incubation protein sample than in the paper one. Similarly, the third score, linked to the interaction effect, was the signs of the interaction coefficients for the facial tissue substrate at 70 hours times one minus the interaction effect p-values:

$$\text{sgn}(\hat{\gamma}_{tissue,70h,l}) \times (1 - \mathbb{P}(\hat{\gamma}_{tissue,70h,l} = 0)) = \text{sgn}(\hat{\gamma}_{tissue,70h,l}) \times \mathbb{P}(\hat{\gamma}_{tissue,70h,l} \neq 0)$$

This score, derived from the p-values of the interaction effect, was therefore close to 1 (respectively -1) for proteins showing a significant positive (respectively negative) interaction coefficient in the tissue incubation protein sample at 70 hours *i.e.* proteins for which level increases more or decreases less between 46 hours and 70 hours in the tissue incubation protein sample than in the paper one.

Statistical significance of the enrichment scores for each protein group was estimated by comparisons with the scores resulting from 10000 permutations of the protein group labels. The p-values thus obtained were not corrected for multiple testing in order to recover maximum knowledge on the

possible trends for all the protein groups, specifically for the pathways containing only a few proteins. However this procedure might lead to false positives and therefore the highlighted protein groups were thereafter systematically checked in more detail to confirm the trends by identifying each of its significantly varying proteins.

Acknowledgments

This work was supported by Région Ile-de-France (R2DS 2010-08 project).

Supplementary materials

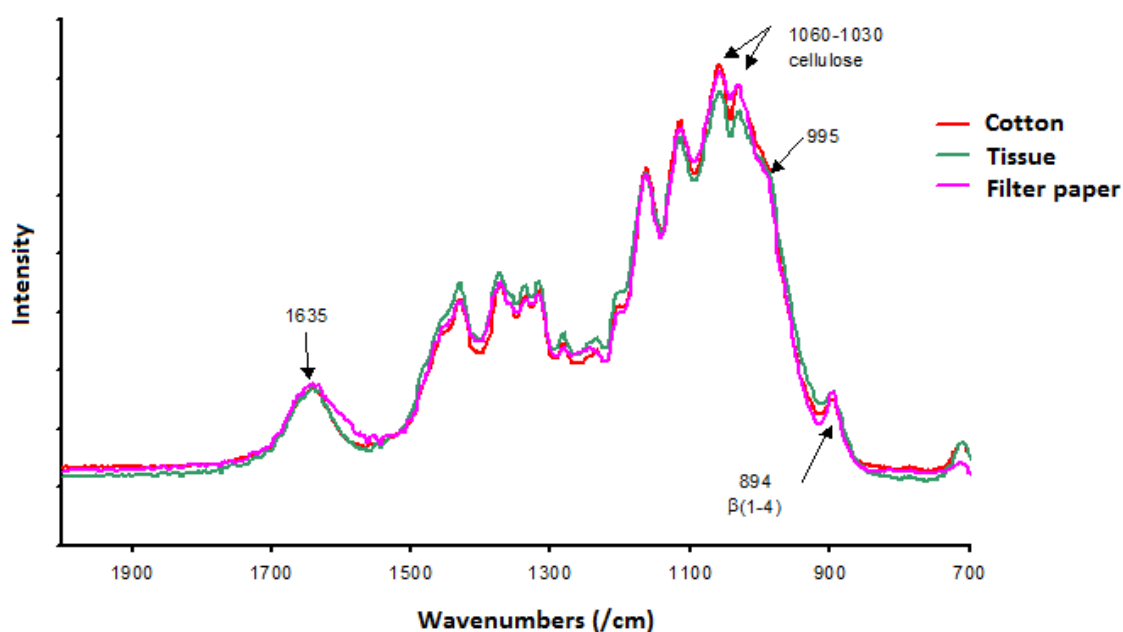


Fig. SII-1. Infrared spectra obtained for cotton, facial tissue and filter paper.

Table SII-1. Quantitative and statistical results from the label-free quantitative approaches based on AUC and SC for all the identified *C. cellulolyticum* CAZymes.

AUC: Area Under Curve, SC: Spectral Counting, in red: cellulosome components, highlight in red: cip-cel gene cluster, highlight in green: xyl-doc gene cluster. For the AUC pipe-line, missing values were replaced by 0 in the shown Protein Abundance Index data.

			AUC pipe-line: Protein Abundance Index based on Area Under Curve															
Gene loci names	UniprotKB Acc	CAZY domains (cellulosomes components)	Pellet - Tissue 46h		Pellet - Tissue 70h		Pellet - Paper 46h		Pellet - Paper 70h		Supernatant - Tissue 46h		Supernatant - Tissue 70h		Supernatant - Paper 46h		Supernatant - Paper 70h	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Ccel_0141	B814H5	GH67	3,3E+06	3,0E+06	7,8E+07	6,5E+07	2,3E+06	4,0E+06	0,0E+00	0,0E+00	2,7E+07	1,5E+07	2,7E+07	3,4E+07	3,7E+07	6,4E+06	1,5E+07	2,7E+07
Ccel_0154	B814I8	GH2	0,0E+00	0,0E+00	0,0E+00	0,0E+00	2,9E+07	5,1E+07	3,7E+07	6,3E+07	0,0E+00	0,0E+00	8,6E+06	8,4E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0203	B81510	GH3	1,8E+07	6,4E+06	1,1E+08	2,4E+07	6,3E+07	2,1E+07	2,7E+07	1,6E+07	1,0E+08	8,4E+07	8,7E+07	7,9E+07	1,6E+08	4,2E+07	8,1E+07	2,5E+07
Ccel_0210	B81517	CE7	8,1E+06	7,0E+06	2,2E+07	3,9E+07	2,8E+07	2,8E+07	2,2E+07	3,0E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	7,6E+06	1,3E+07	2,3E+07	4,0E+07
Ccel_0231	Q0PRN4	GH9,CBM3	8,7E+07	2,2E+07	7,7E+07	3,0E+07	2,9E+08	1,6E+08	1,8E+08	1,1E+08	1,5E+06	2,6E+06	3,9E+07	3,2E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0286	B81593	CE4	5,7E+07	6,7E+06	1,3E+08	1,7E+08	8,7E+07	1,5E+08	1,7E+08	2,2E+07	5,4E+06	9,3E+06	6,2E+07	8,9E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0335	B815Q4	GH15	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,1E+07	2,7E+07	1,3E+07	2,3E+07	4,4E+06	7,7E+06	7,9E+06	1,4E+07	0,0E+00	0,0E+00	6,7E+06	1,2E+07
Ccel_0374	B815U2	GH1	3,2E+06	2,8E+06	1,9E+07	2,3E+07	1,1E+07	1,3E+07	3,1E+06	5,4E+06	2,0E+07	8,4E+06	2,1E+07	3,0E+07	3,0E+07	7,6E+06	2,1E+07	3,5E+06
Ccel_0379	B815U7	GH5,CBM32	1,3E+08	3,1E+07	1,1E+08	1,3E+07	3,1E+08	1,1E+07	1,8E+08	1,0E+08	2,1E+07	1,3E+07	3,6E+07	4,0E+07	6,0E+07	2,1E+07	1,7E+06	3,0E+06
Ccel_0417	B815Y4	PL11	2,2E+07	8,3E+06	6,5E+06	1,1E+07	1,1E+08	7,5E+07	3,6E+07	2,6E+07	1,2E+06	2,1E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0428	B815Z4	GH5,CBM17,CBM28	9,2E+07	2,6E+07	2,2E+08	6,0E+07	2,0E+08	7,0E+07	8,8E+07	5,6E+07	0,0E+00	0,0E+00	5,3E+07	6,0E+07	8,1E+07	4,3E+07	3,8E+07	1,5E+07
Ccel_0429	B815Z5	GH44,CBM44	8,2E+08	1,2E+08	8,9E+08	2,7E+08	1,5E+09	4,7E+08	8,5E+08	5,3E+08	4,7E+07	1,8E+07	9,7E+07	2,2E+07	7,5E+07	2,4E+07	1,3E+07	1,1E+07
Ccel_0486	B816H6	GT28	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	2,2E+07	4,3E+06
Ccel_0567	B81724	CE4	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,3E+07	5,7E+07	3,0E+06	5,2E+06	0,0E+00	0,0E+00	7,1E+06	1,2E+07	6,7E+07	1,2E+08	1,0E+07	1,8E+07
Ccel_0643	B817A0	GH18	0,0E+00	0,0E+00	9,3E+06	9,1E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	5,2E+08	9,0E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0649	B817A6	GH30	1,8E+08	3,3E+07	3,6E+08	1,1E+08	5,4E+08	1,6E+08	3,2E+08	2,4E+08	5,2E+07	3,6E+06	2,1E+08	2,3E+08	1,3E+08	1,1E+07	8,1E+07	3,0E+07

Ccel_0728	B817V0	CBM3	5,7E+09	4,1E+09	4,2E+09	7,8E+08	1,4E+10	3,9E+09	1,2E+10	1,3E+10	6,6E+08	2,3E+08	1,4E+09	1,1E+09	7,0E+08	1,9E+08	8,3E+08	5,1E+08
Ccel_0729	P37698	GH48	1,8E+09	2,0E+08	2,1E+09	6,8E+08	5,2E+09	1,2E+09	4,2E+09	1,9E+09	5,6E+07	1,1E+07	2,1E+08	1,3E+08	1,1E+08	1,0E+08	3,2E+06	5,6E+06
Ccel_0730	P37699	GH8	1,0E+08	1,0E+07	1,3E+08	2,2E+07	2,9E+08	8,2E+07	1,6E+08	1,2E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0731	P37700	GH9,CBM3	1,5E+08	9,0E+07	4,3E+08	6,9E+07	4,8E+08	2,2E+08	3,1E+08	1,2E+08	1,7E+07	1,8E+07	6,1E+07	6,8E+07	4,0E+07	9,3E+06	3,2E+07	2,8E+07
Ccel_0732	B817V4	CBM4,GH9	1,2E+09	1,6E+08	1,4E+09	2,5E+08	4,1E+09	1,2E+09	2,6E+09	1,4E+09	3,8E+07	8,6E+06	8,0E+07	8,7E+07	1,8E+09	2,9E+09	2,7E+07	2,2E+07
Ccel_0733	B817V5	None	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,1E+07	5,4E+07	1,3E+08	1,0E+08	8,6E+07	1,5E+08	0,0E+00	0,0E+00
Ccel_0734	B817V6	GH9,CBM3	7,6E+07	1,2E+07	1,5E+08	2,4E+07	3,1E+08	1,0E+08	1,7E+08	1,0E+08	4,0E+07	6,9E+07	9,4E+07	1,1E+08	4,0E+08	3,8E+08	1,0E+08	9,8E+07
Ccel_0735	B817V7	GH9,CBM3	7,0E+07	1,5E+07	1,9E+07	1,7E+07	2,0E+08	3,9E+07	1,2E+08	9,0E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,9E+07	3,2E+07
Ccel_0736	B817V8	GH5	0,0E+00	0,0E+00	0,0E+00	0,0E+00	5,4E+07	4,7E+07	5,8E+06	1,0E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0737	B817V9	GH9	1,2E+08	8,1E+06	1,9E+08	4,6E+07	3,2E+08	8,2E+07	1,9E+08	1,1E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,2E+07	1,1E+07	0,0E+00	0,0E+00
Ccel_0740	B817W1	GH5	1,1E+08	1,9E+07	5,1E+07	4,7E+07	3,3E+08	8,6E+07	1,9E+08	1,3E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0750	B817X1	GH11	1,9E+08	1,1E+07	3,6E+08	9,1E+07	4,8E+08	1,3E+08	3,7E+08	2,1E+08	0,0E+00	0,0E+00	6,5E+07	6,9E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0752	Q0PRN0	CBM35,GH26	3,6E+07	5,7E+06	3,0E+07	2,1E+07	6,0E+07	2,7E+07	3,9E+07	1,2E+07	0,0E+00	0,0E+00	1,0E+07	1,8E+07	1,8E+07	1,6E+07	0,0E+00	0,0E+00
Ccel_0753	Q0PRM9	GH9,CBM3	1,4E+08	2,9E+07	1,1E+08	4,4E+07	3,3E+08	6,9E+07	2,5E+08	1,6E+08	6,9E+06	7,3E+06	5,7E+06	5,5E+06	3,5E+07	1,0E+07	8,9E+06	9,3E+06
Ccel_0755	B817X6	GH9	2,3E+08	4,0E+07	3,8E+08	6,9E+07	3,6E+08	5,9E+07	2,9E+08	1,8E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0840	P25472	GH5,CBM11	7,9E+06	6,9E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,0E+07	1,8E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0881	B818M0	CBM16	0,0E+00	0,0E+00	0,0E+00	0,0E+00	5,9E+06	1,0E+07	0,0E+00	0,0E+00	1,3E+08	8,0E+07	1,1E+08	9,7E+07	1,3E+09	3,5E+08	3,3E+08	8,4E+07
Ccel_0920	B818Q5	CE1	1,0E+08	5,4E+06	1,9E+08	8,2E+07	1,4E+08	4,8E+07	3,8E+08	6,6E+08	1,1E+08	1,6E+08	1,0E+08	1,3E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_0931	Q0PRN5	GH10	4,1E+08	1,4E+07	5,5E+08	7,8E+07	1,4E+09	9,6E+07	1,1E+09	7,8E+08	6,5E+07	8,9E+06	2,7E+08	2,8E+08	1,7E+08	3,9E+07	8,1E+07	2,9E+07
Ccel_1004	B819A6	GH31	0,0E+00	0,0E+00	1,2E+07	1,1E+07	1,1E+07	1,9E+07	7,3E+06	6,7E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1011	B819B3	GH43	3,8E+06	6,5E+06	4,1E+07	1,1E+07	5,4E+07	1,8E+07	2,2E+07	1,9E+07	1,2E+07	1,1E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	4,1E+06	7,0E+06
Ccel_1036	B819D8	GH51	6,3E+06	3,4E+06	3,6E+07	1,4E+07	3,6E+07	1,1E+07	1,9E+07	1,8E+07	4,0E+07	1,7E+07	5,9E+07	5,4E+07	1,2E+08	2,2E+07	1,2E+08	3,6E+07
Ccel_1060	B819G0	None	2,7E+07	1,1E+06	0,0E+00	0,0E+00	1,4E+08	4,5E+07	8,4E+07	7,6E+07	1,1E+07	1,0E+07	3,7E+07	6,3E+07	2,3E+07	4,0E+07	0,0E+00	0,0E+00
Ccel_1099	P17901	GH5	1,6E+08	4,9E+06	3,0E+08	4,6E+07	4,9E+08	1,5E+08	2,7E+08	1,4E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	4,7E+07	6,8E+07
Ccel_1101	B8HZW0	None	2,7E+07	2,7E+07	1,4E+07	2,3E+07	1,5E+08	7,9E+07	7,0E+07	3,0E+07	0,0E+00	0,0E+00	2,5E+07	4,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1139	B8HZZ6	GH3	1,2E+07	1,6E+06	6,4E+07	2,1E+07	1,8E+07	4,9E+06	1,6E+07	6,3E+06	6,6E+07	7,9E+07	4,5E+07	4,0E+07	1,9E+07	1,7E+07	3,0E+07	1,1E+07
Ccel_1207	B8I0I9	GH74	1,2E+08	2,5E+07	2,4E+08	6,5E+07	4,5E+08	1,6E+08	2,3E+08	1,7E+08	0,0E+00	0,0E+00	1,3E+07	2,2E+07	0,0E+00	0,0E+00	9,3E+06	1,6E+07
Ccel_1221	B8I0K2	GH51	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,2E+06	5,6E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1229	B8I0L0	GH43,CBM6	2,7E+06	4,6E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1230	B8I0L1	GH10,CBM6	2,4E+07	3,5E+06	2,1E+07	6,1E+06	8,9E+07	4,7E+07	3,0E+07	1,7E+07	1,2E+06	2,1E+06	4,8E+07	5,6E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00

Ccel_1232	B8I0L3	CE1,CBM6	4,2E+07	6,3E+06	3,0E+07	2,6E+07	8,1E+07	1,8E+07	3,5E+07	3,9E+07	0,0E+00	0,0E+00	1,8E+07	3,1E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1235	B8I0L6	GH43,CBM6	7,1E+06	1,0E+06	0,0E+00	0,0E+00	1,4E+07	1,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1237	B8I0L8	GH27,CBM6	2,2E+06	3,9E+06	1,2E+07	2,1E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1239	B8I0M0	GH2,CBM6	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,2E+07	2,1E+07	0,0E+00	0,0E+00	6,3E+07	5,5E+07	0,0E+00	0,0E+00
Ccel_1240	B8I0M1	GH62,CBM6,CE6	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,0E+07	1,8E+07	2,6E+07	2,3E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1242	B8I0M3	GH30,CBM6	2,0E+08	8,4E+07	0,0E+00	0,0E+00	1,3E+07	2,3E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1246	B8I0M7	CE8	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,3E+07	3,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1249	B8I0N0	GH9,CBM3	1,3E+08	1,8E+07	1,4E+08	3,6E+07	4,6E+08	1,3E+08	2,6E+08	1,9E+08	0,0E+00	0,0E+00	6,8E+06	1,2E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1258	B8I0N9	GH8	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,2E+07	2,0E+07	0,0E+00	0,0E+00
Ccel_1259	B8I0P0	GH39	6,3E+06	5,5E+06	1,2E+07	1,2E+07	1,3E+07	1,2E+07	8,4E+06	6,6E+06	1,3E+07	5,0E+06	3,6E+07	2,8E+07	3,6E+07	3,1E+06	2,4E+07	6,1E+06
Ccel_1260	B8I0P1	CE1	0,0E+00	0,0E+00	6,9E+06	1,2E+07	1,7E+06	2,9E+06	0,0E+00	0,0E+00	1,5E+07	1,5E+07	2,0E+07	3,4E+07	3,0E+07	7,6E+06	5,5E+06	9,5E+06
Ccel_1307	B8I0T6	GT2	0,0E+00	0,0E+00	0,0E+00	0,0E+00	5,5E+07	5,0E+07	2,7E+07	4,7E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,9E+07	3,3E+07	3,4E+07	5,9E+07
Ccel_1439	B8I1W7	GH94	6,0E+06	5,4E+06	5,9E+07	5,4E+07	5,4E+07	3,1E+07	1,1E+07	1,3E+07	2,0E+07	1,1E+07	1,9E+07	2,3E+07	7,6E+07	2,3E+07	4,6E+07	2,7E+07
Ccel_1491	B8I2I8	CBM3	1,7E+07	1,3E+07	1,8E+07	4,8E+06	4,2E+07	1,3E+07	2,7E+07	2,8E+07	1,7E+08	7,2E+07	1,1E+08	4,1E+07	6,3E+08	2,4E+08	1,6E+08	4,1E+07
Ccel_1543	B8I269	None	0,0E+00	0,0E+00	4,0E+07	3,5E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,9E+08	1,6E+08	5,6E+08	3,9E+08	3,7E+08	1,6E+08	1,2E+08	7,3E+07
Ccel_1549	B8I275	GH120	0,0E+00	0,0E+00	7,2E+07	1,7E+07	7,5E+06	1,3E+07	7,3E+06	9,7E+06	1,7E+07	4,7E+06	2,5E+07	1,0E+07	2,2E+07	1,1E+07	1,2E+07	3,2E+06
Ccel_1550	B8I276	GH18	1,7E+07	1,6E+07	1,2E+07	2,1E+07	2,1E+07	3,7E+07	1,7E+07	3,0E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1597	B8I2F6	GH27	7,8E+07	1,3E+07	6,9E+07	6,5E+07	1,5E+08	4,4E+07	9,6E+07	6,0E+07	3,4E+07	2,4E+05	1,3E+08	1,2E+08	2,6E+07	4,5E+07	1,4E+07	2,4E+07
Ccel_1648	B8I2K6	GH9,CBM3,CBM3	9,1E+07	2,1E+07	8,1E+07	2,0E+07	2,9E+08	5,1E+07	1,7E+08	1,4E+08	2,0E+07	5,8E+06	8,3E+07	7,4E+07	6,2E+07	2,0E+07	2,9E+07	1,1E+07
Ccel_1655	B8I2L3	None	4,9E+07	8,5E+06	6,6E+07	2,0E+07	1,5E+08	4,0E+07	1,1E+08	5,5E+07	0,0E+00	0,0E+00	4,5E+07	6,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1656	B8I2L4	CBM6	4,2E+06	3,7E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,8E+07	3,1E+07	5,4E+06	9,4E+06	0,0E+00	0,0E+00	9,9E+06	1,7E+07	0,0E+00	0,0E+00
Ccel_1809	B8I3I6	None	2,4E+07	2,4E+07	2,1E+07	3,6E+07	2,4E+07	4,2E+07	1,8E+07	1,8E+07	0,0E+00	0,0E+00	9,9E+07	8,7E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1888	B8I393	GT5	3,7E+06	3,8E+06	1,4E+07	1,2E+07	3,3E+06	5,7E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_1972	B8I3H7	GH43	1,3E+07	1,1E+07	2,8E+07	3,3E+07	1,2E+07	2,0E+07	2,8E+07	2,4E+07	2,8E+08	1,7E+08	7,6E+07	5,5E+07	1,4E+08	3,4E+07	4,9E+07	5,2E+07
Ccel_2017	B8I3M0	None	3,7E+07	1,8E+06	6,2E+07	1,1E+07	8,7E+07	2,8E+07	5,9E+07	4,0E+07	7,1E+06	6,6E+06	4,8E+07	6,3E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2109	B8I421	GH94	5,9E+07	2,3E+07	2,6E+08	6,7E+07	2,5E+08	7,3E+07	2,0E+08	1,2E+08	4,2E+08	1,6E+08	3,7E+08	2,1E+08	9,3E+08	1,1E+08	6,5E+08	2,4E+08
Ccel_2123	B8I434	GH53	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,5E+07	2,6E+07	4,3E+06	7,5E+06	0,0E+00	0,0E+00	1,1E+07	1,9E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2162	B8I471	CE2	1,2E+08	1,6E+07	2,1E+08	5,4E+07	5,8E+08	1,8E+08	3,0E+08	2,4E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2169	B8I478	CBM25	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	2,0E+07	3,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2215	B8I4Q3	GT35	6,5E+07	1,6E+07	2,6E+08	8,9E+07	2,6E+08	8,9E+07	1,3E+08	7,4E+07	6,4E+07	3,2E+07	1,6E+08	1,6E+08	1,1E+08	3,8E+07	1,6E+08	7,0E+07

Ccel_2226	B8I4R4	GH9	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2243	B8I4T1	PL1	5,1E+07	2,7E+06	4,1E+07	9,5E+06	1,3E+08	3,7E+07	8,4E+07	5,5E+07	2,3E+07	2,0E+07	1,2E+08	1,4E+08	5,6E+07	1,9E+07	1,4E+07	1,3E+07
Ccel_2319	B8I5B9	CBM22,CBM22,CBM22,GH10,CBM9	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,9E+07	3,5E+07	0,0E+00	0,0E+00	9,6E+06	9,9E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2320	B8I5C0	CBM22,CBM22,GH10,CBM9	0,0E+00	0,0E+00	1,1E+07	1,9E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	2,1E+06	3,6E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	6,8E+06	1,2E+07
Ccel_2337	B8I5D7	GH5	1,1E+08	4,7E+07	1,7E+08	2,7E+07	1,0E+08	3,3E+06	8,3E+07	2,7E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2354	B8I5F2	GH94	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	8,3E+06	1,4E+07	6,2E+06	1,1E+07	0,0E+00	0,0E+00	4,1E+06	7,2E+06
Ccel_2392	B8I5I9	CBM4,GH9	6,0E+07	1,2E+07	3,6E+07	8,0E+06	1,3E+08	3,4E+07	6,1E+07	4,3E+07	1,0E+07	1,8E+07	0,0E+00	0,0E+00	2,4E+07	4,1E+07	0,0E+00	0,0E+00
Ccel_2395	B8I5J2	CE3	0,0E+00	0,0E+00	5,8E+06	1,0E+07	3,4E+07	3,1E+07	6,9E+06	1,2E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2417	B8I5L1	GT39	0,0E+00	0,0E+00	1,6E+08	2,8E+08	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	5,2E+07	2,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2442	B8I606	None	2,3E+07	1,1E+07	1,4E+07	2,4E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,2E+09	8,5E+08	6,1E+08	5,3E+08	3,4E+09	8,8E+08	7,5E+08	1,6E+08
Ccel_2455	B8I619	GH31	0,0E+00	0,0E+00	4,0E+07	3,2E+07	3,6E+07	2,0E+07	1,2E+07	1,3E+07	3,7E+07	5,2E+07	6,8E+07	8,8E+07	2,5E+08	2,9E+07	9,0E+07	4,3E+07
Ccel_2530	B8I693	CBM13	3,2E+07	6,6E+06	8,9E+07	2,0E+07	1,0E+08	4,0E+07	5,7E+07	4,4E+07	0,0E+00	0,0E+00	4,9E+06	8,4E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2577	B8I6R3	GT84,GH94	8,9E+06	8,0E+06	1,0E+08	4,0E+07	2,8E+07	4,7E+07	4,6E+06	8,0E+06	3,0E+07	4,5E+07	7,4E+06	8,8E+06	1,3E+08	7,8E+07	1,2E+07	2,0E+07
Ccel_2614	B8I6V0	GH43	0,0E+00	0,0E+00	4,0E+07	3,8E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	7,3E+06	1,3E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_2621	B8I6V7	CBM30,GH9	2,0E+07	3,2E+06	2,8E+07	1,4E+07	2,8E+07	1,5E+07	4,0E+06	6,9E+06	0,0E+00	0,0E+00	9,3E+06	1,6E+07	6,1E+07	5,3E+07	4,2E+07	7,3E+07
Ccel_2669	B8I6Z9	CBM9,CE1	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,2E+07	1,2E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	4,7E+06	8,1E+06
Ccel_2876	B8I870	GT51	1,5E+07	5,8E+06	5,1E+07	5,8E+06	1,5E+07	2,6E+07	1,6E+07	2,8E+07	0,0E+00	0,0E+00	6,2E+06	1,1E+07	0,0E+00	0,0E+00	1,8E+07	1,8E+07
Ccel_2883	B8I877	GT4	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,4E+08	2,4E+08
Ccel_2893	B8I887	GH18	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	2,9E+07	4,3E+06	0,0E+00	0,0E+00
Ccel_2921	B8I8B5	GT26	0,0E+00	0,0E+00	5,2E+06	4,6E+06	7,8E+06	2,8E+06	3,2E+06	3,9E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_3181	B8I0E5	GT1	0,0E+00	0,0E+00	4,4E+07	7,6E+07	7,3E+09	1,3E+10	3,3E+09	5,7E+09	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00
Ccel_3240	B8I0X3	None	4,5E+06	7,7E+06	1,6E+07	2,8E+07	0,0E+00	0,0E+00	0,0E+00	0,0E+00	1,4E+07	1,2E+07	1,2E+08	7,5E+07	1,3E+07	2,3E+07	5,0E+07	6,0E+07
Ccel_3410	B8I1R3	GH3	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	7,9E+06	1,4E+07
Ccel_3412	B8I1R5	GH94	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	0,0E+00	3,1E+07	3,4E+07	2,1E+07	2,0E+07	2,4E+07	4,2E+07	1,0E+08	2,8E+07
Ccel_3438	B8I1U1	GH43	9,6E+06	2,4E+06	6,2E+07	1,4E+07	3,7E+07	1,2E+07	1,2E+07	2,0E+07	2,7E+07	1,2E+07	4,3E+07	3,5E+07	3,4E+07	1,3E+07	4,1E+07	1,2E+07
Ccel_3463	B8I288	CBM9,CE1	2,5E+07	7,4E+06	3,1E+07	2,0E+07	5,7E+07	1,6E+07	6,1E+06	6,1E+06	0,0E+00	0,0E+00	0,0E+00	0,0E+00	4,9E+07	5,6E+07	0,0E+00	0,0E+00

Gene loci names	UniprotKB Acc	CAZY domains (cellulosomes components)	SC pipe-line: number of unique peptides															
			Pellet - Tissue 46h		Pellet - Tissue 70h		Pellet - Paper 46h		Pellet - Paper 70h		Supernatant - Tissue 46h		Supernatant - Tissue 70h		Supernatant - Paper 46h		Supernatant - Paper 70h	
			Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Ccel_0141	B8I4H5	GH67	3	3	8	2			2	0	13	4	8	2	4	1	8	1
Ccel_0154	B8I4I8	GH2																
Ccel_0203	B8I5I0	GH3	8	4	10	1	7	2	8	2	17	1	14	7	11	3	13	2
Ccel_0210	B8I5I7	CE7	5	1	5	1	4	0	4	1	5	1	3	2	4	1	6	2
Ccel_0231	Q0PRN4	GH9,CBM3	20	1	7	1	17	5	20	2			8	3			1	1
Ccel_0286	B8I593	CE4	8	1	5	1	5	2	6	0	4	2	5	3			2	0
Ccel_0335	B8I5Q4	GH15																
Ccel_0374	B8I5U2	GH1			1	0					4	2	4	1	2	1	4	0
Ccel_0379	B8I5U7	GH5,CBM32	17	2	11	1	16	1	17	2	4	1	7	1	5	0	4	1
Ccel_0417	B8I5Y4	PL11	11	1	7	2	13	0	12	1	2	1			2	1	1	1
Ccel_0428	B8I5Z4	GH5,CBM17,CBM28	32	1	28	1	28	1	28	2	12	3	16	2	16	3	17	1
Ccel_0429	B8I5Z5	GH44,CBM44	42	3	28	5	36	3	37	2	8	4	13	3	8	1	6	2
Ccel_0486	B8I6H6	GT28																
Ccel_0567	B8I724	CE4																
Ccel_0643	B8I7A0	GH18																
Ccel_0649	B8I7A6	GH30	31	3	29	4	27	1	31	4	14	5	24	2	13	1	13	2
Ccel_0728	B8I7V0	CBM3	70	1	41	3	62	6	72	1	29	8	33	5	22	2	23	2
Ccel_0729	P37698	GH48	52	7	40	2	54	2	61	8	17	7	22	8	14	2	14	2
Ccel_0730	P37699	GH8	16	0	10	2	15	2	17	1					1	0		
Ccel_0731	P37700	GH9,CBM3	42	3	33	3	42	0	43	2	6	4	14	3	7	2	5	1
Ccel_0732	B8I7V4	CBM4,GH9	69	7	48	3	65	4	74	4	15	4	17	3	11	3	7	1
Ccel_0733	B8I7V5	None	8	1	9	1	9	0	8	1	8	0	9	2	9	0	8	1

Ccel_0734	B8I7V6	GH9,CBM3	25	1	15	2	24	1	26	1	1	0	2	1				
Ccel_0735	B8I7V7	GH9,CBM3	13	1	3	1	16	2	12	1								
Ccel_0736	B8I7V8	GH5	7	2	2	1	6	2	7	2								
Ccel_0737	B8I7V9	GH9	17	2	12	1	15	2	16	2			3	2	2	2	3	0
Ccel_0740	B8I7W1	GH5	16	2	8	3	15	1	15	1								
Ccel_0750	B8I7X1	GH11	18	2	14	1	15	1	18	3	7	1	8	2	6	0	5	1
Ccel_0752	Q0PRN0	CBM35,GH26	13	2	8	0	9	2	10	1	2	1	3	2	2	1	2	1
Ccel_0753	Q0PRM9	GH9,CBM3	31	3	14	1	27	3	29	1	8	2	8	1	8	1	7	2
Ccel_0755	B8I7X6	GH9	26	1	19	2	18	1	25	3	7	4	12	2	9	2	8	1
Ccel_0840	P25472	GH5,CBM11	7	2	3	0			5	1	1	0						
Ccel_0881	B8I8M0	CBM16	5	2	4	1	8	1	8	1	16	2	14	1	22	3	22	1
Ccel_0920	B8I8Q5	CE1	11	1	10	1	11	1	11	1	7	2	10	5	5	1	4	1
Ccel_0931	Q0PRN5	GH10	14	1	8	1	14	2	18	2	7	1	7	1	8	1	7	2
Ccel_1004	B8I9A6	GH31																
Ccel_1011	B8I9B3	GH43	6	1	7	1	5	2	7	2	2	1	4	1	1	0	4	1
Ccel_1036	B8I9D8	GH51	3	1	5	1	3	2	4	0	6	2	7	1	7	1	8	2
Ccel_1060	B8I9G0	None	5	1	3	1	5	1	7	1			3	1				
Ccel_1099	P17901	GH5	17	1	15	5	20	2	19	1	2	1	2	2	3	1	2	1
Ccel_1101	B8HZW0	None	11	1	7	1	13	1	16	1	2	1	4	3				
Ccel_1139	B8HZZ6	GH3	6	1	10	2	3	1	4	1	10	2	11	2	5	3	8	2
Ccel_1207	B8I0I9	GH74	28	3	23	2	28	1	27	1	4	2	6	2	5	2	3	0
Ccel_1221	B8I0K2	GH51	2	0	3	0	3	1	2	1	7	1	4	1	7	0	3	0
Ccel_1229	B8I0L0	GH43,CBM6	6	1	2	1	5	1	4	1			1	1				
Ccel_1230	B8I0L1	GH10,CBM6	13	1	6	3	12	2	10	1	2	2	7	4	2	1	1	1
Ccel_1232	B8I0L3	CE1,CBM6	9	1	6	2	8	1	7	2	5	1	5	2	1	0		
Ccel_1235	B8I0L6	GH43,CBM6	3	2	1	0	3	1	2	1								
Ccel_1237	B8I0L8	GH27,CBM6	3	1			1	0	1	0								
Ccel_1239	B8I0M0	GH2,CBM6	2	1	1	0	1	0	1	1	1	0	1	0				
Ccel_1240	B8I0M1	GH62,CBM6,CE6	15	1	5	1	7	2	7	1								
Ccel_1242	B8I0M3	GH30,CBM6	5	1	1	0	2	0	3	1	1	0	1	1	1	1	1	0

Ccel_1246	B8I0M7	CE8																
Ccel_1249	B8I0N0	GH9,CBM3	25	1	17	2	26	3	26	2	2	0	4	1	4	2	3	2
Ccel_1258	B8I0N9	GH8																
Ccel_1259	B8I0P0	GH39			2	2	2	1			4	1	3	2	3	2	4	0
Ccel_1260	B8I0P1	CE1	2	1	4	1	2	2	2	1	5	1	5	2	4	1	4	1
Ccel_1307	B8I0T6	GT2																
Ccel_1439	B8I1W7	GH94			3	1	7	1	5	2	9	1	4	2	12	2	10	1
Ccel_1491	B8I2I8	CBM3	19	6	11	2	16	7	18	4	45	5	37	1	47	1	37	5
Ccel_1543	B8I269	None	7	4	5	2	7	2	6	2	21	3	17	1	13	1	13	2
Ccel_1549	B8I275	GH120	3	1	10	2	2	1	2	1	4	1	8	1	4	1	5	2
Ccel_1550	B8I276	GH18	11	2	3	1	7	2	9	1								
Ccel_1597	B8I2F6	GH27	13	1	10	1	9	1	11	2	11	4	11	2	7	1	5	2
Ccel_1648	B8I2K6	GH9,CBM3,CBM3	27	2	11	4	26	3	27	1	9	2	15	4	11	3	11	1
Ccel_1655	B8I2L3	None	11	1	7	1	10	2	11	1	5	2	9	1	5	1	4	1
Ccel_1656	B8I2L4	CBM6	7	1			3	2	4	1	2	0	5	1				
Ccel_1809	B8I3I6	None	17	2	11	1	16	1	15	0	9	2	13	3	6	1	7	2
Ccel_1888	B8I393	GT5	3	2			4	1	2	0								
Ccel_1972	B8I3H7	GH43	11	1	6	1	8	2	10	2	20	4	15	2	17	1	12	2
Ccel_2017	B8I3M0	None	7	1	4	2	4	1	7	2	2	0	5	0				
Ccel_2109	B8I421	GH94	21	3	21	2	22	2	21	1	32	2	27	4	33	3	31	3
Ccel_2123	B8I434	GH53	5	1			3	2	4	1								
Ccel_2162	B8I471	CE2	15	1	13	1	15	2	18	2								
Ccel_2169	B8I478	CBM25	2	1	3	0	3	1	3	1								
Ccel_2215	B8I4Q3	GT35	21	1	28	1	23	1	26	1	16	2	21	2	12	2	19	1
Ccel_2226	B8I4R4	GH9									3	1			2	1	2	1
Ccel_2243	B8I4T1	PL1	15	2	4	2	12	3	13	2	9	5	15	4	5	1	4	1
Ccel_2319	B8I5B9	CBM22,CBM22,CBM22,GH10,CBM9																
Ccel_2320	B8I5C0	CBM22,CBM22,GH10,CBM9																
Ccel_2337	B8I5D7	GH5	18	1	10	3	12	2	16	4								
Ccel_2354	B8I5F2	GH94	3	1	4	1	2	1			10	3	8	2	10	4	12	3

Ccel_2392	B8I5I9	CBM4,GH9	25	1	12	1	18	2	17	2	3	2	4	1	3	2	3	1
Ccel_2395	B8I5J2	CE3	3	1	2	1	2	1	3	1			2	0				
Ccel_2417	B8I5L1	GT39																
Ccel_2442	B8I6O6	None	13	2	6	2	10	2	11	1	34	2	27	2	34	1	29	4
Ccel_2455	B8I6I9	GH31	4	0	7	1	7	1	5	1	5	1	3	2	6	1	9	2
Ccel_2530	B8I693	CBM13	7	1	6	1	6	1	5	1			2	1	1	1		
Ccel_2577	B8I6R3	GT84,GH94																
Ccel_2614	B8I6V0	GH43	7	1	9	2	9	1	7	2	9	1	7	3	10	1	10	1
Ccel_2621	B8I6V7	CBM30,GH9	25	2	9	5	16	3	11	3	4	3	7	2	3	1	4	1
Ccel_2669	B8I6Z9	CBM9,CE1																
Ccel_2876	B8I870	GT51																
Ccel_2883	B8I877	GT4																
Ccel_2893	B8I887	GH18																
Ccel_2921	B8I8B5	GT26			2	1												
Ccel_3181	B8I0E5	GT1					6	1	4	1								
Ccel_3240	B8I0X3	None																
Ccel_3410	B8I1R3	GH3																
Ccel_3412	B8I1R5	GH94	13	2	16	1	14	2	12	3	22	1	18	2	15	2	20	2
Ccel_3438	B8I1U1	GH43	4	1	5	0	4	1	4	2	4	1	6	1	2	1	4	2
Ccel_3463	B8I288	CBM9,CE1	12	1	9	1	12	1	8	1					7	5	2	1

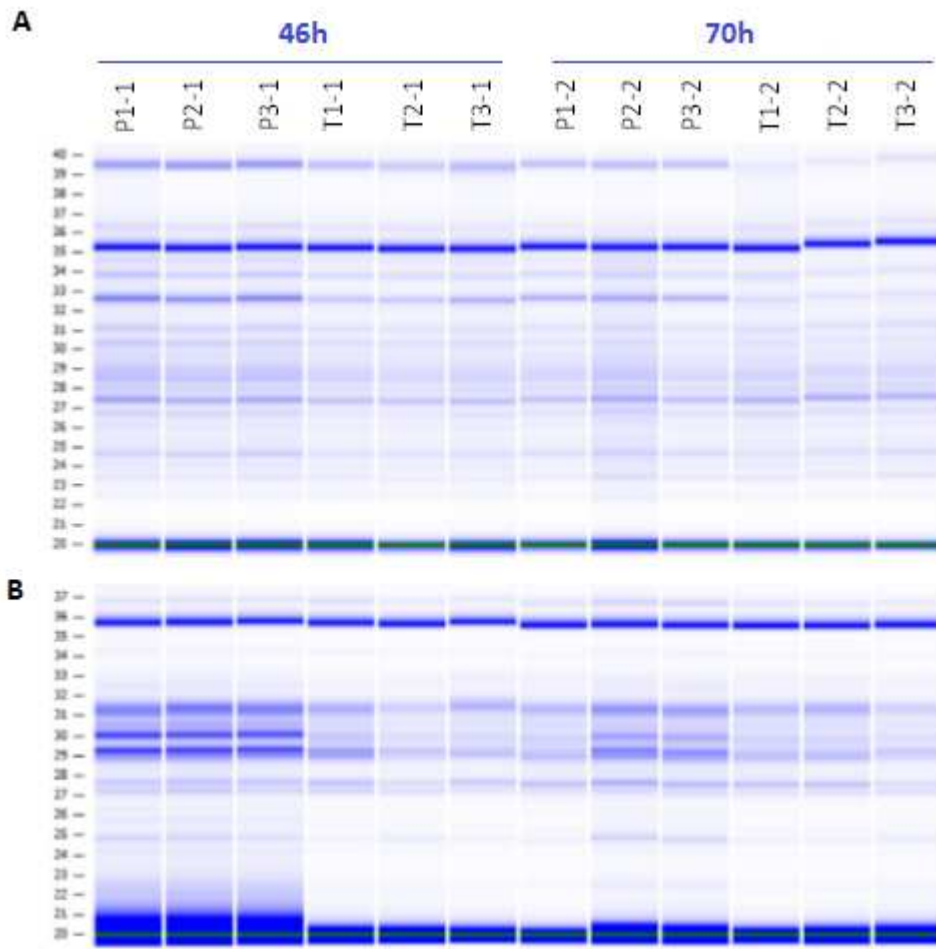


Fig. SII-2. Gel-like images of proteins from total proteomes (A) and exoproteomes (B). Profiles were obtained using Agilent 2100 Bioanalyzer with a High Sensitivity Protein 250 kit. P : filter paper, T : facial tissue.

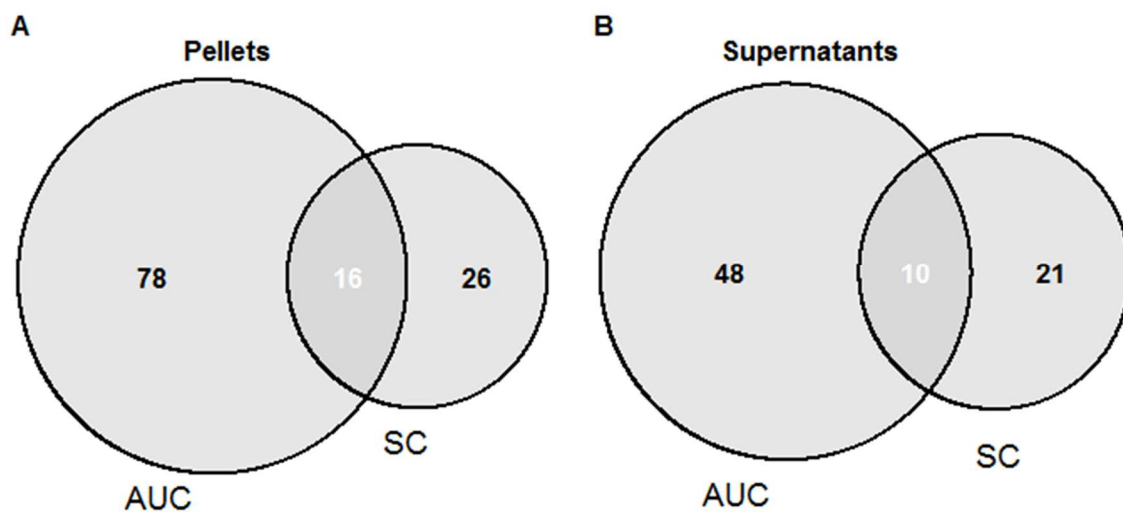


Fig. SII-3. Proteins with a significant effect based on AUC and SC approaches.

Point-clés

Confirmations :

- Rôle important du flux carboné entrant, *via* la valeur de pH ou la saturation des voies métaboliques, dans les profils fermentaires de *C. cellulolyticum* (Desvaux *et al.*, 2000; M Desvaux *et al.*, 2001b)

Nouveaux éléments :

- Premier aperçu en protéomique, et protéomique quantitative, du protéome total de *C. cellulolyticum*
- Influence de caractéristiques physico-chimiques du substrat sur la colonisation
- Influence de caractéristiques physico-chimiques du substrat sur les profils fermentaires
- Rôle probablement important du degré de polymérisation de la cellulose, mais certainement en combinaison avec d'autres facteurs

Complément concernant les expériences réalisées sur cultures de souches pures

Afin de mettre en évidence les interactions entre micro-organismes durant la méthanisation de déchets cellulotiques simples et de faire un lien avec l'étude précédente, des incubations en co-culture ont été réalisées avec *C. cellulolyticum* et *M. barkeri*. Cette archée méthanogène est capable d'utiliser à la fois l'acétate, le dihydrogène et le dioxyde de carbone produits par la bactérie cellulolytique au cours de la fermentation de la cellulose pour la production de méthane (méthanogènes acétoclastiques et hydrogénotrophes). Malgré plusieurs tentatives de co-cultures, la production de méthane n'a jamais été observée bien que la cellulose ait été dégradée avec production d'acétate. Durant les premiers jours de l'incubation, le pH évoluait dans des gammes non inhibitrices pour la méthanogénèse. D'après ces résultats, il semblerait que la production de méthane soit inhibée en co-culture avec *C. cellulolyticum*.

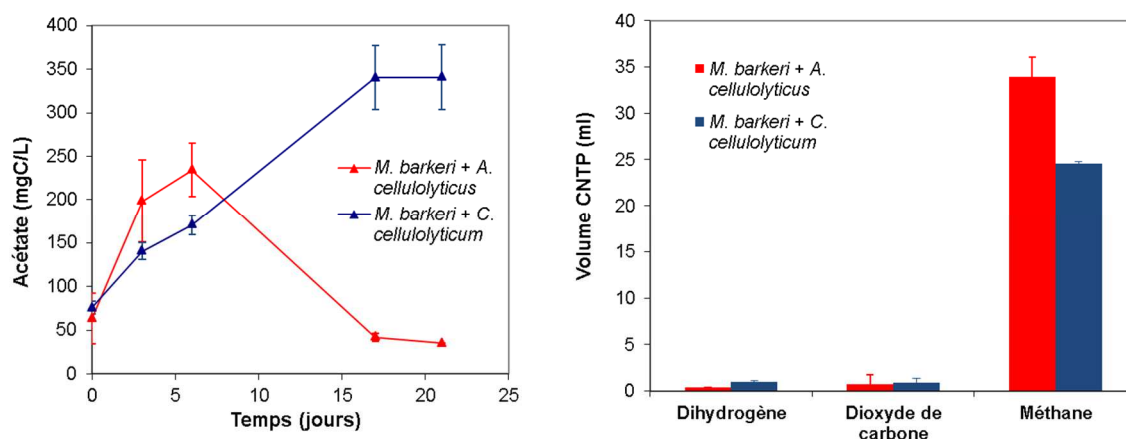


Figure 15. Suivi de la production d'acétate (gauche) et composition finale du biogaz (droite) lors de la co-culture de *M. barkeri* avec *C. cellulolyticum* ou *A. cellulolyticus*.

Cette expérience a été répétée avec chaque souche cellulolytique, *C. cellulolyticum* et *A. cellulolyticus*, afin de tester cette hypothèse. En comparaison avec les essais précédents, la durée de pré-culture de l'archée a été plus longue et une importante quantité de méthane a été produite pendant cette période mais n'a pas pu être mesurée. Néanmoins, les résultats ont montré qu'en présence d'*A. cellulolyticus* plus de méthane et moins de dihydrogène étaient produits et que la concentration en acétate diminuait après une phase d'accumulation (Figure 15). Ceci montre que durant la dégradation du substrat par *A. cellulolyticus*, l'archée méthanogène était capable d'utiliser les produits de fermentation de la cellulose, acétate et H_2/CO_2 , pour produire du méthane. En effet les résultats semblent cohérents puisqu'environ 0.42 mmol d'acétate ont été consommées lors de la co-culture avec *A. cellulolyticus* et la différence de production de méthane entre les deux co-cultures est de 0.45

mmol, en considérant en première approximation que les quantités de méthane mesurées sur la co-culture avec *C. cellulolyticum* correspondent à une production survenue exclusivement durant la pré-culture de l'archée seule. Cependant, les faibles quantités finales en H₂ et CO₂ observées lors des deux co-cultures, en comparaison avec celles observées en cultures pures, suggèrent qu'une partie du méthane, environ 4 ml, a pu être produite par la voie hydrogénotrophe. Des analyses de la composition isotopique du méthane par IRMS ont permis de confirmer que la voie majoritaire pour la production de méthane était bien la voie acétoclaste.

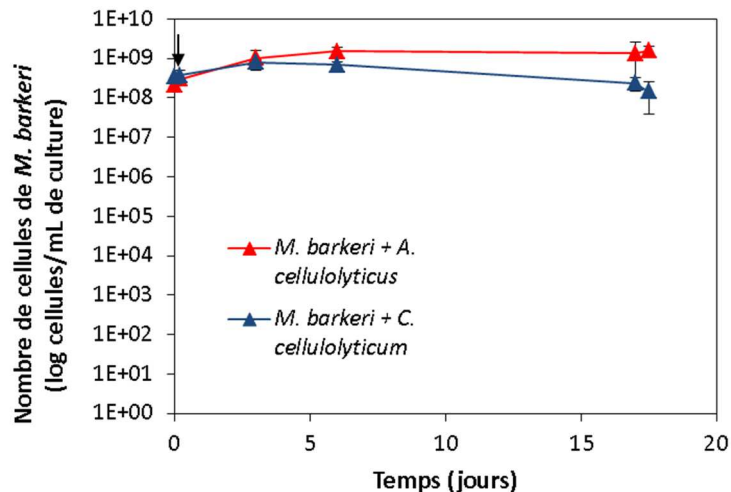


Figure 16. Suivi de la croissance de *M. barkeri* en co-culture avec une bactérie cellulolytique par PCR quantitative.

Des analyses en qPCR ont été effectuées afin de suivre la croissance des souches en co-cultures (Figure 16). D'après les résultats obtenus, tandis que l'ensemble des souches bactériennes présentaient une courbe de croissance classique, la quantité d'archée diminuait fortement après 3 jours d'incubation (d'un facteur 10 environ). L'ensemble de ces résultats suggère que la présence de *C. cellulolyticum* pourrait inhiber l'activité de *M. barkeri* entraînant finalement une diminution de sa population suite à un épuisement énergétique. Il pourrait aussi s'agir d'un phénomène plus direct de toxicité. Des expériences complémentaires sont nécessaires afin de valider cette hypothèse et les résultats ne donneront donc pas lieu à la rédaction d'un article. En particulier, il serait nécessaire de déterminer si l'inhibition est due à des conditions physico-chimiques modifiées pendant la croissance de *C. cellulolyticum* et non favorables pour l'activité de *M. barkeri* comme par exemple l'épuisement d'une ressource, ou de la production d'une molécule inhibitrice, voire toxique. Les résultats obtenus sont cohérents avec ceux de la littérature puisque *M. barkeri* a déjà été co-cultivée avec succès avec *A. cellulolyticus* (Laube and Martin, 1981). En revanche aucune étude n'a été réalisée sur une co-culture entre *M. barkeri* et *C. cellulolyticum*. Si ce phénomène d'inhibition était confirmé ces résultats

contribueraient à un cadre de référence pour la compréhension des structures des communautés complexes issues de méthaniseurs lors de la dégradation de composés lignocellulosiques.

En parallèle des co-cultures entre les souches cellulolytiques et l'archée méthanogène, une expérience de co-culture entre les deux souches cellulolytiques a été réalisée. Malheureusement, ni la colonisation ni la croissance des deux souches n'a pu être suivie puisqu'aucun des marqueurs classiquement utilisés, gènes des ARNr 16S ou type de coloration de Gram, n'était réellement spécifique de l'une des deux souches. Ces résultats ne seront pas présentés ici mais pourraient être analysés lorsque des marqueurs spécifiques auront été validés.

CHAPITRE 3

Avant-propos

Afin de confronter certaines hypothèses ou conclusions de l'article présenté précédemment à des résultats provenant de conditions expérimentales plus proches d'une méthanisation industrielle, les trois substrats, mouchoir en papier, papier filtre et coton en disques, ont été étudiés lors de leur méthanisation par des communautés microbiennes complexes dans des incubations en culture discontinues (« batch ») non agitées de type BMP (Figure 17). Le milieu réactionnel consistait en une solution minérale correspondant à la norme BMP NF EN ISO 11734, supplémentée d'une solution de carbonates (0.25M), des différents substrats et, d'un inoculum stabilisé, filtré et concentré et conservé à -80°C issu du digesteur mésophile situé à Varennes-Jarcy (91). En comparaison avec l'étude présentée dans le chapitre I, le ratio inoculum/substrat (basé sur les teneurs en matières volatiles) utilisé ici est plus faible. En effet, les substrats utilisés semblent moins récalcitrant à la dégradation, comme le montrent les teneurs en matières volatiles.

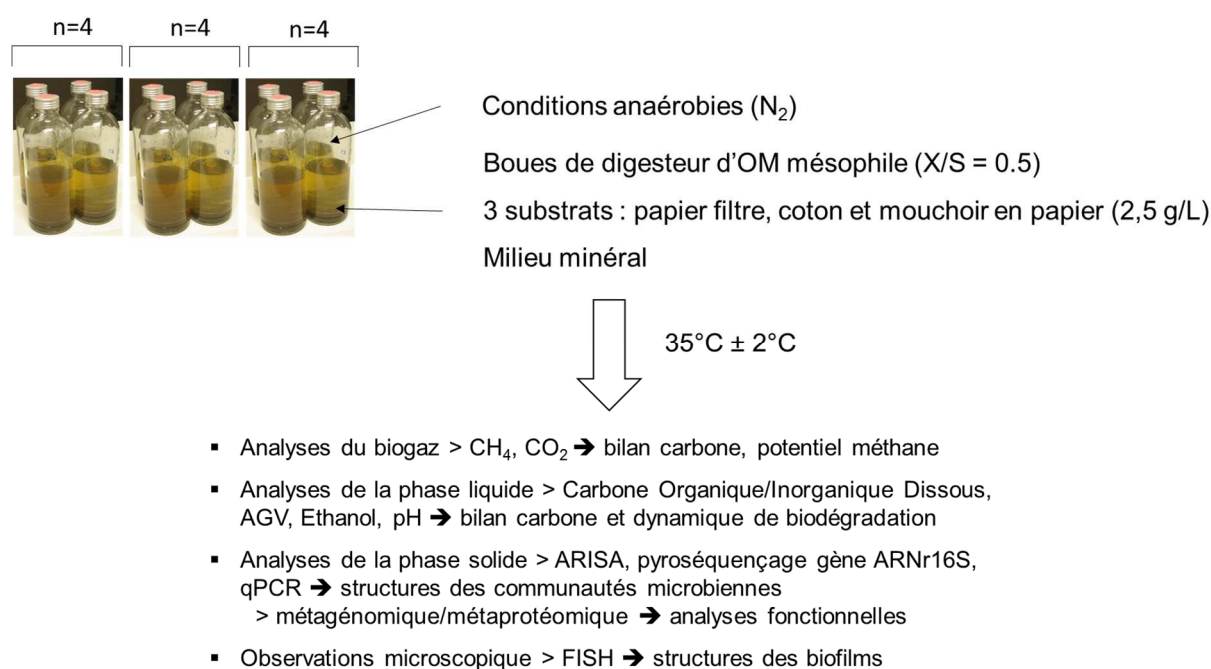


Figure 17. Dispositif expérimental réalisé pour l'étude de la dégradation de trois substrats cellulosiques par une biomasse complexe.

Les analyses réalisées sur ces incubations consistent en un suivi de la production et de la composition du biogaz afin principalement d'évaluer la bioconversion des substrats en méthane. Un suivi de la production d'AGV et de carbone organiques/inorganiques dissous est également réalisé afin d'identifier une possible accumulation d'AGV et de mesurer les quantités carbone encore non converti en biogaz. L'analyse du compartiment microbien est réalisée par ARISA, qPCR, pyroséquençage des gènes ARNr16S et FISH, afin d'avoir un aperçu de la structure des communautés microbiennes en lien avec les performances de dégradation obtenues. Plus précisément, les analyses de qPCR effectuées ont ciblé d'une part plusieurs populations taxonomiques, archées, bactéries, *Bacteroidetes* et *Firmicutes*. D'autre part, elles ont également ciblé des populations fonctionnelles, cellulolytiques, grâce aux gènes codant des cellulases possédant des domaines GH48, et hydrolytiques, grâce aux gènes codant des hydrolases à fer (gène *HydA*). Des analyses de métaprotéomique ont également été réalisées sur un point de temps de chaque condition pour identifier les fonctions exprimées par les communautés microbiennes. Afin d'améliorer l'identification des protéines détectées en métaprotéomique, les métagénomomes de deux échantillons provenant d'incubations de mouchoir en papier et de papier filtre ont été séquencés et analysés (Figure 18).

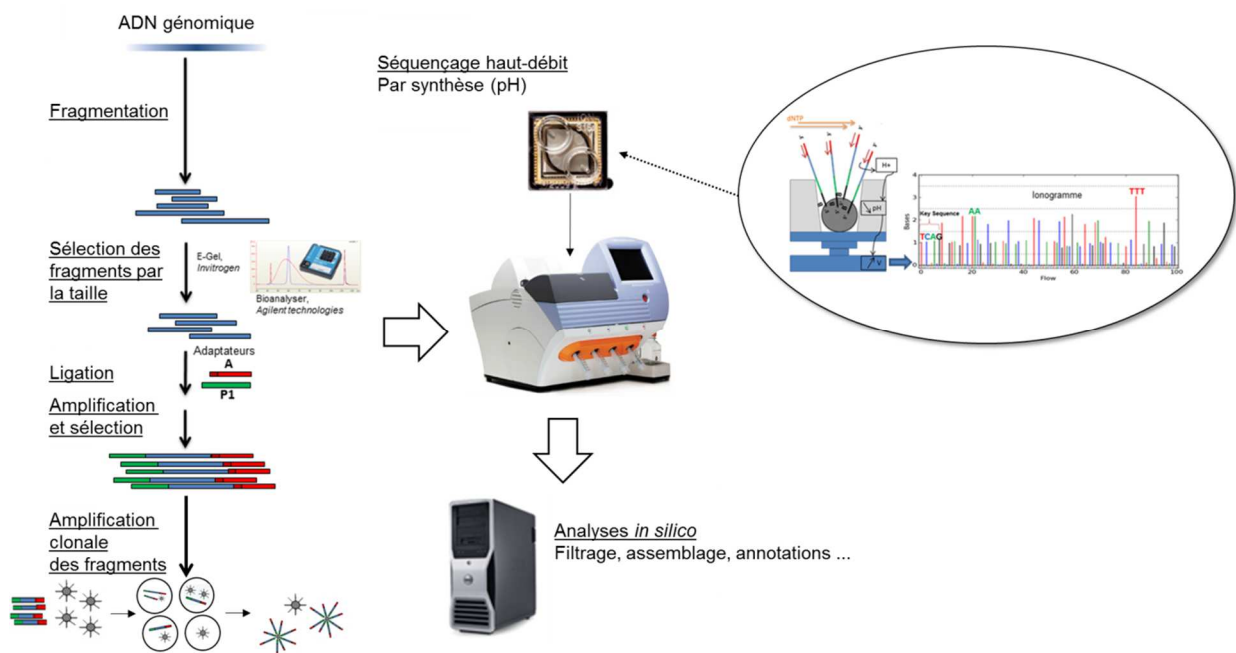


Figure 18. Aperçu du schéma expérimental pour l'analyse métagénomique.

L'analyse des métagénomomes est réalisée par l'utilisation de la suite d'instrument et des kits Ion Torrent pour le séquenceur PGM. Cela nécessite en premier lieu la création d'une librairie à partir de l'ADN extrait d'un échantillon. Cette étape consiste à fragmenter l'ADN, à liguer l'extrémité de chaque fragment à des séquences adaptatrices, connues et spécifiques, à sélectionner les fragments d'environ 400 paires de bases (la taille est déterminée en fonction du kit de séquençage utilisé) et à quantifier

ces fragments. L'étape suivante consiste en la préparation des fragments pour le séquençage. Les fragments se fixent, *via* l'une des séquences adaptatrices, à des billes, idéalement un par bille, et sont amplifiés pendant la PCR en émulsion (procédé de génération de microréacteurs de PCR dans de l'huile) jusqu'à l'obtention de billes portant un grand nombre de fragments clonaux. Une étape d'enrichissement permet ensuite d'augmenter la quantité de billes dont les fragments ont été correctement amplifiés, avant de réaliser le séquençage de ces fragments. Le séquençage est réalisé grâce à l'utilisation de puces spécifiques qui contiennent une grande densité de puits pouvant contenir au plus une bille. Lors du séquençage, l'intégration de chaque type de base est détectée séquentiellement grâce à l'émission d'un proton lors de l'intégration d'une base. L'ensemble de ces étapes est répété pour chaque échantillon.

Afin d'étudier plus finement les dynamiques de colonisation, des incubations en microplaques ont également été réalisées en parallèle, selon un schéma expérimental similaire à celui présenté dans le chapitre II. Concernant ces incubations réalisées en microplaques, les premiers résultats étaient prometteurs puisque durant les premiers jours d'incubations une dynamique de colonisation similaire à celle observée en souches pures est apparue (Figure 19). Cependant, ces résultats étaient peu reproductibles, surtout après 8 jours d'incubations et n'ont donc pas été inclus dans le manuscrit destiné à être publié. Cette variabilité pourrait être liée à la variabilité de structure des communautés microbiennes dans les différents puits de la microplaque, résultant probablement d'un effet d'échantillonnage lors de l'inoculation de chaque puit.

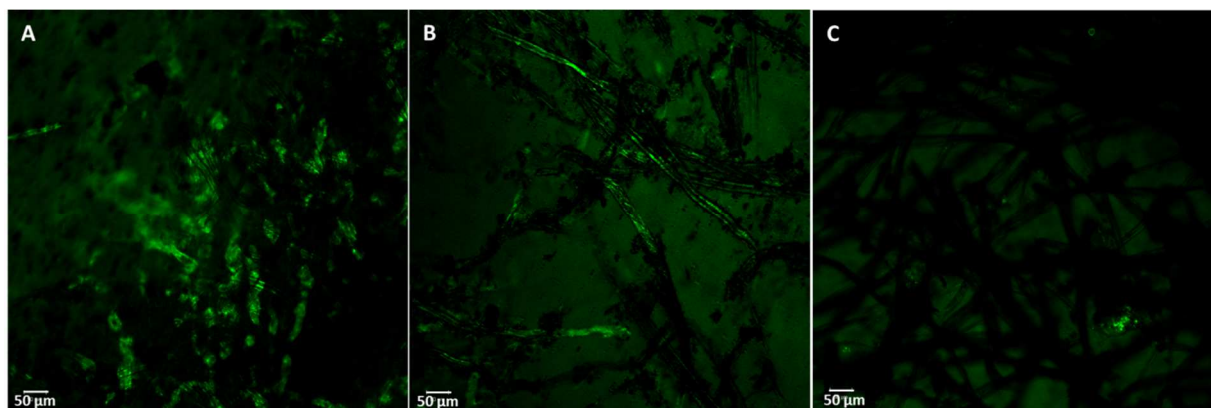


Figure 19. Photos de colonisation de mouchoir en papier, papier filtre et coton par une biomasse complexe obtenues après 8 jours d'incubations. Coloration ChemChrome V8 (Chemunex). A : Mouchoir en papier, B : Papier filtre, C : Coton en disques.

Les résultats obtenus sont présentés ci-après sous forme d'un manuscrit d'article qui sera encore retravaillé et étoffé avant soumission à un journal international à comité de lecture. En effet, la partie métagénomique correspond aux derniers développements effectués pendant la thèse. Dans sa

version actuelle, le manuscrit inclut une présentation synthétique et globale des résultats de métagénomique. Il sera complété par des analyses plus approfondies sur ces aspects-là, en particulier, le métagénome sera assemblé. L'assemblage du métagénome en grands contigs offre la possibilité d'explorer plus finement les gènes et voies métaboliques présents dans la communauté microbienne et couplé à la métagénomique pourrait favoriser une meilleure assignation des protéines identifiées par les analyses de spectrométrie de masse. Ces compléments d'analyses permettront de discuter plus en détails des résultats obtenus, notamment concernant le rôle des principaux groupes microbiens identifiés, et de les situer par rapport à ceux décrits dans la littérature.

III. Overview of methanization from distinct cellulosic materials based on comparative metaproteogenomics

Nelly Badalato¹, Alain Guillot², Chrystelle Bureau¹, Elie Desmond¹, Mathilde Lagesse¹, Angeline Guenne¹, Véronique Monnet², Ariane Bize¹

¹ Irstea, UR HBAN, F-92761 Antony, France

² INRA, UMR 1319 MICALIS, PAPPSO, 78352 Jouy-en-Josas cedex, France

Abstract

Municipal solid waste (MSW) biodegradable fraction is mainly composed of lignocellulosic materials in more developed countries, which can be exploited during anaerobic treatment processes. This renewable energetic material is however recalcitrant to degradation due to its insolubility and a highly complex and heterogeneous structure. The present study aims at assessing and understanding the effects of cellulose nature on its biodegradation and associated microbial population. Structural characteristics can vary according to the materials and may affect the accessibility and release of growth substrates. Experiments consisted in anaerobic digestion of three cellulosic materials – filter paper, paper tissue and cotton – closely related to MSW and already studied in a previous study with a model cellulolytic strain, *Clostridium cellulolyticum*. Each material was incubated in replicated mesophilic batch methanization microcosms, inoculated with a biomass originating from a solid-waste mesophilic methanizer. The main results showed that tissue paper was the most easily degradable material whereas cotton was the most recalcitrant, as it had been the case for the model system. Main results from Automated Ribosomal Intergenic Spacer Analysis (ARISA), 16S rDNA pyrosequencing analysis and preliminary results from high-throughput metaproteogenomics showed that substrate characteristics also led to the selection of specific microbial communities. In particular, it appeared that *Treponema sp.* and *Ruminococcaceae* members are quite abundant and active during filter paper and cotton biodegradation. By contrast, *Bacteroidia* members dominated the microbial community and were the more active during facial tissue biodegradation. Further analyses are required to fully understand their function within the community. This work brings new insights into the solid waste methanization processes by studying the effects of substrates' specific characteristics on the microbial populations, increasing the core knowledge required to develop more robust and efficient bioprocesses, more specifically optimized for abundant lignocellulosic waste.

Keywords

Anaerobic digestion, colonization, lignocellulose, metaproteomics, metagenomics

Introduction

Cellulosic materials are the major components of municipal solid waste (MSW) in more developed countries. It is widely accepted that cellulose hydrolysis is an important rate-limiting step during MSW anaerobic digestion. Many factors account for this recalcitrance such as accessible surface area, presence of other recalcitrant compounds such as lignin, degree of polymerization, porosity and cellulose crystallinity (Mansfield *et al.*, 1999). The precise importance of each of these factors is difficult to assess owing to the interdependence between these structural parameters (Lynd *et al.*, 2002). Cellulose degraders in anaerobic environment often use cell-bound complexed cellulase systems and are found in close contact with their substrates (Lynd *et al.*, 2002). In this context it is important to characterize the links between cellulose structure characteristics and biodegradation efficiency. Numerous studies have tackled this issue but only a few integrate analyses of the microbial compartment (Song *et al.*, 2005; Weimer *et al.*, 1990). Hence, more detailed examinations of the microbial compartment in addition to standard biodegradation studies can generate invaluable information to improve cellulose anaerobic biodegradation processes.

To better understand the links between substrate properties, biodegradation dynamics and microbial populations, the anaerobic digestion of three cellulosic manufactured materials – facial tissue, filter paper, cotton – was studied by an integrated approach combining biodegradation dynamics follow-up and microbial community characterizations. A strong focus on functional aspects was noticeably achieved through high-throughput metaproteogenomics analyses. Moreover, the cellulosic substrates used in the study were characterized in detail elsewhere (Badalato *et al.*, in preparation) and these previous results also constitute a reference for data interpretation.

Materials and methods

Reactor setup and analytical methods

Each of three cellulosic materials, cotton, filter paper and facial tissue, was digested in quadruplicate anaerobic batch microcosms incubated at $35\pm 2^\circ\text{C}$. The reactors consisted in 1 L glass bottles filled with 1.4 g dry matter of substrate, 660 mL of a mineral solution containing carbonate (0.14 M) and an inoculum. The stabilized inoculum originated from an anaerobic mesophilic solid-waste digester and was added to the reactors to a final inoculum-to-substrate ratio of 0.5 volatile solids. Biogas production was monitored by pressure measurements and biogas composition was analyzed with a gas chromatograph as described in Qu *et al.* (2009b). For liquid and solid phase analyses, fresh samples (2 mL) were collected through the septum and centrifuged at 10 000g for 10 min. The pH was measured on the liquid phase just after sampling. Pellets and supernatants were stored separately at -80°C until subsequent analyses. Dissolved organic carbon (DOC) and dissolved inorganic carbon (DIC) concentrations were measured in the supernatants with a BIORITECH 700 analyzer. Volatile fatty acid

(VFA) concentrations were measured using a DX 120 Ion Chromatograph (Dionex) with an IonPac ICE-AS1 column.

Carbon mass balance calculations

Carbon mass balance was calculated from total amounts of carbon present as DIC, DOC, CH₄ and CO₂ at each sampling time compared to the total initial carbon content. The total carbon removed by sampling was calculated from the measures of DIC, DOC and an estimation of cellulose concentrations. The remaining solid carbon included undigested cellulose, biomass and other undissolved carbon and was deduced by difference with the total initial carbon. Total initial carbon was a sum of carbon present at day 0 as DIC, DOC, CH₄, CO₂, cellulose (carbon content of 42 % for all the materials) and inoculum (carbon content of 24 %) based on analyses realized in Badalato *et al.* (in preparation).

Microbial community analyses

Total DNA was extracted from the pellets using the PowerSoil™ DNA isolation kit (MoBio Laboratories) according to the manufacturer's instructions. Bacterial fingerprinting analyses were realized by Automated Ribosomal Intergenic Spacer Analysis (ARISA). For the polymerase chain reaction (PCR), the ITSf/ITSr bacterial primer set (Cardinale *et al.*, 2004) was used to amplify the ITS region in the rRNA gene operon between the 16S and 23S rRNA genes. PCR conditions and amplified ITS fragment separation by capillary electrophoresis were realized as described in Fechner *et al.* (2010). Profile analyses were performed as described in Fechner *et al.* (2010), using the R package StatFingerprints (Michelland *et al.*, 2009).

For tag pyrosequencing analyses, the primer set 27F - 519R was chosen to target bacterial populations and amplify variable regions (V3-V4) of the 16S rRNA genes. Multiplexed bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP) was performed using the Titanium platform (Roche Applied Science) in a commercial facility (Research and Testing Laboratories, Lubbock, TX) and the results were analyzed as previously described in Lü *et al.* (2013) with the open source software package QIIME (Caporaso *et al.*, 2010). Most abundant OTUs sequences were identified using the EzTaxon-e server (<http://eztaxon-e.ezbiocloud.net/>; Kim *et al.*, 2012) on the basis of 16S rRNA gene sequence data. Rarefaction curves were calculated using the R package Vegan (Oksanen *et al.*, 2011).

Quantitative PCR targeting *Bacteria* and *Archaea* populations and *GH48* and *hydA* genes were realized as described in Carlei *et al.* (in preparation). Briefly, primer sets described in Schuurman *et al.* (2004) for *Bacteria* and in Gilbert *et al.* (Gilbert *et al.*, 2010) for *Archaea*, and in Pereyra *et al.* (2010) for *GH48* and *hydA* genes and iQ Sybr Green Supermix (Bio-Rad) were employed. The amplifications were performed in a CFX96 real-time thermal cycler (Bio-Rad).

For Fluorescent *in situ* hybridization analyses (FISH), 1 mL freshly-collected liquid samples were fixed with 4% paraformaldehyde immediately after collection and stored at -20°C in 50% ethanol in phosphate buffer saline 1X until further analyses. Fixed samples were hybridized following a standard protocol (Manz *et al.*, 1992). Combinations of specific or general probes were used: 5'-FITC EUB338mix/5'-Cy3 Arc915, targeting *Bacteria* (Amann *et al.*, 1990) and *Archaea* (Stahl, 1991) respectively, 5'-FITC BAC303/5'-Cy3 UCL284, targeting *Bacteroidetes* (Manz *et al.*, 1996) and *Ruminococcaceae* (Li *et al.*, 2009). Visualization of hybridized samples and image acquisition were realized using an inverted confocal laser scanning microscope (LSM510, Zeiss).

Metagenomic analyses

Metagenomes were obtained from DNA extracted from tissue sample (T 3) and paper sample (P1) at day 18. Sequencing was realized using the *Ion Torrent PGM* platform. All methods were realized according to the manufacturer's instructions. Libraries were prepared from total DNA using the Ion Xpress Plus Fragment Library Kit, template preparation were realized with the Ion OneTouch 400 Template v2 Kit and sequencing were performed using Ion PGM 400 Sequencing Kit and Ion 316 v2 Chip Kit. Quality filtered FastQ files were analyzed using the MG-RAST server (Meyer *et al.*, 2008) with default settings for analyses.

Metaproteomic analyses

For protein extraction, 50 mL freshly-collected incubation samples, were supplemented with 1 mM PMSF (Phenylmethylsulfonyl Fluoride) to inhibit protease activity and were centrifuged 20 min at 6000 g. The obtained pellets were immediately frozen in liquid nitrogen and stored at -80°C until subsequent analyses. Proteins from the pellets were extracted as described in Hanreich *et al.* (2012) with minor modifications. Briefly, cell disruption was performed by an ultrasonication step. Proteins were extracted using liquid phenol and precipitated with ammonium acetate in methanol. After several washing steps, the protein pellets were resuspended in buffer (7 M urea, 2 M thiourea, 4% (w/v) CHAPS) as described in Lü *et al.* (2013) and then stored at -80°C until subsequent analyses. A total of twelve protein samples were obtained (3 different substrates, 4 replicates, 1 time point). Protein profiles were visualized, and protein concentrations were measured, using an Agilent 2100 Bioanalyzer with a High Sensitivity Protein 250 kit following the manufacturer's instructions. Seven protein samples were selected for metaproteomics analyses (Table III-1). For each of them, 35 µg of purified proteins were fixed in SDS-PAGE gel (NuPAGE Bis-Tris Mini Gels, Invitrogen). After a short-duration migration, gel fragments (~5 mm) containing the proteins were excised. Washing and In-gel trypsin digestion were classically performed on each gel fraction.

Four μl of tryptic peptide liquid mixtures from each sample were analyzed by shotgun liquid chromatography-tandem mass spectrometry (LC-MS/MS). They were separated by nano chromatography on C18 column (50 cm, 75 μm i.d., Nanoseparation) by using the NanoLC-Ultra system (nano2DUltra, Eksigent). Buffers were 0.1% formic acid in water) and 0.1% formic acid in acetonitrile (buffer A. Peptide separation was achieved using a linear gradient of buffer A from 1% to 30% in 162 min and 30-40% in 10 min at a flow rate of 300 $\text{nl}\cdot\text{min}^{-1}$ for a total acquisition time of 188 min. Eluted peptides were analyzed on-line with a Q Exactive mass spectrometer (Thermo Electron) using a nanoelectrospray interface (1.4 kV ionisation potential) (non-coated capillary probe, 10 μm i.d.; New Objective). Peptide ions were analysed using Xcalibur2.07 with the following data-dependent acquisition steps: (1) full MS scan on a 300 to 1400 range of mass-to-charge ratio (m/z) with a resolution of 15000 and (2) MS/MS fragmented and detected in the linear trap (normalized collision energy: 35%, activation time 50 ms, isolation window of 2 th, dynamic exclusion time of 60 sec). Step 2 was repeated for the 10 major ions detected in step 1. Q Exactive raw files were converted to mzXML files with ProteoWizard 3.0.3706 (Kessner *et al.*, 2008). Protein identification was realized with X!Tandem software (X!Tandem CYCLONE 2010.12.01.1 [<http://www.thegpm.org>]) (Fenyő and Beavis, 2003) against Uniprot database (version January 2012, <http://uniprot.org>) associated to a classical contaminant database (trypsin, keratins, ...). The X!Tandem search parameters were as follows: trypsin specificity with one missed cleavage, fixed alkylation of cysteine, and variable oxidation of methionine. The mass tolerance was fixed to 10 ppm for precursor ions and 0.5 Da for fragment ions. For all proteins identified with a protein E-value of <0.01 in the first step, additional peptides were searched to reinforce identification, using similar parameters except that semitryptic peptides and protein N-terminal acetylations were accepted. All peptides identified with an E-value lower than <0.1 were conserved. A minimum of two different peptides with an E-value lower than 0.03 and a protein E-value (product of unique peptide E-values) lower than 10^{-3} were required for validation. The false discovery rate were estimated to 0.4% by decoy database search using the “include reverse” option of X!Tandem. Identified proteins were filtered by using the X!tandem pipeline (<http://pappso.inra.fr/bioinfo/xtandempipeline>). Protein databases were managed with the free public java software “Database Manager” (PAPPSO platform, INRA [<http://moulon.inra.fr/PAPPSO>]).

Results and discussion

General overview of the experimental design

The anaerobic digestion of each of three cellulosic materials used as sole carbon source was realized in quadruplicate batch anaerobic microcosms operated in mesophilic conditions (see Materials and Methods, Table III-1). The reactors were inoculated with a complex biomass prepared from a digestion sludge collected in a mesophilic methanization plant treating residual MSW and biowaste. The inoculum was previously characterized (Badalato *et al.*, in preparation), with 35.49/14.26 for TS (%)/VS (%) and a 23.6% total carbon content. The studied materials were facial tissue, filter paper and cotton; for convenience, they are referred to hereafter as tissue, paper and cotton, respectively. They were selected for their subtle differences in chemical and structural features which could result in contrasted biodegradation dynamics, as assessed by a previous study performed in a context of biodegradation by a pure cellulolytic bacterial strain (Badalato *et al.*, in preparation). Briefly, tissue is the most reactive substrate. It contains cellulose and hemicellulose and it is associated to the lowest degree of polymerization. Paper shows an intermediate reactivity and degree of polymerization, while cotton is the less reactive and has the higher degree of polymerization. Cotton also contains a large proportion of high molecular weight polymers and/or aggregates. Estimated crystallinity index does not appear as the main factor for the recalcitrance of these materials (Badalato *et al.*, in preparation).

The degradation and methanization dynamics were monitored by measuring the biogas production and composition, the DOC, DIC and VFA concentrations and the pH. Global microbial community changes were assessed by qPCR assays targeting wide taxonomic groups (*Archaea* and *Bacteria* domains) and by ARISA fingerprinting of the bacterial populations. To achieve a detailed functional insight in the bioprocesses, 16S rRNA gene tag pyrosequencing, metagenomics and coupled metaproteogenomics were applied for each substrate at incubation times corresponding to the well-established cellulolysis and to similar degradation stages: day 18 was selected for tissue and paper, while day 25 was selected for cotton to account for the slower degradation dynamics for this latter substrate (see next result section). Finally, targeted complementary functional analyses were performed, such as qPCR assays targeting functional gene families involved in cellulolysis (GH48) and hydrolysis (HydA). A few FISH observations targeting the taxonomic groups *Archaea*, *Bacteria*, or *Ruminococcaceae* were performed to visualize the biomass structuration around cellulose fibers. *Ruminococcaceae* family contains (not exclusively) widespread cellulolytic bacteria.

For clarity, the results are presented mainly by methodology in the next result sections.

Table III-1. Overview of the experimental design.

Replicates	Content	Sampling time points (days)							
		0	12	18	25	32	40	50	67
Inoc	Inoculum	(time points irrelevant) TS/VS, Carbon content, 16S rRNA							
T1-T4	Tissue	Physico-chemical analyses, qPCR, on all replicates							
		ARISA on all replicates							
				FISH (<i>Bacteroidetes- Ruminococcaceae</i>)					
				16S rRNA on T3, T4					
				Metagenomics on T3					
		Protein extraction and gel migration on all replicates Metaproteomics on T3, T4							
P1-P4	Paper	Physico-chemical analyses, ARISA, qPCR, on all replicates							
		ARISA on all replicates							
				FISH on P1 (<i>Archaea- Bacteria</i>)					
				16S rRNA on P1, P2, P3					
				Metagenomics on P1					
		Protein extraction and gel migration on all replicates Metaproteomics on P1, P2, P3							
C1-C4	Cotton	Physico-chemical analyses, qPCR, on all replicates							
		ARISA on all replicates							
					16S rRNA on C1, C3				
					Protein extraction and gel migration on all replicates Metaproteomics on C1, C3				

Inoc: inoculum – T: facial tissue – P: filter paper – C: cotton. TS/VS: Total solids/Volatile solids. 16S rRNA: 16S rRNA gene tag pyrosequencing. Physico-chemical analyses: measurements of biogas production and composition, of DOC, DIC and VFA concentrations and of pH.

Carbon mass balance over the course of degradation

To follow the distribution of carbon compounds and the kinetics of carbon release, the carbon mass balance was calculated at each sampling time (Fig. III-1). Results showed that carbon flux dynamics were partly different for each cellulosic material tested. For all 3 substrates, the degradation led to a transitory accumulation of DOC, mainly as acetate and also as propionate (data not shown). This was associated to a non-inhibitory pH decrease from 8.5 to 7.2 (data not shown), consistent with the use of a carbonate solution limiting pH fall. The methanogenesis started progressively after a lag phase of a dozen days, probably owing to low methanogen abundance at the beginning of the incubation time. Comparable final methane yields were observed and the final DOC levels were low, showing that the 3 materials tested probably have very similar energetic potentials (data not shown). However, distinct kinetic behaviors were associated to each substrate. Facial tissue was hydrolyzed faster as shown by the greater VFA accumulation and the faster methane production. By contrast, cotton degradation was the slowest with a limited DOC accumulation and a delayed methanogenesis phase. Tissue thus appeared as a more reactive substrate. Altogether, results indicate that the energetic potential was independent from the cellulose waste origin and properties.

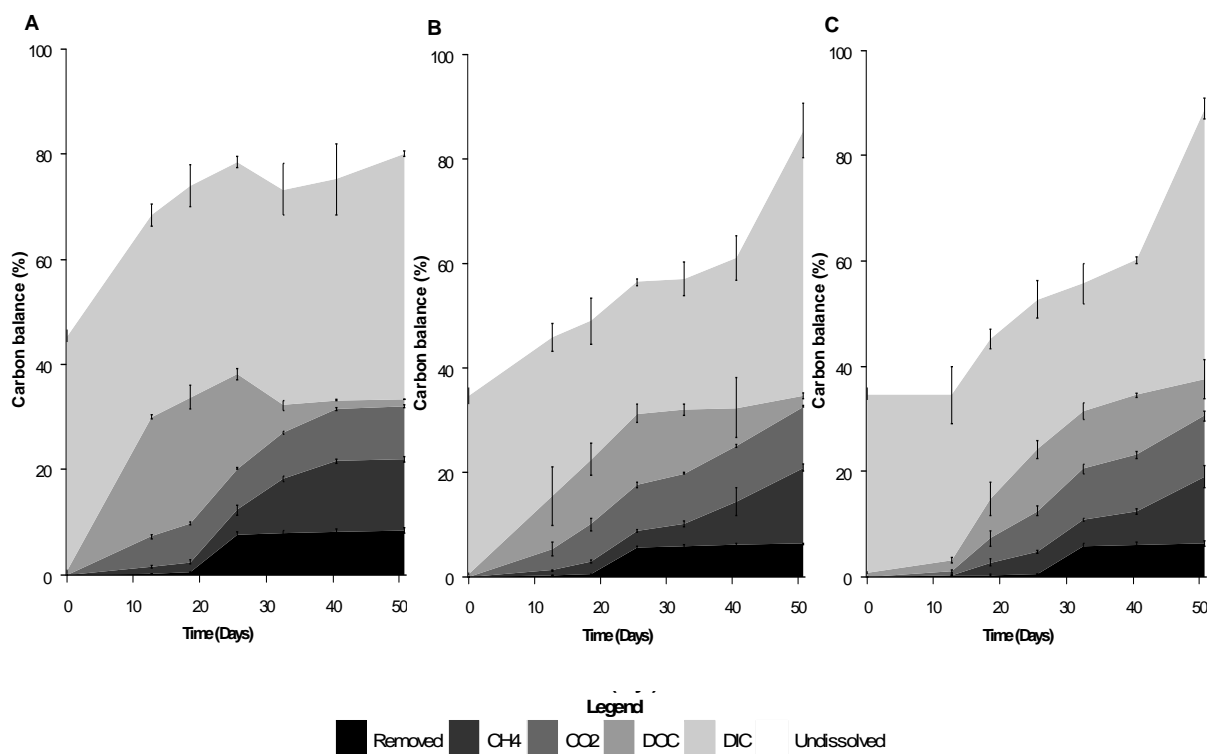


Fig. III-1 Carbon mass balances for tissue (A), paper (B) and cotton (C) incubations. At each time point, the total carbon is distributed among dissolved inorganic carbon and dissolved organic carbon in the liquid phase, carbon dioxide and methane in the accumulated biogas, total carbon removed by sampling and the estimated remaining solid carbon. A large portion of the dissolved inorganic carbon corresponds to carbonate in the mineral solution. Error bars are the standard deviations for the replicate samples.

Microbial community structures

Changes in bacterial community populations were analysed using the fingerprinting technique ARISA for the first phases of the digestion, at days 0, 12, 18, and 25 of the incubation time (Fig. III-2). Overall, a certain degree of variability was observed among the replicates, as shown by the presence/absence of bands or by variations in band intensity among some replicates (e.g. P2 and P3 profiles at day 18). General trends could however be discerned.

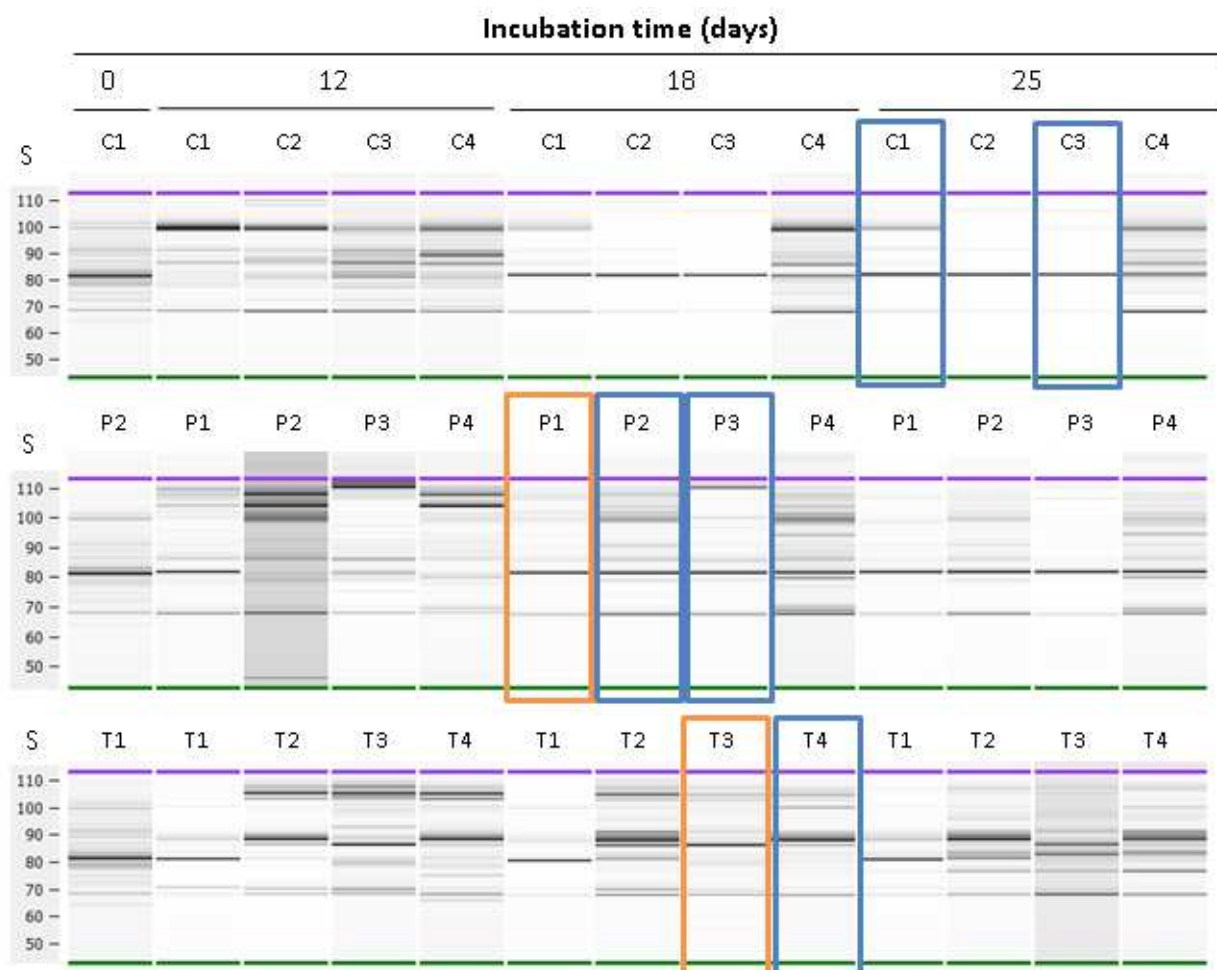


Fig. III-2. ARISA profiles obtained with a primer set targeting the bacteria.

The scale bars(S) indicate the migration time in seconds. Green and purple bands correspond to internal size markers. **Blue rectangles** indicate the replicates and time points for which 16rRNA gene pyrosequencing and metaproteomic analyses were performed. **Orange rectangles** indicate the replicates and time points for which metagenomic analyses were performed in addition to 16rRNA gene pyrosequencing and metaproteomics.

Two population shifts were visible based on the profiles, for most reactors: the first between days 0 and 12, the second between days 18 and 26 (tissue) or days 12 and 18 (paper, cotton). The first shift most probably corresponded to the establishment of the communities adapted to cellulose

degradation and to the operating conditions, and the second one could be related to population selection according to the level of substrate degradation. Except for the tissue, the variability and profile complexity seemed to decrease over time, possibly indicating the emergence of dominant stable functional groups. Finally, the profiles for cotton seemed to show an overall lower complexity than the other two substrates.

Hierarchical clustering analyses (data not shown) of the bacterial ARISA profiles at days of incubation showed that the replicates clustered as expected, except for C4 and T3, which suggests the presence of shared features in the bacterial populations, among replicates.

To identify abundant groups from the microbial communities, 16S rRNA gene tag pyrosequencing analyses were performed on the inoculum and on a selection of samples and replicates (Table III-1), with primer sets specific for bacteria. As previously described (section Overview of the experimental design), time points associated to similar substrate degradation levels and to a supposedly well-established cellulolysis activity were selected: two replicates from day 18 were analysed for tissue incubations; 3 replicates from day 18 were analysed for paper incubations; and 2 replicates from day 25 were analysed for cotton.

Rarefaction curves (not shown) and calculated diversity index (Table III-2) suggested that an important fraction of the sample diversity had been covered, sufficient to identify some major bacterial groups (Observed OTUs/Chao1 > 50% for all of the samples). Nevertheless, a non-negligible fraction of the diversity probably remained uncharacterized. Moreover, PCR biases causing artefacts and distortions of the results can of course never be ruled out, such that these tag pyrosequencing results should be interpreted with the usual caution.

Table III-2. Number of identified OTUs and diversity index

	Observed		
	OTUs	Chao1	1/Simpson
Inoculum	308	365	15.86
T4	75	123	5.05
T3	105	156	2.04
P1	156	205	15.32
P2	97	148	8.32
P3	63	103	2.06
C1	82	100	4.08
C3	82	107	6.15

From 63 to 308 OTUs were identified according to the samples (Table III-2). Based on the sequencing results, the bacterial population from the inoculum was more diverse than the bacterial populations from the reactors. This trend could result from the combined effects of the inoculum sampling at the inoculation step and the strong selection of bacteria adapted to the operating conditions during the incubations. The Venn diagrams showed that an important fraction of the OTUs was identified in all three conditions (core OTUs, ~17%) and that a non-less important fraction of the identified OTUs was detected only for one type of substrate (e.g. tissue, ~17%) (data not shown). This observation is not surprising in the context of bioprocesses based on complex microbial communities (e.g. Rivière *et al.*, 2009).

The distinct composition structure of the inoculum compared the other samples was consistent with a strong selection occurring during the initial stages of the incubation, as also observed by ARISA. By contrast to the observations at the OTU level, most of the identified dominant groups were shared among the samples (Fig. III-3) at the class and genus taxonomic levels. The composition profiles were rather similar among replicates, except for P3 which showed a very different structure compared to P1 and P2.

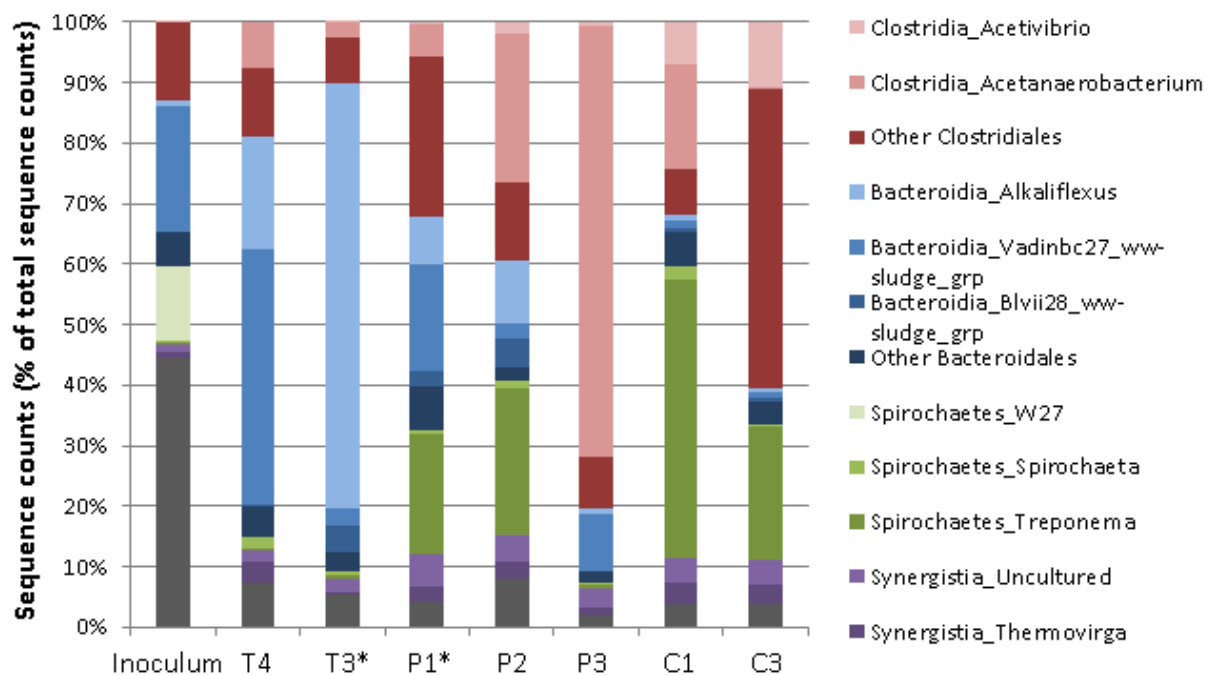


Fig. III-3. Taxonomic distributions obtained by 16S rRNA gene tag pyrosequencing, at the genus level. While metaproteomics data were generated for all these corresponding samples, * indicate the samples that were used for metagenome sequencing (T3, P1).

In more details, at the class level, *Bacteroidia* members were highly dominant in the paper reactor samples, followed by members of the class *Clostridia*. By contrast, three main groups were present in similar proportions in the paper reactor samples (P1, P2): *Clostridia*, *Bacteroidia* and *Spirochaetes*. For P3, however, sequence reads were by far mainly attributed to the class *Clostridia*, and *Spirochaetes* were detected at a very low level. For cotton reactor samples, finally, the two dominant classes were *Spirochaetes* and *Clostridia*. Members of the class *Synergistia* were identified in all samples, although in limited proportions (~2-7%). They seemed more abundant in the paper and cotton reactor samples than in the tissue reactor samples.

At the genus level, in a similar way, the number of dominant genera appeared as limited. The proportion of the most abundant genera within each class showed a rather high variability among the replicates of each substrate. A noticeable exception concerned the genus *Treponema*, since *Spirochaetes* members, when abundant, were almost exclusively assigned to the genus *Treponema* in the incubation samples (P1, P2, C1, C3). Within *Bacteroidetes*, dominant genera were *Alkaliflexus* and *Vadinbc27* for the tissue reactor samples, and *Alkaliflexus*, *Vadinbc27* and *Blvii28* for the paper reactor samples (P1, P2). Within the class *Clostridia*, *Acetanaerobacterium* and other *Clostridiales* were the most represented for the tissue and paper reactor samples, while *Acetivibrio* was abundant in the cotton reactor samples, in addition to *Acetanaerobacterium* and other *Clostridiales*.

ARISA and 16S rRNA gene tag pyrosequencing results interestingly suggest that distinct microbial groups and distinct community assemblages could have emerged in relation with the specific chemical and structural features of the cellulosic materials used as substrates. Cotton seems more associated to *Treponema* genus, various members of the *Clostridiales* order and *Acetivibrio* genus. Tissue seems more associated to high proportions of *Alkaliflexus* and *Vadinbc27*. Paper shows a somewhat mixed pattern, with 3 dominant taxonomic groups detected for this substrate (*Treponema*, several *Bacteroidia* and *Clostridia* genera).

The underlying selection mechanisms resulting in these distinct assemblages could encompass: the recruitment of distinct cellulolytic strain according to the substrate's features; the modification of the carbon flows entering the cellulolytic bacteria, due to the variable degree of recalcitrance of the substrates; the modification of amount and proportions of fermentation products released by the cellulolytic bacteria *C. cellulolyticum* (Badalato *et al.*, in preparation); the impact of these modified fluxes on the microorganisms downstream of the degradation chain; direct interactions or compatibility/incompatibility networks among microorganisms, for instance involving cellulolytic bacteria.

The community structure analyses also showed the great extent of biological variability in these ecosystems. The case of the paper, sole substrate for which 3 replicates were chosen for pyrosequencing, suggest that functional redundancy could lead to the emergence of several different community networks associated to comparable degradation efficiencies. It highlights the importance of increasing the number of replicates to reach sufficient statistical power.

Metagenomic analyses for filter paper and facial tissue

The metagenomic analyses were primarily employed to create a database sufficiently specific for the samples, during the database search and protein identification of the metaproteomics approaches. A synthetic overview of the metagenomic results is presented below.

Table III-3. MG-RAST metagenome analysis indicators.

	Tissue (T3)	Paper (P1)
Sequences Count	3,098,332	2,927,077
Mean Sequence Length	249 ± 94 bp	197 ± 78 bp
Mean GC percent	46 ± 8 %	45 ± 8 %
Post QC: Sequences Count	2,663,234	2,636,623
Identified Protein Features	400,074	490,054
Identified rRNA Features	1,946	2,420

QC: quality control

Two DNA samples (from T3 and P1 at day 18) were used for metagenome sequencing with *an Ion Torrent PGM* sequencer. The obtained reads were analysed using the MG-RAST server (Meyer *et al.*, 2008) and results are shown in Table III-3. A total of roughly 400,000 and 500,000 protein features were obtained for T3 and P1 respectively. In addition, roughly 2000 and 2400 rRNA features were identified for T3 and P1 respectively. The protein features were used to construct the database for metaproteomics (see next section).

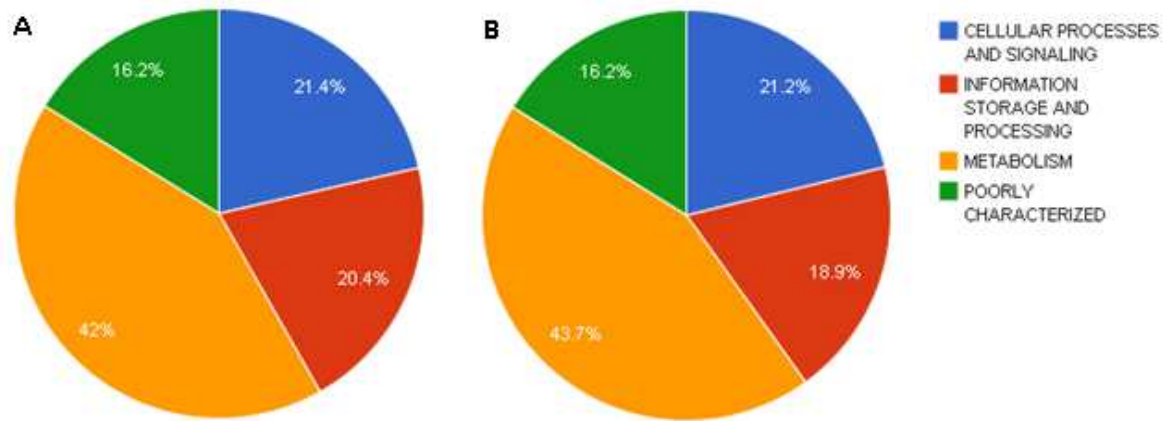


Fig. III-4. COG repartition of the annotated proteins based on the MG-RAST metagenome analysis results. A: tissue (T3). B: paper (P1).

Concerning the functional aspects, COG repartitions of the annotated metagenome proteins were highly similar between T3 and P1 (Fig. III-4). Functions related to metabolism appeared as dominant, followed by cellular processes and signalling, information storage and processing, and then poorly characterized functions. In more details (data not shown), it seems that the identified functions covered a wide range of COG functions and were rather evenly distributed among the different categories. Unsurprisingly for microbial communities comprising mainly bacteria and archaea, the less represented categories included RNA processing and modification, chromatin structure and dynamics, extracellular structures, cytoskeleton.

Concerning the taxonomic composition of the metagenome, several important trends were consistently retrieved when comparing these results with the 16S rRNA gene tag pyrosequencing analyses (Fig. III-5). For T3, *Bacteroidetes* was strongly dominant (63.9%). For P1, *Firmicutes* and *Bacteroidetes* were the most represented (69.6% in total); *Spirochaetes* were also an abundant group, although to a lesser extent (7.2%). Interestingly, significant discrepancies also appeared, mainly the significant proportion of *Proteobacteria* sequences (10.3% and 9.8% for T3 and P1 respectively), which was not the case for 16S rRNA gene tag pyrosequencing analyses. These results could be explained by several factors, such as primer set specificity properties and different taxonomic references used in MG-RAST versus Qiime.

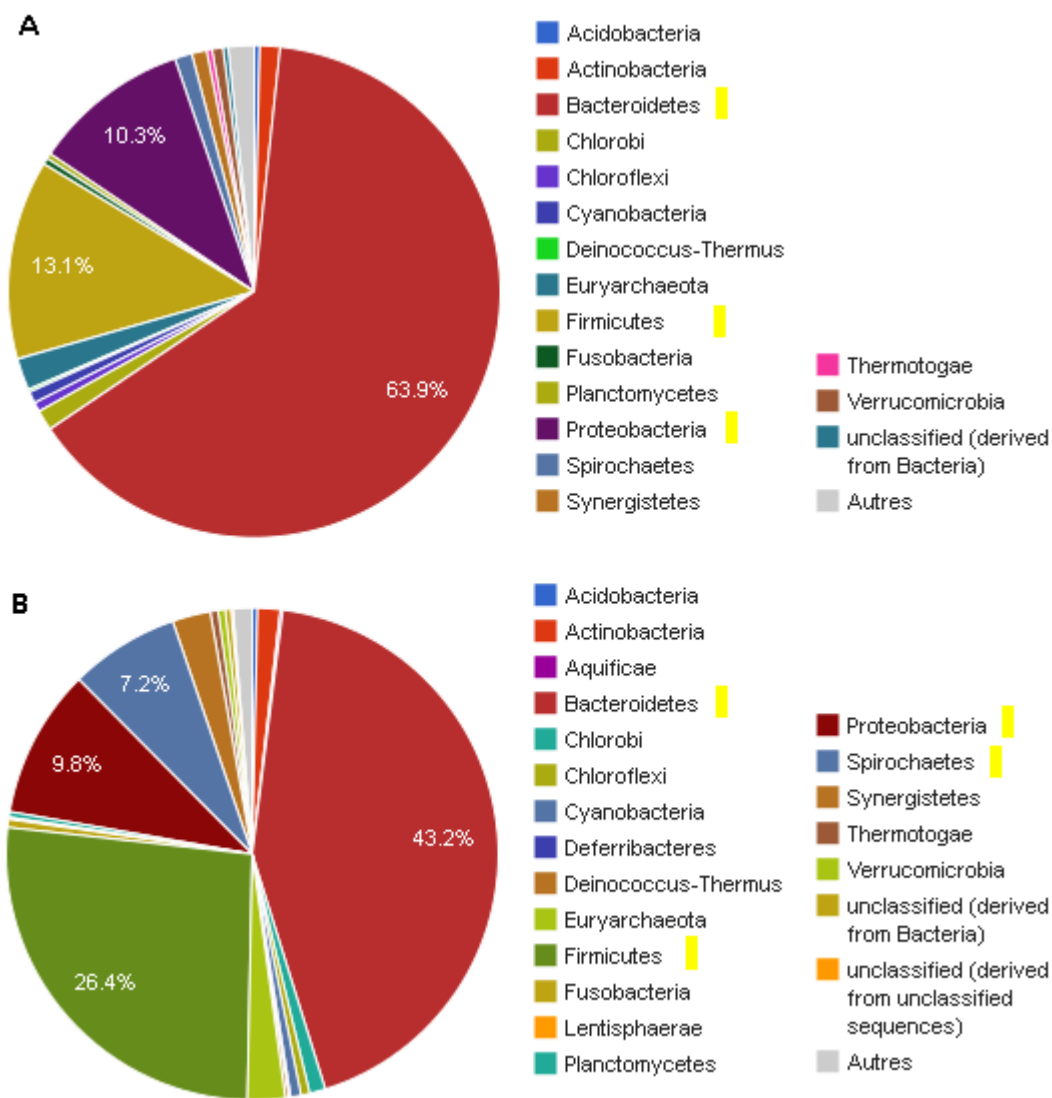


Fig. III-5. Taxonomic composition based on the MG-RAST metagenome analysis results (Identified proteins). A: tissue (T3). B: paper (P1).

To conclude this section, two low-coverage (*Ion torrent PGM* is a bench top sequencer) but good quality metagenomes were obtained, showing overall homogeneous function distributions according to the COG classification, and showing taxonomic distributions (based on identified rRNA sequences) in overall good agreement with the 16S rRNA gene analyses. Using these data, a protein database comprising 89 0128 protein sequences was to help identify dominant components from the metaproteomes.

Coupled metaproteogenomics analyses

Before performing metaproteomics analyses, protein profiles were generated for all the replicates for the tissue and paper reactors at day 18 and for the cotton reactor at day 25 (Fig. III-6). These profiles were based on protein extracted from pellets of liquid samples and they indicated the good quality and the sufficient amount of extracted proteins. Similar to the profiles obtained by ARISA fingerprinting, these profiles suggested a significant level of biological variability among the replicates, here especially for the cotton and paper reactors. The 8 protein samples corresponding to the same replicates and time points as for the 16S rRNA gene pyrosequencing analyses were used for metaproteomic analyses.

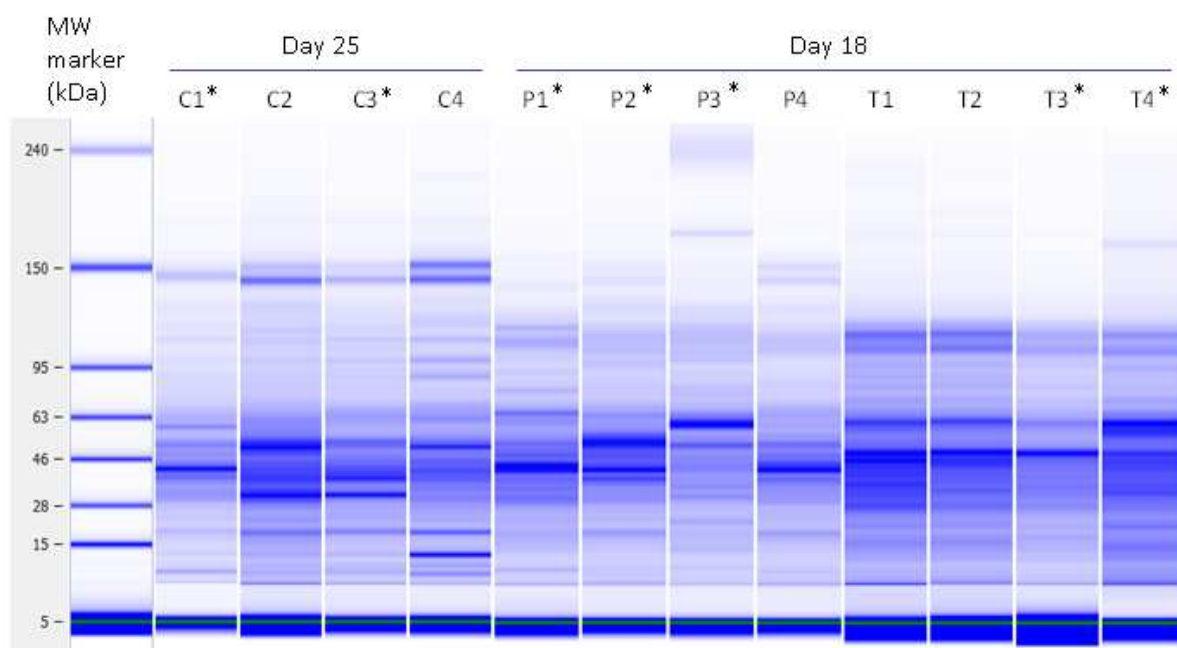


Fig. III-6. Molecular weight profiles of the proteins extracted from sample pellets. Results are shown for all the replicates of the considered incubation time point. The separation was realized on a lab-on-a-chip equipment. * indicate the samples used for metaproteomics analyses.

Over 18 000 MS/MS spectra were generated for each sample on average (Table III-4). A first step of the analysis was the evaluation of the relevance and interest of the metagenomic database for protein identification. The percentages of spectra assigned to a peptide from the database were compared when using a reference public database, UniprotKB (right part of the Table III-4) and the database generated from the metagenomic data (left part of the Table III-4). This approach clearly demonstrated the interest of relying on the metagenome, specific database, since the average percentage of assigned spectra was roughly multiplied by 5 and reached values closer to those usually obtained on pure model strains (~25%, Beganovic *et al*, 2010).

Table III-4. Comparison of protein identified with two different databases

Samples	DB : metagenomic protein sequences			DB: UniprotKB		
	Spectra used	Spectra assigned	% assignation	Spectra used	Spectra assigned	% assignation
C3	24032	4001	16.6%	24032	968	4.0%
C1	21679	2930	13.5%	21679	595	2.7%
T4	17785	1828	10.3%	17785	341	1.9%
T3	11090	612	5.5%	11090	80	0.7%
P3	17022	2014	11.8%	17022	358	2.1%
P1	15247	1785	11.7%	15247	279	1.8%
P2	22254	3486	15.7%	22254	584	2.6%
Mean	18444	2379	12.2%	18444	458	2.3%

In addition, the classification of the database into COG categories was determined. Information is shown in Table III-5. Within the COG category “metabolism”, 56714 proteins (8% of the total database) corresponded to the category “Carbohydrate metabolism”, especially relevant in the context of this study. It could thus be expected that the database was suitable to identify relevant protein function with respect to the studied bioprocess.

Table III-5. COG classification of the sequences from the metagenomic database used for protein identification.

COG	Count	%
Information storage and processing	135515	19.1156432
Cellular processes and signaling	172278	24.3014041
Metabolism	400651	56.5155264
Poorly characterized	478	0.06742632
Total cog	708922	100

The sequence were compared to the COG using a maximum e-value of $1e^{-15}$, a minimum identity of 60 %, and a minimum alignment length of 15 measured in aa for protein and bp for RNA databases.

Before developing detailed functional analyses of the identified proteins, the proteins identified among reactor types and replicates were compared, as shown by the Venn diagrams (Fig. III-7). These results mainly showed that the proportion of shared identified proteins was rather limited, which could results, according to the cases, from the biological variability, from distinct functions activated by each substrate type, and most probably, from the limited coverage obtained for the complex metaproteomes (as suggested by the preliminary estimate of rarefaction curves, not shown here) which could increase the influence of sampling factor effects. Following this observations, the results

from replicates were pooled to generate 3 datasets of identified proteins corresponding respectively to tissue, cotton and paper incubation reactors (Fig. III-7 panel A). Surprisingly, more proteins were identified for cotton incubation samples than for tissue incubation samples whereas no specific metagenomic dataset had been generated for cotton reactors (see Result section on Metagenomic analyses). A similar trend was observed at the replicate level for tissue, for which more proteins were identified for T4 whereas metagenomics had been performed on T3 (Fig. III-7 panel D), as well as for paper, for which more proteins were identified for P2, whereas metagenomic datasets had been produced for P1 (Fig. III-7 panel B). Among cotton, paper and tissue reactors, the highest number of identified proteins was obtained for the paper samples, consistent with the higher number of replicates analyzed for this substrate compared to the other two (3 replicates *versus* 2).

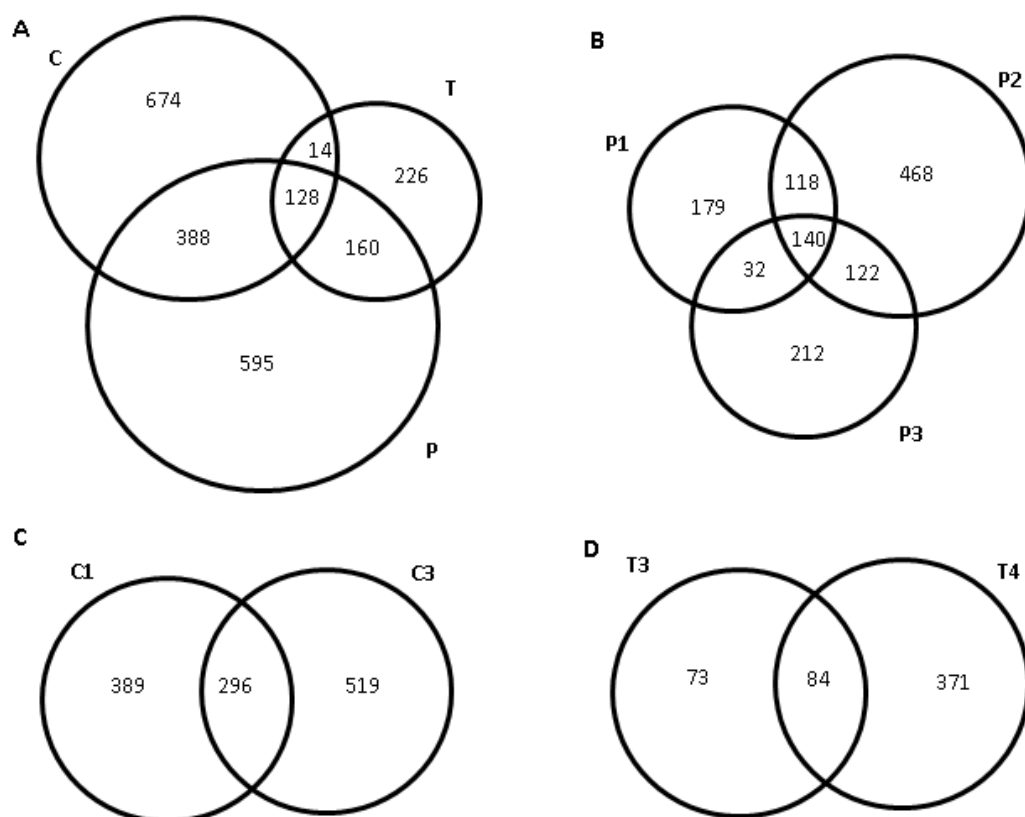


Fig. III-7. Venn diagrams of the identified proteins for each substrate type (panel A, results from the replicates were pooled) and for each replicate (panels B to D). The numbers of non-redundant protein functions are indicated.

To realize functional analyses of the identified proteins, it was required to obtain precise functional annotations. These data were extracted from the MG-RAST pipeline results and the RefSeq database. Among the identified proteins, only a certain proportions of corresponding reads returned a positive BLAT result, indicative of the presence of proteins with similar sequences in the public databases. This was expected since metagenome and newly-sequenced genome almost systematically comprise a

significant proportion of previously unknown genes (ORFans). From here, the annotation of all the proteins with a positive BLAT result should have been retrieved. However, many annotated proteins were finally missing (Fig. III-8 Bars from the right side compared to bars from the middle). This highlighted the necessity to design a more refined annotation procedure. This can be achieved by combining multiple available databases (GenBank, M5NR ...). However, in order to obtain first meaningful results at the functional level, the proteins in this first approach were used for a first preliminary analysis and some illustrative result elements are presented below. Therefore, these results should be examined with caution since half of the retrieved proteins were not analysed. Besides, information about taxonomic affiliation of retrieved proteins should be interpreted with caution since protein assignment was based on BLAT results from DNA sequences.

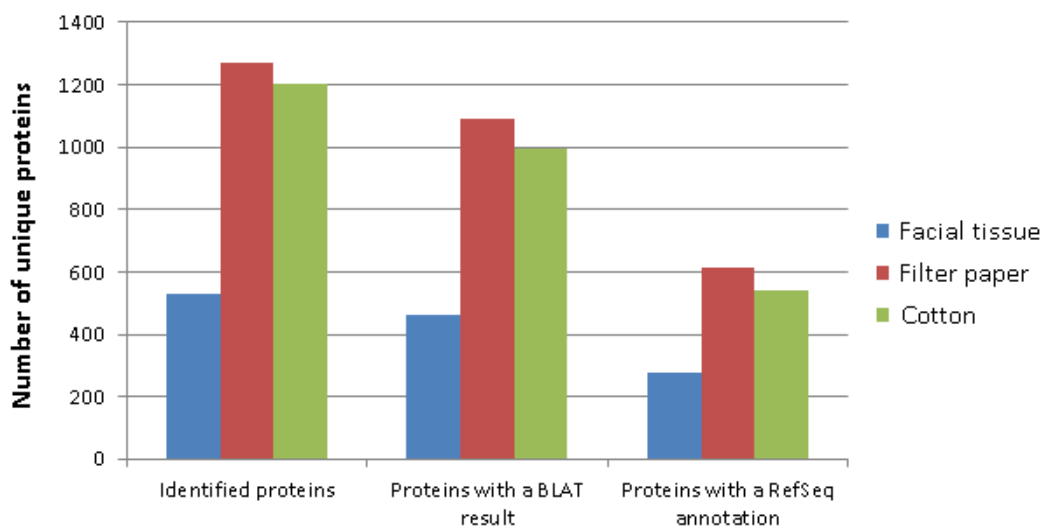


Fig. III-8. Number of non redundant proteins from each condition retrieved at each step along the MG-RAST pipeline.

First examination of the annotated protein dataset, showed that most of the (non redundant) proteins retrieved were assigned to bacteria belonging to the *Spirochaetes* class and the *Ruminococcaceae* (*Clostridia*) family for paper and cotton (Fig. III-9). Regarding the *Ruminococcaceae* family this result is not surprising since many isolated mesophilic cellulolytic bacteria belong to this family (Ludwig *et al.*, 2007) and could be active during cellulose digestion. However, a contrasted result was observed for tissue, since proteins assigned to bacteria belonging to the *Ruminococcaceae* family were much less numerous. This observation is in good agreement with results obtained from 16S rRNA gene pyrosequencing analysis, showing that members of this group were not abundant nor very active. Since the 16S rRNA gene pyrosequencing analysis did not clearly indicated the presence of other possible cellulolytic bacteria (such as bacteria belonging to the genus *Fibrobacteria*), a possible explanation is that cellulolytic bacteria, belonging to the *Ruminococcaceae* family or not, are no longer present and

active at this time point, according to the stage of substrate degradation. Instead, proteins assigned to bacteria belonging to the genus *Anaerophaga* (*Bacteroidetes*), which are possibly fermentative bacteria (Denger and Warthmann, 2002), were quite numerous. Another possibility is the cellulolytic microorganisms for facial tissue correspond to new or poorly described species.

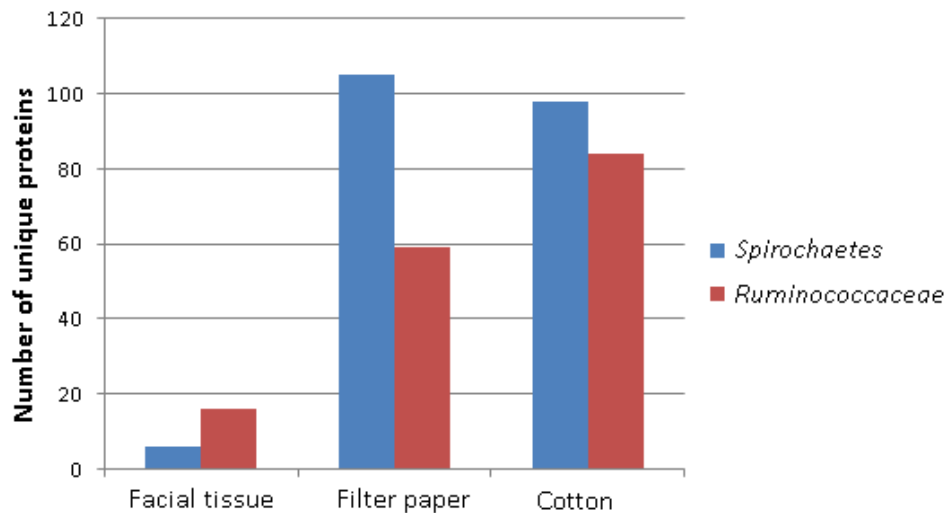


Fig. III-9. Number of non redundant proteins from each condition related to the *Spirochaetes* class and the *Ruminococcaceae* family

The high proportion of proteins assigned to bacteria belonging to the *Spirochaetes* class, and more specifically to the *Treponema* genus, for paper and cotton is quite interesting since this group was already shown as involved in lignocellulose digestion (Warnecke *et al.*, 2007) (data not shown). Indeed, most of these proteins were assigned to *Treponema brennaborensis*. This bacterium was not extensively described but appears as highly motile (Schrank *et al.*, 1999). A similar conclusion also emerged from this dataset since a lot of their identified proteins were related to cell motility, such as chemotaxis and flagellar assembly (data not shown).

To further examine the precise function of *Treponema sp.* and other groups in cellulose bioconversion, more analyses are required. These analyses include, for example, identification of carbohydrate active enzymes (CAZymes) to directly highlight cellulose deconstruction functions, since a lot of CAZymes are already described for bacteria such as *T. brennaborensis*. This can be achieved by database searching, using amino acids sequences of the retrieved proteins, against the CAZy database (Cantarel *et al.*, 2009; Yin *et al.*, 2012).

Targeted functional analyses

To complement or support the first elements obtained by the metaproteogenomic approach as well as by the follow-up of the methanization kinetics through physico-chemical measurements and general microbial community analyses, specific targeted functional analyses were performed.

A few FISH analyzes were performed to observe the microbial biomass colonizing cellulose fibers within the samples, using different probe sets. Overall, the observations showed the intense level of bacterial fiber colonization for paper and tissue (Fig. III-10, right). On tissue, they demonstrated that *Ruminococcaceae* members were abundant at the fiber surface and that *Bacteroidetes*, on the other hand, were not dominant at the fiber surface (Fig. III-10, left). This observation is quite surprising since this group was not revealed as highly abundant nor very active, as shown by 16S rRNA gene pyrosequencing and metaproteogenomic analyses. One possible explanation is that *Ruminococcaceae* members were abundant only at the fiber surface. At this time point (day 18), most of the tissue was digested, as shown by the hydrolysis efficiency (see Carbon mass balance over the course of degradation section) compared to paper and cotton. Thus available fibers could be less abundant and more abundant and active bacteria than *Ruminococcaceae* members were retrieved instead.

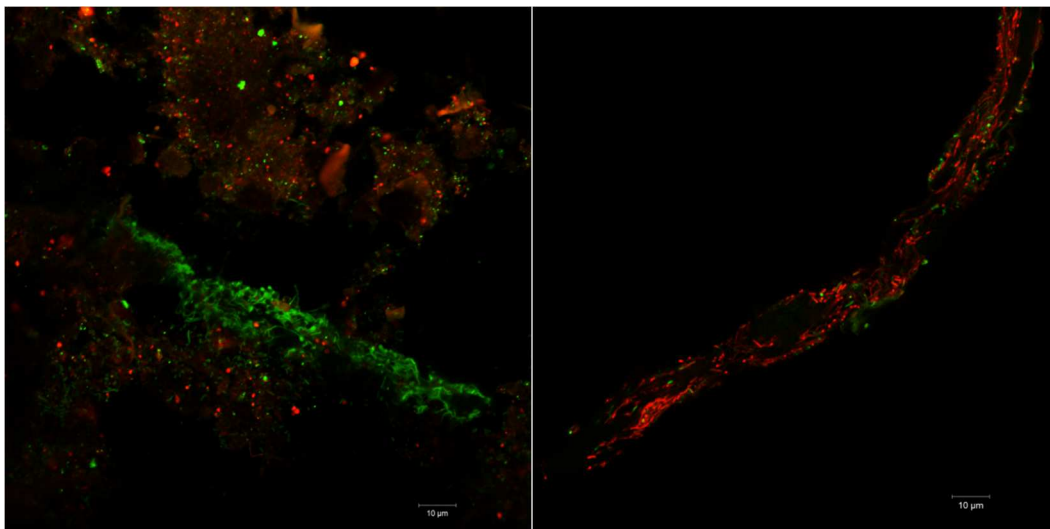


Fig. III-10. FISH-CLSM biofilm images from filter paper (left) and facial tissue (right) incubations at day 18.

Left : The hybridized *Archaea* are indicated in red (probe Cy3-Arc915) and the hybridized *Bacteria* are indicated in green (probe FITC-EUB mix). Right : The hybridized *Ruminococcaceae* are indicated in red (probe Cy3-UCL284) and the hybridized *Bacteroidetes* are indicated in green (probe FITC-BAC303)

To quantify the evolution of archaea, bacteria and specific functional genes, quantitative PCR assays were performed on the extracted DNA (Fig. III-11). Concerning quantitative PCR on complex microbial communities, it must be noted that the level of primer specificity for analysis of a given sample is generally unknown, since every microbial community is different and its composition is not always fully-known. Moreover, different gene copy numbers can occur among different strains. Precise quantifications are thus delicate and in the present work, quantitative PCR was mainly employed to suggest trends.

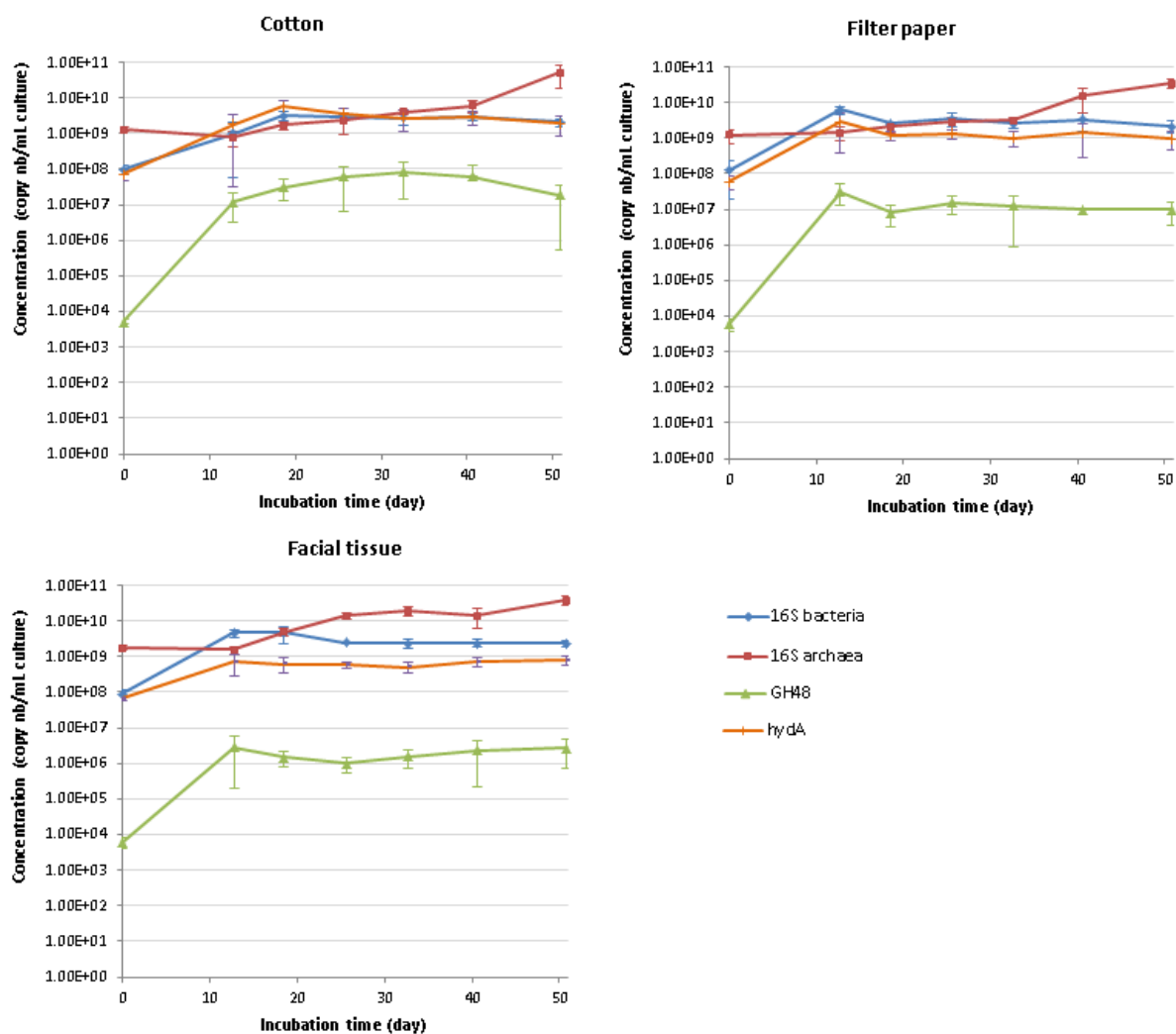


Fig. III-11. Quantification from the extracted DNAs of genes related to taxonomy or to specific gene functions.

Archaea and bacteria were detected at similar ranges for all the substrates ($\sim 10^8$ - 10^{10} copies/mL culture). However, distinct evolution trends were observed between bacteria and archaea: for bacteria, the detected levels at the beginning of the incubation time were of $\sim 10^8$ copies/mL culture in all conditions (consistent with every reactor being inoculated with the same amount of biomass). This initial levels were roughly multiplied by a factor 10 to 100 at the second sampling time point (day 12). Thereafter, the levels slightly decreased over time. It suggests that adapted bacterial communities rapidly develop and that their growth is then limited, possibly due to limiting hydrolysis rate and/or to substrate exhaustion over time.

Concerning archaea, the detected levels at day 0 were 10 times higher than for bacteria. This result is quite unexpected since archeas are generally found as the limited population in similar batch incubations (Lü *et al.*, 2013) and could be due to the primers sets used for quantitation. According to the substrate, these levels remained stable or slightly increased during the first incubation phase. A steeper increase was observed at later stages of the incubation. This general pattern could originate from the slower growth rate of methanogenic archaea compared to many fermentative bacteria. Interestingly, the steeper increase started at earlier stages for tissue (from day 12 onwards) compared to the other substrates (from day 32 onwards for paper and from day 40 onwards for cotton): this observation is in good agreement with the levels of reactivity of the substrates limiting the overall methanization process.

Concerning functional genes, *HydA*, a general marker of hydrolysis activity was targeted, and the detected levels were almost identical to those detected for bacteria: parallel evolutions at similar copy number levels were observed. This is consistent with the hydrolysis being a very general and widespread function, thus strongly linked to the total amount of bacteria.

A second functional assay, targeting genes coding for enzymes harboring a GH48 domain, was performed (Fig. III-11). This family comprises biochemically confirmed cellulases and are considered as key component of various cellulolytic systems (Blouzard *et al.*, 2010; Sukharnikov *et al.*, 2012) This gene family is thus more specific for the cellulolysis function compared to *HydA*. The detected levels for GH48 followed parallel time evolutions compared to bacteria, from day 12 onwards. However, compared to bacteria, the detected levels were lower by a factor 100 or less for cotton and paper, and by a factor 1000 for tissue. At day 0, the difference between the levels detected for GH48 and for bacteria was more pronounced, of a factor 10^4 for all substrates (same amount of inoculum in every reactor). Taken together, these observations interestingly suggest that the abundance of cellulolytic microorganisms is limited within the studied microbial communities, and that they are strongly selected during the initial stages of the incubation, the ratio of detected level for GH48 over detected

level for bacteria being multiplied by 10 to 100 between day 0 and day 12. Besides, the overall lower GH48 levels were detected for tissue incubation samples compared to cotton and paper. Interestingly, most cellulolytic bacteria belonging to *Firmicutes* possesses GH48 cellulases (for example *Clostridium cellulolyticum* and *Ruminococcus albus*) whereas other families of cellulases, such as GH5 or GH9 enzymes, are more widely distributed among bacteria (for example *Fibrobacter succinogenes*) (Cantarel *et al.*, 2009). This result is in good agreement with the 16S rRNA gene tag pyrosequencing analysis which showed that the bacterial community in the paper and cotton reactors was dominated by the *Clostridia* group compared to the tissue reactor. Therefore, it could suggest that tissue degradation involved more cellulases belonging to other families, compared to paper and cotton, and thus bacteria possessing GH48 enzymes are not as strongly selected. Further examination of CAZymes are needed to identify these cellulases (Yin *et al.*, 2012).

Conclusions

Methanization of three cellulosic materials – filter paper, paper tissue and cotton – closely related to MSW, and associated microbial population were studied. The main results showed that facial tissue appeared as the more reactive substrate compared to filter paper and cotton, which is consistent with the results obtained on the model system with *C. cellulolyticum* previously investigated. For functional analysis of the microbial community, the metaproteogenomic approach was successfully applied compared to common metaproteomic approaches. Altogether the results primarily showed that *Treponema sp.* and *Ruminococcaceae* members were abundant and active bacterial groups during filter paper and cotton degradation. By contrast, *Bacteroidetes* members largely dominated the bacterial community from facial tissue and were also an active group. Further analyses are required to fully understand the role of identified groups during the methanization of the three studied materials. This work brings new insights into the solid waste methanization processes by studying the effects of substrates' specific characteristics on the microbial populations, increasing the core knowledge required to develop more robust and efficient bioprocesses, more specifically optimized for abundant lignocellulosic waste.

Acknowledgments

This work was supported by Région Ile-de-France (R2DS 2010-08 project). The authors acknowledge Franck Seara and Said El Adami for providing digestion sludge from a French mesophilic waste methanization plant (SYVOM, Urbasys SAS).

Points-clés

Confirmations :

- Les différences de réactivité du mouchoir en papier, du papier filtre et du coton mises en évidence sur le système modèle (Badalato *et al.*, in preparation)
- Sensibilité des communautés microbiennes aux des caractéristiques physico-chimiques du substrat (Badalato *et al.*, in preparation)
- Apports de la démarche de métaprotéogénomique pour l'étude des bioprocédés de digestion anaérobie des composés lignocellulosiques (Hanreich *et al.*, 2013)

Nouveaux éléments :

- Rôle probable des bactéries appartenant au genre *Treponema* dans la dégradation des substrats récalcitrants
- La probable proportion limitée de microorganismes cellulolytiques au sein des communautés microbiennes dégradant les substrats lignocellulosiques

DISCUSSION GENERALE

La mise en place d'une stratégie de travail transversale pour traiter le sujet de thèse a permis d'aborder la problématique sous différents angles et de bénéficier à la fois d'informations obtenues sur des systèmes bien caractérisés, tel que la souche pure *C. cellulolyticum*, et d'informations obtenues sur des systèmes de digestion anaérobie plus réalistes dans un contexte de valorisation de déchets organiques. L'objectif de cette discussion est de confronter les principales conclusions issues des différentes études menées pendant ce travail de thèse. L'analyse des données de métaprotéomiques étant toujours en cours, cette discussion n'a pas vocation à être exhaustive. Dans ce chapitre, les références à ces études sont signalées par le numéro de chapitre correspondant (I, II et III).

I. Origine de la récalcitrance

Lors de la thèse, une partie du travail a été consacrée à l'analyse physico-chimique des principaux substrats utilisés – paille de blé, carton plat, mouchoir en papier, papier filtre, coton en disques. Ces caractérisations avaient pour but d'établir un lien entre les propriétés de ces substrats et les performances de biodégradabilité observées en conditions anaérobies et d'émettre des hypothèses sur leurs effets sur le compartiment microbien réalisant cette bioconversion. Les performances de biodégradabilité, à la fois en termes de cinétiques ou de taux de bioconversion, peuvent en particulier être affectées par des facteurs tels que le taux de cristallinité de la cellulose, la présence de lignine, la surface accessible et le degré de polymérisation (Mansfield *et al.*, 1999). Ces facteurs ont été mesurés sur les substrats « simples », tels que le mouchoir en papier, le papier filtre et le coton en disques (II), et/ou sur les substrats « complexes », tels que la paille et le carton broyés (I).

Les résultats ont tout d'abord révélé que les différences de biodégradabilité observées ne semblent pas totalement liées aux taux de régions cristallines de la cellulose (I et II) bien que ce facteur soit souvent désigné comme prépondérant. D'autre part, il a déjà été montré que le taux de cristallinité de la cellulose n'était pas modifié au cours de la dégradation de cellulose pure par *C. cellulolyticum*, ce qui pourrait suggérer que les bactéries cellulolytiques n'ont pas de préférence pour les régions amorphes ou cristallines de la cellulose (Desvaux *et al.*, 2000). Cette hypothèse pourrait s'expliquer par la présence d'enzymes très actives sur la cellulose cristalline.

Les substrats « simples » utilisés ne comportent pas de lignine. En revanche, des taux de lignine différents entre la paille de blé et le carton ont été notés (I), avec un plus fort taux de lignine dans le carton par rapport à la paille de blé. Dans cette étude, le carton était associé à une biodégradabilité plus faible, si bien que le taux de lignine pourrait être un facteur explicatif. Dans les conditions expérimentales utilisées il n'est pas possible de conclure quant au comportement de la lignine, s'il s'agit uniquement d'un composé difficilement biodégradable (effet sur le rendement) et/ou un

composé « écran » (effet sur la cinétique) puisqu'aucune expérience faisant varier la teneur en lignine n'a été réalisée. D'autre part, il ne peut pas être exclu que ce résultat soit également dû à la présence d'additifs (par exemple des colles) ajoutés lors de la fabrication du carton plat. Bien que deux cas ne soient pas suffisants pour conclure sur l'effet du taux de lignine, une autre étude a mis en évidence cette corrélation entre taux de lignine et biodégradabilité (Zhu *et al.*, 2008).

Concernant la surface accessible, l'étude de la littérature a mis en évidence que l'effet de ce facteur dépend notamment du matériau de départ en conditions de type « BMP » (Pommier *et al.*, 2010; Sharma *et al.*, 1988). Les broyages par voie sèche appliqués à la paille de blé et au carton ont permis de générer deux gammes de tailles de particules (~50 µm et ~170 µm) et donc deux gammes de valeurs de surface accessible (I). Les résultats obtenus ont confirmé un effet limité mais positif de la réduction à des tailles de particules fines de la paille de blé et l'absence d'effet de la réduction de taille du carton, même à des tailles de particules très fines. Ce résultat pourrait en partie s'expliquer par la nature de ces matériaux. La paille de blé est un matériau végétal extrêmement structuré à différents niveaux – parois cellulaires, tissus, organes – dont les caractéristiques structurales et chimiques ne sont pas homogènes (Silva and Rouau, 2011). Le carton plat est un matériau fabriqué à partir de composés végétaux ayant subi de nombreux traitements, lors d'un procédé passant par la fabrication d'une pâte homogène. Ainsi, la réduction de taille des particules de paille de blé pourrait découvrir des structures hétérogènes en termes de biodégradabilité, tandis que la réduction de taille des particules de carton n'aurait d'effets que sur la surface accessible et ce dernier facteur n'apparaît donc pas comme suffisant pour modifier les performances de biodégradation.

Le degré de polymérisation des sucres a également été estimé pour les substrats « simples » (II). La méthode utilisée n'a cependant pas pu être appliquée aux substrats « complexes » car elle nécessite plusieurs stades de délignification qui ne peuvent actuellement pas garantir la préservation des propriétés initiales de la cellulose. Les résultats ont montré que le degré de polymérisation le plus faible était observé pour le mouchoir en papier, par ailleurs associé à une meilleure biodégradabilité tandis que le degré de polymérisation le plus élevé était mesuré pour le coton, par ailleurs associé à une biodégradabilité plus faible (II et III). Ceci pourrait suggérer un rôle important de ce facteur dans la biodégradabilité des substrats. Il est cependant rarement mesuré dans le contexte de la digestion anaérobie, certainement parce que des substrats « complexes » sont étudiés la plupart du temps, et que la méthode de mesure ne s'y applique actuellement pas. En revanche, son effet a déjà été étudié dans le cadre de la digestion enzymatique de cellulose pure, *via* l'utilisation d'enzymes pures, et il a été montré que plus le degré de polymérisation est élevé et plus le taux de dégradation est faible (Puri, 1984). D'autre part, les mesures de degré de polymérisation intègrent les signaux issus simultanément de la cellulose et des hémicelluloses présentes, qui possèdent un degré de polymérisation plus faible

que la cellulose. Dans le cas du mouchoir en papier, substrat le plus réactif, la présence d'hémicelluloses a été détectée (II). Il apparaît donc que la présence de ces biopolymères pourrait également avoir un effet positif sur la bioconversion des composés lignocellulosiques.

La contribution de chacun de ces facteurs à la récalcitrance globale des substrats est difficilement évaluable en raison de leurs interactions. L'importance de leur contribution peut en effet dépendre de celle des autres facteurs, comme cela a été montré pour les taux de lignine et de cristallinité (Zhu *et al.*, 2008). Il apparaît donc que l'origine de la récalcitrance des composés lignocellulosiques est une combinaison de propriétés micro- et macroscopiques.

II. Impact de la récalcitrance sur le compartiment microbien

II.1. Effets de la récalcitrance sur la structure des communautés

L'effet de la récalcitrance des composés lignocellulosiques sur les comportements microbiens a particulièrement été étudié pendant ce travail de thèse. Les études réalisées en biomasses complexes ont montré que lors de la dégradation du carton plat et du coton en disques, des bactéries appartenant au phylum des *Spirochaetes*, et en particulier du genre *Treponema*, étaient abondantes dans la communauté microbienne (I et III). Cette observation pourrait suggérer qu'il existe un lien entre la récalcitrance du substrat et l'abondance de ce groupe. Dans ce but, l'abondance relative des bactéries appartenant au phylum des *Spirochaetes* a été reliée au temps de demi-production de méthane pour un ratio inoculum/substrat donné. Le ratio inoculum/substrat fait en effet partie des facteurs influençant le démarrage et la cinétique de production de méthane, il est donc difficile de comparer des études réalisées avec des ratios différents (Boulangier, 2011; Motte *et al.*, 2013). On notera que le temps de demi-production de méthane ne semble pas être un indicateur idéal de la récalcitrance puisqu'il n'est pas seulement lié à l'étape d'hydrolyse du substrat récalcitrant, mais également à toutes les étapes intervenant en aval jusqu'à la méthanogénèse. Il a cependant été utilisé ici à défaut de disposer d'un indicateur plus adapté. Les résultats n'ont montré aucune corrélation significative entre ces paramètres mais une tendance non-significative est apparue, montrant que plus le temps de demi-production de méthane est élevé et plus les bactéries de ce groupe sont abondantes. Il a déjà été montré que des bactéries du genre *Treponema*, telle que *Treponema bryantii*, sont impliquées dans des interactions avec des bactéries cellulolytiques dans le rumen (Stanton and Canale-Parola, 1980), en particulier avec *Fibrobacter succinogenes* (Shinkai *et al.*, 2010). Cependant ici, le genre *Fibrobacter* n'a été que peu (I) ou pas détecté (III) (< 5% du total des séquences d'ARNr 16S) dans les échantillons, ce qui suggère que, dans les conditions expérimentales utilisées, les bactéries du genre *Treponema*

sont impliquées dans des interactions non décrites à ce jour. Il a également été montré que ce taxon est très abondant dans le tube digestif des termites (Warnecke *et al.*, 2007). Les analyses de pyroséquençage des gènes d'ARNr 16S effectuées sur les incubations avec le carton et le coton en disques ont désigné *Treponema lecithinolyticum* comme l'espèce cultivée la plus proche de l'OTU majoritaire assignée à ce taxon. Cette souche est décrite comme saccharolytique si bien que les bactéries du genre *Treponema* présentes dans les microcosmes pourraient participer à la fermentation des sucres libérés lors de l'hydrolyse des substrats. Les analyses de métagénomique réalisées ont montré que, dans les incubations avec le coton, les protéines assignées à des bactéries appartenant au phylum *Spirochaetes* étaient majoritairement assignées à *Treponema brennaboreense*. Peu d'informations sont disponibles sur cette bactérie, cependant elle est décrite comme une bactérie très motile. L'annotation des protéines a été déterminée en se basant sur le niveau de similarité entre la séquence issue des métagénomiques analysés et les séquences présentes dans la banque de données RefSeq et les séquences les plus proches étaient celles de *T. brennaboreense*. Le ou les microorganisme(s) dont les protéines ont été détectées partagent donc un certain niveau de similarité de séquences avec celles *T. brennaboreense*. Parmi ces protéines détectées, de nombreuses étaient impliquées dans la motilité cellulaire ce qui suggère que, comme *T. brennaboreense*, ce/ces microorganisme(s) sont très motiles. Des protéines de transport des sucres ont également été identifiées, ce qui suggère qu'ils sont également impliqués dans l'utilisation des sucres présents dans le milieu, donc très probablement des produits d'hydrolyse des substrats. Des analyses plus approfondies sont en cours, notamment sur l'identification de CAZymes dans les données de métagénomique et de métagénomique, afin de déterminer si ces microorganismes peuvent également être directement impliqués dans la dégradation des lignocelluloses, en particulier des hémicelluloses, comme cela a été montré chez les termites (Warnecke *et al.*, 2007).

II.2. Effets de la récalcitrance sur la colonisation microbienne

En plus de l'effet sur la structure des communautés microbiennes, un lien semble également exister entre la récalcitrance et la vitesse et l'étendue de colonisation des substrats « simples » par les microorganismes (II et III). Malheureusement les difficultés techniques rencontrées pour l'étude de la colonisation en biomasses complexes n'ont pas permis d'obtenir ces données de fiabilité suffisante pour les substrats « complexes » ni pour les substrats « simples », en raison de la variabilité observée. De même, le comptage du nombre de bactéries adhérentes et la quantification de la biomasse adhérente, par analyse d'images, n'ont pas pu être réalisés en raison de la complexité de structure du biofilm sur la plupart des échantillons (amas de cellules et présence de colonies au sein des reliefs du substrat). Des grandes tendances ont cependant pu être dégagées. Ces études ont ainsi révélé que le mouchoir en papier, qui était le plus rapidement dégradé, était également plus rapidement et plus largement

colonisé, en comparaison avec le papier filtre et le coton, ce dernier étant *a contrario* dégradé et colonisé le plus lentement. Ces résultats sont cohérents avec l'hypothèse selon laquelle la concentration en sites accessibles pour l'adhésion et la colonisation bactérienne est un facteur plus déterminant que la concentration en masse du substrat, lors de la dégradation d'un substrat insoluble (II). Les caractéristiques de ces sites d'adhésion n'ont pas été déterminées mais semblent être en lien avec celles impliquées dans la récalcitrance des substrats (II). L'identification de ces caractéristiques pourrait en particulier être réalisée *via* l'identification des mécanismes mis en jeu lors de l'adhésion et la colonisation du substrat et qui restent peu connus. L'implication du cellulosome dans l'adhésion suggère que cette étape dépend de l'interaction entre les domaines de liaisons au substrat (CBM) portés par les cellulosomes et des motifs particuliers du substrat, en raison de leurs affinités particulières (Cantarel *et al.*, 2009). En revanche, l'implication de structures de type pili, suggère une interaction moins motif-spécifique et peut-être une plus grande importance de facteurs physico-chimiques, tels que les reliefs du substrat ou des interactions non spécifiques, dans cette étape d'adhésion des bactéries à leur substrat. Les protéines impliquées dans la formation des pilis n'ont cependant été détectées dans aucune des deux études de protéomique réalisées (II et III).

Les données obtenues sur la structuration spatiale des communautés colonisant les substrats « simples » lors de l'étude réalisée avec la biomasse complexe sont limitées (III). Néanmoins les résultats ont mis en évidence la forte proportion de bactéries appartenant à la famille *Ruminococcaeae* à la surface du substrat en comparaison avec les bactéries appartenant au phylum *Bacteroidetes*. Ces bactéries de la famille *Ruminococcaeae*, pourraient être phylogénétiquement proches d'*Acetivibrio cellulolyticus*, comme l'ont révélé les analyses de pyroséquençage des gènes ARNr 16S. En nombre de séquences, les OTUs affiliées à cette espèce ne représentent qu'une faible proportion de toutes les séquences obtenues, et cela pourrait suggérer que les bactéries appartenant à ces OTUs ne sont abondantes qu'à la surface du substrat, ce qui expliquerait leur faible proportion au sein de la communauté totale. Les résultats obtenus par qPCR mettent également en évidence que les microorganismes cellulolytiques représentent une proportion limitée de la biomasse totale. Cette proximité entre substrat et bactéries cellulolytiques apparaît finalement comme incontournable, puisque même les bactéries dont le système enzymatique de cellulolyse ne repose pas sur les cellulosomes s'associent souvent étroitement aux substrats (Wang *et al.*, 2011). Ceci suggère donc que l'application de conditions opératoires favorisant la colonisation des matériaux lignocellulosiques représente un fort intérêt pour l'amélioration de leur dégradation d'une part mais également pour leur fermentation en méthane grâce à tous les avantages qu'offre le mode de vie en biofilm, telles que, par exemple, la résilience et la résistance des communautés aux variations physico-chimiques.

III. Éléments sur les stratégies de (pré)traitement des déchets

Les résultats de l'étude sur les broyages très fins de paille de blé et de carton plat par broyage par voie sèche ont permis d'explorer les limites des effets de ce type de prétraitement, et ont confirmé les effets modérés de ces technologies (I). L'augmentation de la production de méthane de 10% observée pour un broyage très fin de la paille ne constitue pas un résultat suffisamment marqué pour que ce prétraitement précis soit retenu dans un contexte de valorisation de déchets, d'autant plus en considérant les dépenses énergétiques liées à son utilisation. Il semble donc y avoir une certaine cohérence entre les contraintes liées à l'efficacité (un broyage en particules ultra fines n'apporte pas de gains de production significatifs par rapport à certains broyages plus grossiers) et aux aspects économiques (en moyenne, les broyages intenses permettant d'atteindre des tailles de particules plus fines sont plus coûteux). En comparaison, les prétraitements d'autre nature décrits comme les plus efficaces permettent de doubler la production de méthane. D'autre part, les matériaux lignocellulosiques étudiés ici ne comptent pas parmi les moins facilement biodégradables, en comparaison par exemple avec du bois. Néanmoins, les avantages des technologies de prétraitement mécanique à sec résident dans le fait qu'elles ne consomment pas d'eau, ne produisent pas d'effluents, qu'elles peuvent être utilisées seules (contrairement aux méthodes chimiques qui nécessitent un broyage préalable), et que certaines pourraient être intégrées avec succès au sein d'installations existantes (Lindmark *et al.*, 2012), car associées à une balance énergétique globale positive. D'autre part, les technologies de broyage par voie sèche sont nombreuses et les résultats ont montré que selon la technologie employée il est possible d'obtenir des effets différents sur les structures des composés (I). Dans la mesure où il est dans tous les cas souhaitable de broyer les déchets solides avant leur valorisation (homogénéisation facilitée et meilleure homogénéité, limitation des problèmes de colmatage, etc), l'étude et l'optimisation des technologies de broyage offre donc de nouvelles possibilités pour l'amélioration des performances de dégradation des composés lignocellulosiques, en s'orientant vers des méthodes rudimentaires.

L'ensemble des résultats obtenus pendant la thèse a mis en évidence la sensibilité des communautés microbiennes aux différences parfois fines de composition et de structure des substrats, qui se manifeste par l'émergence de communautés parfois différentes (I, II et III). Il est déjà établi que des paramètres opératoires peuvent avoir un effet sur la structure des communautés microbiennes et sur les performances de dégradation. Ceci suggère que des conditions opératoires particulières adaptées à la bioconversion des composés lignocellulosiques ou prenant mieux en compte la présence de tels substrats au sein des déchets et résidus pourraient être recherchées.

CONCLUSION ET PERSPECTIVES

Dans le cadre général de la méthanisation des composés lignocellulosiques et des problématiques liées à la récalcitrance de la lignocellulose à la dégradation, le sujet de thèse traitait de la colonisation des déchets lignocellulosiques en conditions anaérobies en lien avec l'efficacité de leur dégradation. En particulier, le déterminisme de cette étape de colonisation en lien avec le démarrage de l'hydrolyse a été abordé, notamment concernant les caractéristiques physico-chimiques des substrats avec ou sans prétraitements. L'ensemble du processus a été pris en compte afin de mieux appréhender les effets sur l'ensemble de la bioconversion de ces composés et les mécanismes et interactions mis en jeu au sein de la biomasse ont été étudiés.

Afin d'atteindre les objectifs de thèse, une approche transversale a été développée, combinant des modèles de cultures de souches pures proches des souches présentes dans les systèmes complexes, tel que *C. cellulolyticum*, modèle des *Clostridia* cellulolytiques mésophiles, et des systèmes de méthanisation en laboratoire par des communautés complexes. Cette approche permet une étude plus approfondie des processus, en particulier des mécanismes cellulaires, influencés par les différentes conditions, grâce aux informations déjà disponibles sur les modèles biologiques, et également d'évaluer la cohérence possible de ces résultats issus de modèles biologiques avec ceux obtenus dans des conditions plus réalistes dans un contexte de méthanisation des composés lignocellulosiques. Plus précisément, différents types de déchets lignocellulosiques réalistes ont été étudiés et caractérisés (en particulier de la paille de blé et du carton plat prétraités par broyages par voie sèche ou natifs, du mouchoir en papier, du papier filtre et du coton en disques). Des approches intégratives ont été appliquées à l'étude de ces systèmes, couplant des analyses haut-débit (métagénomique/(méta)protéomique) afin d'étudier les mécanismes biologiques et/ou les interactions écologiques mises en jeu, avec un suivi physico-chimique de la biodégradation et enfin des caractérisations physico-chimiques de ces composés.

Les principales conclusions à l'issue de ce travail de thèse sont :

- le rôle prépondérant des propriétés chimiques et structurales des composés lignocellulosiques sur les profils de colonisation et de dégradation ;
- le lien probable entre des propriétés micro- et macro-structurales des composés lignocellulosiques et leur récalcitrance ;
- les limites des broyages fins appliqués à la paille de blé et au carton plat sur les performances de leur dégradation, avec un effet positif modéré du broyage fin de la paille;

- la sensibilité des communautés microbiennes aux changements parfois fins de surface du substrat, qui se manifeste par l'émergence de communautés parfois différentes en fonction du prétraitement mécanique appliqué ;
- le lien possible entre l'émergence de bactéries appartenant au genre *Treponema* et la récalcitrance du substrat ;
- la mise en évidence d'une modulation quantitative du système cellulasique chez *Clostridium cellulolyticum* lors de la dégradation de mouchoir en papier accompagné d'un profil métabolique particulier qui est probablement lié aux performances de dégradation ;
- la bonne concordance des résultats obtenus sur le système modèle utilisé et sur les systèmes réalistes pour l'étude de la biodégradation des déchets lignocellulosiques simples.

Ce travail a permis de traiter sous un angle nouveau les questions liées à la récalcitrance des déchets lignocellulosiques en abordant à la fois les aspects structuraux, écologiques et fonctionnels. Ces résultats alimentent le corps de connaissances fondamentales sur les bioprocédés. Ils confirment que les matériaux lignocellulosiques sont particuliers parmi les déchets non-dangereux, de part leur nature et leur effets sur le compartiment microbien, et nécessitent la mise en œuvre de procédés adaptés.

Les visées applicatives de ces travaux sont assez larges et concernent l'optimisation des procédés de prétraitement en amont et la conception de bioprocédés de digestion et de fermentation adaptés au traitement de grandes quantités de matériaux lignocellulosiques, en particulier pour la production de méthane mais également pour d'autres molécules d'intérêt telle que le bioéthanol.

RÉFÉRENCES

- Abbassi-Guendouz, A., Brockmann, D., Trably, E., Dumas, C., Delgenès, J.-P., Steyer, J.-P., Escudié, R., 2012. Total solids content drives high solid anaerobic digestion via mass transfer limitation. *Bioresour. Technol.* 111, 55–61.
- Abbassi-Guendouz, A., Trably, E., Hamelin, J., Dumas, C., Steyer, J.P., Delgenès, J.-P., Escudié, R., 2013. Microbial community signature of high-solid content methanogenic ecosystems. *Bioresour. Technol.* 133, 256–62.
- Adams, J.J., Pal, G., Jia, Z., Smith, S.P., 2006. Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A.* 103, 305–10.
- ADEME, 2009. Campagne nationale de caractérisation des ordures ménagères.
- Akin, D.E., Benner, R., 1988. Degradation of polysaccharides and lignin by ruminal bacteria and fungi. *Appl. Environ. Microbiol.* 54, 1117–25.
- Amann, R., Fuchs, B.M., Behrens, S., 2001. The identification of microorganisms by fluorescence in situ hybridisation. *Curr. Opin. Biotechnol.* 12, 231–6.
- Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., Stahl, D. a, 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 56, 1919–25.
- AMORCE, 2011. Méthanisation des déchets ménagers, Etat des lieux.
- Arantes, V., Saddler, J.N., 2011. Cellulose accessibility limits the effectiveness of minimum cellulase loading on the efficient hydrolysis of pretreated lignocellulosic substrates. *Biotechnol. Biofuels* 4, 3.
- ATEE, 2011. Etat des lieux de la filiere methanisation en France.
- Ayers, W. a, 1958. Nutrition and physiology of *Ruminococcus flavefaciens*. *J. Bacteriol.* 76, 504–9.
- Baba, Y., Tada, C., Fukuda, Y., Nakai, Y., 2013. Improvement of methane production from waste paper by pretreatment with rumen fluid. *Bioresour. Technol.* 128, 94–9.
- Bacchetti De Gregoris, T., Aldred, N., Clare, A.S., Burgess, J.G., 2011. Improvement of phylum- and class-specific primers for real-time PCR quantification of bacterial taxa. *J. Microbiol. Methods* 86, 351–6.
- Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., Wolfe, R.S., 1979. Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* 43, 260–96.
- Barakat, A., de Vries, H., Rouau, X., 2013. Dry fractionation process as an important step in current and future lignocellulose biorefineries: a review. *Bioresour. Technol.* 134, 362–73.

- Bayané, A., Guiot, S.R., 2010. Animal digestive strategies versus anaerobic digestion bioprocesses for biogas production from lignocellulosic biomass. *Rev. Environ. Sci. Bio/Technology* 10, 43–62.
- Bayer, E. a, Belaich, J.-P., Shoham, Y., Lamed, R., 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* 58, 521–54.
- Bayer, E. a, Chanzy, H., Lamed, R., Shoham, Y., 1998. Cellulose, cellulases and cellulosomes. *Curr. Opin. Struct. Biol.* 8, 548–57.
- Bayer, E. a, Kenig, R., Lamed, R., 1983. Adherence of *Clostridium thermocellum* to cellulose. *J. Bacteriol.* 156, 818–27.
- Bayer, E. a, Lamed, R., 1986. Ultrastructure of the cell surface cellulosome of *Clostridium thermocellum* and its interaction with cellulose. *J. Bacteriol.* 167, 828–36.
- Bayer, E. a, Lamed, R., White, B. a, Flint, H.J., 2008. From cellulosomes to cellulosomics. *Chem. Rec.* 8, 364–77.
- Béguin, P., Lemaire, M., 1996. The cellulosome: an exocellular, multiprotein complex specialized in cellulose degradation. *Crit. Rev. Biochem. Mol. Biol.* 31, 201–36.
- Bélaich, J.P., Tardif, C., Bélaich, A., Gaudin, C., 1997. The cellulolytic system of *Clostridium cellulolyticum*. *J. Biotechnol.* 57, 3–14.
- Belanche, A., Doreau, M., Edwards, J.E., Moorby, J.M., Pinloche, E., Newbold, C.J., 2012. Shifts in the Rumen Microbiota Due to the Type of Carbohydrate and Level of Protein Ingested by Dairy Cattle Are Associated with Changes in Rumen Fermentation. *J. Nutr.* 142, 1684–1692.
- Beloin, C., Houry, A., Froment, M., Ghigo, J.-M., Henry, N., 2008. A short-time scale colloidal system reveals early bacterial adhesion dynamics. *PLoS Biol.* 6, e167.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *R. Stat. Soc.* 57, 289–300.
- Bernalier, A., Fonty, G., Bonnemoy, F., Gouet, P., 1993. Inhibition of the cellulolytic activity of *Neocallimastix frontalis* by *Ruminococcus flavefaciens*. *J. Gen. Microbiol.* 139, 873–880.
- Beukes, N., Chan, H., Doi, R.H., Pletschke, B.I., 2008. Synergistic associations between *Clostridium cellulovorans* enzymes XynA, ManA and EngE against sugarcane bagasse. *Enzyme Microb. Technol.* 42, 492–498.
- Blouzard, J.-C., Coutinho, P.M., Fierobe, H.-P., Henrissat, B., Lignon, S., Tardif, C., Pagès, S., de Philip, P., 2010. Modulation of cellulosome composition in *Clostridium cellulolyticum*: adaptation to the polysaccharide environment revealed by proteomic and carbohydrate-active enzyme analyses. *Proteomics* 10, 541–54.
- Blumer-Schuetz, S.E., Kataeva, I., Westpheling, J., Adams, M.W., Kelly, R.M., 2008. Extremely thermophilic microorganisms for biomass conversion: status and prospects. *Curr. Opin. Biotechnol.* 19, 210–7.
- Bomble, Y.J., Beckham, G.T., Matthews, J.F., Nimlos, M.R., Himmel, M.E., Crowley, M.F., 2011. Modeling the self-assembly of the cellulosome enzyme complex. *J. Biol. Chem.* 286, 5614–23.

- Boulanger, A., 2011. Préparation d'un déchet ménager pour l'optimisation du potentiel et de la cinétique méthanogène.
- Bridier, A., Meylheuc, T., Briandet, R., 2013. Realistic representation of *Bacillus subtilis* biofilms architecture using combined microscopy (CLSM, ESEM and FESEM). *Micron* 48, 65–9.
- Brulc, J.M., Antonopoulos, D. a, Miller, M.E.B., Wilson, M.K., Yannarell, A.C., Dinsdale, E. a, Edwards, R.E., Frank, E.D., Emerson, J.B., Wacklin, P., Coutinho, P.M., Henrissat, B., Nelson, K.E., White, B. a, 2009. Gene-centric metagenomics of the fiber-adherent bovine rumen microbiome reveals forage specific glycoside hydrolases. *Proc. Natl. Acad. Sci. U. S. A.* 106, 1948–53.
- Brune, A., Emerson, D., Breznak, J. a., 1995. The Termite Gut Microflora as an Oxygen Sink: Microelectrode Determination of Oxygen and pH Gradients in Guts of Lower and Higher Termites. *Appl. Environ. Microbiol.* 61, 2681–7.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 37, D233–8.
- Cardinale, M., Brusetti, L., Quatrini, P., Borin, S., Puglia, A.M., Rizzi, A., Zanardini, E., Sorlini, C., Corselli, C., Daffonchio, D., 2004. Comparison of different primer sets for use in automated ribosomal intergenic spacer analysis of complex bacterial communities. *Appl. Environ. Microbiol.* 70, 6147–56.
- Carvalho, A.L., Dias, F.M. V, Prates, J. a M., Nagy, T., Gilbert, H.J., Davies, G.J., Ferreira, L.M. a, Romão, M.J., Fontes, C.M.G. a, 2003. Cellulosome assembly revealed by the crystal structure of the cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A.* 100, 13809–14.
- Caspi, R., Altman, T., Dale, J.M., Dreher, K., Fulcher, C.A., Gilham, F., Kaipa, P., Karthikeyan, A.S., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Paley, S., Popescu, L., Pujar, A., Shearer, A.G., Zhang, P., Karp, P.D., 2010. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 38, D473–D479.
- Cavalier-Smith, T., 2004. Only six kingdoms of life. *Proc. Biol. Sci.* 271, 1251–62.
- Chang, V.S., Holtzapple, M.T., 2000. Fundamental factors affecting biomass enzymatic reactivity. *Appl. Biochem. Biotechnol.*
- Chapleur, O., 2012. Ingénierie écologique des communautés microbiennes de méthanisation des déchets ligno-cellulosiques.
- Chen, J., Stevenson, D.M., Weimer, P.J., 2004. Albusin B, a bacteriocin from the ruminal bacterium *Ruminococcus albus* 7 that inhibits growth of *Ruminococcus flavefaciens*. *Appl. Environ. Microbiol.* 70, 3167–70.
- Chen, J., Weimer, P., 2001. Competition among three predominant ruminal cellulolytic bacteria in the absence or presence of non-cellulolytic bacteria. *Microbiology* 147, 21–30.
- Cho, W., Jeon, S.D., Shim, H.J., Doi, R.H., Han, S.O., 2010. Cellulosomic profiling produced by *Clostridium cellulovorans* during growth on different carbon sources explored by the cohesin

- marker. *J. Biotechnol.* 145, 233–9.
- Costerton, J.W., Lewandowski, Z., Caldwell, D.E., Korber, D.R., Lappin-Scott, H.M., 1995. Microbial biofilms. *Annu. Rev. Microbiol.* 49, 711–45.
- Craig, R., Beavis, R.C., 2004. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 20, 1466–7.
- Crow, V.L., Pritchard, G.G., 1977. Fructose 1,6-diphosphate-activated L-lactate dehydrogenase from *Streptococcus lactis*: kinetic properties and factors affecting activation. *J. Bacteriol.* 131, 82–91.
- Dam, P., Kataeva, I., Yang, S.-J., Zhou, F., Yin, Y., Chou, W., Poole, F.L., Westpheling, J., Hettich, R., Giannone, R., Lewis, D.L., Kelly, R., Gilbert, H.J., Henrissat, B., Xu, Y., Adams, M.W.W., 2011. Insights into plant biomass conversion from the genome of the anaerobic thermophilic bacterium *Caldicellulosiruptor bescii* DSM 6725. *Nucleic Acids Res.* 39, 3240–3254.
- Dassa, B., Borovok, I., Lamed, R., Henrissat, B., Coutinho, P., Hemme, C.L., Huang, Y., Zhou, J., Bayer, E. a, 2012. Genome-wide analysis of *acetivibrio cellulolyticus* provides a blueprint of an elaborate cellulosome system. *BMC Genomics* 13, 210.
- Denger, K., Warthmann, R., 2002. *Anaerophaga thermohalophila* gen. nov., sp. nov., a moderately thermohalophilic, strictly anaerobic fermentative bacterium. *J. Syst.* 173–178.
- Desvaux, M., 2006. Unravelling carbon metabolism in anaerobic cellulolytic bacteria. *Biotechnol. Prog.* 22, 1229–38.
- Desvaux, M., 2005. *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol. Rev.* 29, 741–64.
- Desvaux, M., 2004. Mapping of carbon flow distribution in the central metabolic pathways of *Clostridium cellulolyticum*: Direct comparison of bacterial metabolism with a soluble versus an insoluble carbon source. *J. Microbiol. Biotechnol.* 14, 1200–1210.
- Desvaux, M., 2001. La Fermentation de la Cellulose par *Clostridium cellulolyticum* : Métabolisme Modèle d'un *Clostridium Cellulolytique* Mésophile.
- Desvaux, M., Dumas, E., Chafsey, I., Hébraud, M., 2006. Protein cell surface display in Gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol. Lett.* 256, 1–15.
- Desvaux, M., Guedon, E., Petitdemange, H., 2001a. Kinetics and metabolism of cellulose degradation at high substrate concentrations in steady-state continuous cultures of *Clostridium cellulolyticum* on a chemically defined medium. *Appl. Environ. Microbiol.* 67, 3837–45.
- Desvaux, M., Guedon, E., Petitdemange, H., 2001. Carbon flux distribution and kinetics of cellulose fermentation in steady-state continuous cultures of *Clostridium cellulolyticum* on a chemically defined medium. *J. Bacteriol.* 183, 119–30.
- Desvaux, M., Guedon, E., Petitdemange, H., 2001b. Metabolic flux in cellulose batch and cellulose-fed continuous cultures of *Clostridium cellulolyticum* in response to acidic environment. *Microbiology* 147, 1461–71.

- Desvaux, M., Guedon, E., Petitdemange, H., 2000. Cellulose catabolism by *Clostridium cellulolyticum* growing in batch culture on defined medium. *Appl. Environ. Microbiol.* 66, 2461–70.
- Devaux, M.F., Barakat, A., Robert, P., Bouchet, B., Guillon, F., Navez, B., Lahaye, M., 2005. Mechanical breakdown and cell wall structure of mealy tomato pericarp tissue. *Postharvest Biol. Technol.* 37, 209–221.
- Devillard, E., Goodheart, D., 2004. Ruminococcus albus 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel. *J. Bacteriol.* 186, 136–145.
- Ding, S.Y., Bayer, E. a, Steiner, D., Shoham, Y., Lamed, R., 2000. A scaffoldin of the *Bacteroides cellulosolvens* cellulosome that contains 11 type II cohesins. *J. Bacteriol.* 182, 4915–25.
- Ding, S.Y., Bayer, E. a, Steiner, D., Shoham, Y., Lamed, R., 1999. A novel cellulosomal scaffoldin from *Acetivibrio cellulolyticus* that contains a family 9 glycosyl hydrolase. *J. Bacteriol.* 181, 6720–9.
- Dinsdale, D., Morris, E.J., Bacon, J.S., 1978. Electron microscopy of the microbial populations present and their modes of attack on various cellulosic substrates undergoing digestion in the sheep rumen. *Appl. Environ. Microbiol.* 36, 160–8.
- Doi, R.H., 2008. Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. *Ann. N. Y. Acad. Sci.* 1125, 267–79.
- Doi, R.H., Kosugi, A., 2004. Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol.* 2, 541–51.
- Doi, R.H., Kosugi, A., Murashima, K., Tamaru, Y., Han, S.O., 2003. Cellulosomes from mesophilic bacteria. *J. Bacteriol.* 185, 5907–14.
- Dumitrache, A., Wolfaardt, G., Allen, G., Liss, S.N., Lynd, L.R., 2013. Form and function of *Clostridium thermocellum* biofilms. *Appl. Environ. Microbiol.* 79, 231–9.
- Dykstra, A.B., Rodriguez, M., Raman, B., Cook, K.D., Hettich, R.L., 2013. Characterizing the range of extracellular protein post-translational modifications in a cellulose-degrading bacteria using a multiple proteolytic digestion/peptide fragmentation Approach. *Anal. Chem.* 85, 3144–3151.
- Ellis, L.D., Holwerda, E.K., Hogsett, D., Rogers, S., Shao, X., Tschapinski, T., Thorne, P., Lynd, L.R., 2011. Closing the carbon balance for fermentation by *Clostridium thermocellum* (ATCC 27405). *Bioresour. Technol.* 103, 293–299.
- EPA, 2011. Municipal Solid Waste in The United States.
- Fechner, L.C., Vincent-Hubert, F., Gaubert, P., Bouchez, T., Gourlay-Francé, C., Tusseau-Vuillemin, M.-H., 2010. Combined eukaryotic and bacterial community fingerprinting of natural freshwater biofilms using automated ribosomal intergenic spacer analysis. *FEMS Microbiol. Ecol.* 74, 542–53.
- Fenyö, D., Beavis, R.C., 2003. A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. *Anal. Chem.* 75, 768–74.

- Ferdinand, P.-H., Borne, R., Trotter, V., Pagès, S., Tardif, C., Fierobe, H.-P., Perret, S., 2013. Are cellulosome scaffolding protein CipC and CBM3-containing protein HycP, involved in adherence of *Clostridium cellulolyticum* to cellulose? *PLoS One* 8, e69360.
- Fondevila, M., Dehority, B., 1996. Interactions between *Fibrobacter succinogenes*, *Prevotella ruminicola*, and *Ruminococcus flavefaciens* in the digestion of cellulose from forages. *J. Anim. Sci.* 74, 678.
- Foyle, T., Jennings, L., Mulcahy, P., 2007. Compositional analysis of lignocellulosic materials: evaluation of methods used for sugar analysis of waste paper and straw. *Bioresour. Technol.* 98, 3026–36.
- Frigon, J.J.-C., Mehta, P., Guiot, S.S.R., 2012. Impact of mechanical, chemical and enzymatic pre-treatments on the methane yield from the anaerobic digestion of switchgrass. *Biomass and Bioenergy* 36, 1–11.
- García-Martínez, J., Rodríguez-Valera, F., 2000. Microdiversity of uncultured marine prokaryotes: the SAR11 cluster and the marine Archaea of Group I. *Mol. Ecol.* 9, 935–948.
- Giallo, J., Gaudin, C., Belaich, J.P., 1985. Metabolism and Solubilization of Cellulose by *Clostridium cellulolyticum* H10. *Appl. Environ. Microbiol.* 49, 1216–21.
- Giallo, J., Gaudin, C., Belaich, J.P., Petitdemange, E., Caillet-Mangin, F., 1983. Metabolism of glucose and cellobiose by cellulolytic mesophilic *Clostridium* sp. strain H10. *Appl. Environ. Microbiol.* 45, 843–9.
- Gilad, R., Rabinovich, L., Yaron, S., 2003. Cell, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *J. Biol. Chem.* 278, 391–398.
- Gilbert, H.J., 2007. Cellulosomes: microbial nanomachines that display plasticity in quaternary structure. *Mol. Microbiol.* 63, 1568–76.
- Gilbert, J. a, Field, D., Swift, P., Thomas, S., Cummings, D., Temperton, B., Weynberg, K., Huse, S., Hughes, M., Joint, I., Somerfield, P.J., Mühlhng, M., 2010. The taxonomic and functional diversity of microbes at a temperate coastal site: a “multi-omic” study of seasonal and diel temporal variation. *PLoS One* 5, e15545.
- Gloster, T.M., Turkenburg, J.P., Potts, J.R., Henrissat, B., Davies, G.J., 2008. Divergence of catalytic mechanism within a glycosidase family provides insight into evolution of carbohydrate metabolism by human gut flora. *Chem. Biol.* 15, 1058–67.
- Godon, J.J., Arcemisbèhère, L., Escudié, R., Harmand, J., Miambi, E., Steyer, J.P., 2013. Overview of the Oldest Existing Set of Substrate-optimized Anaerobic Processes: Digestive Tracts. *Bioenergy Res.* 6, 1063–1081.
- Gold, N.D., Martin, V.J.J., 2007. Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.* 189, 6787–95.
- Graham, J.E., Clark, M.E., Nadler, D.C., Huffer, S., Chokhawala, H. a, Rowland, S.E., Blanch, H.W., Clark, D.S., Robb, F.T., 2011. Identification and characterization of a multidomain

- hyperthermophilic cellulase from an archaeal enrichment. *Nat. Commun.* 2, 375.
- Guedon, E., Desvaux, M., Payot, S., Petitdemange, H., 1999a. Growth inhibition of *Clostridium cellulolyticum* by an inefficiently regulated carbon flow. *Microbiology* 145 (Pt 8, 1831–8.
- Guedon, E., Desvaux, M., Petitdemange, H., 2002. Improvement of Cellulolytic Properties of *Clostridium cellulolyticum* by Metabolic Engineering Improvement of Cellulolytic Properties of *Clostridium cellulolyticum* by Metabolic Engineering.
- Guedon, E., Payot, S., Desvaux, M., Petitdemange, H., 2000. Relationships between cellobiose catabolism, enzyme levels, and metabolic intermediates in *Clostridium cellulolyticum* grown in a synthetic medium. *Biotechnol. Bioeng.* 67, 327–35.
- Guedon, E., Payot, S., Desvaux, M., Petitdemange, H., 1999b. Carbon and electron flow in *Clostridium cellulolyticum* grown in chemostat culture on synthetic medium. *J. Bacteriol.* 181, 3262–9.
- Guillén, D., Sánchez, S., Rodríguez-Sanoja, R., 2010. Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.* 85, 1241–9.
- Hamilton-Brehm, S.D., Mosher, J.J., Vishnivetskaya, T., Podar, M., Carroll, S., Allman, S., Phelps, T.J., Keller, M., Elkins, J.G., 2010. *Caldicellulosiruptor obsidiansis* sp. nov., an anaerobic, extremely thermophilic, cellulolytic bacterium isolated from Obsidian Pool, Yellowstone National Park. *Appl. Environ. Microbiol.* 76, 1014–20.
- Han, S.O., Yukawa, H., Inui, M., Doi, R.H., 2005. Effect of carbon source on the cellulosomal subpopulations of *Clostridium cellulovorans*. *Microbiology* 151, 1491–7.
- Hanreich, A., Heyer, R., Benndorf, D., Rapp, E., Pioch, M., Reichl, U., Klocke, M., 2012. Metaproteome analysis to determine the metabolically active part of a thermophilic microbial community producing biogas from agricultural biomass. *Can. J. Microbiol.* 58, 917–22.
- Hanreich, A., Schimpf, U., Zakrzewski, M., Schlüter, A., Benndorf, D., Heyer, R., Rapp, E., Pühler, A., Reichl, U., Klocke, M., 2013. Metagenome and metaproteome analyses of microbial communities in mesophilic biogas-producing anaerobic batch fermentations indicate concerted plant carbohydrate degradation. *Syst. Appl. Microbiol.* 36, 330–338.
- Hansen, M. a T., Kristensen, J.B., Felby, C., Jørgensen, H., 2011. Pretreatment and enzymatic hydrolysis of wheat straw (*Triticum aestivum* L.)--the impact of lignin relocation and plant tissues on enzymatic accessibility. *Bioresour. Technol.* 102, 2804–11.
- Hansen, M.C., Palmer, R.J., White, D.C., 2000. Flowcell culture of *Porphyromonas gingivalis* biofilms under anaerobic conditions. *J. Microbiol. Methods* 40, 233–9.
- Harazono, K., Yamashita, N., Shinzato, N., Watanabe, Y., Fukatsu, T., Kurane, R., 2003. Isolation and characterization of aromatics-degrading microorganisms from the gut of the lower termite *Coptotermes formosanus*. *Biosci. Biotechnol. Biochem.* 67, 889–92.
- Haruta, S., Cui, Z., Huang, Z., Li, M., Ishii, M., Igarashi, Y., 2002. Construction of a stable microbial community with high cellulose-degradation ability. *Appl. Microbiol. Biotechnol.* 59, 529–34.
- Heidelberger, P., Welch, P.D., 1983. Simulation Run Length Control in the Presence of an Initial

- Transient. *Oper. Res.* 31, 1109–1144.
- Heinecke, N.L., Pratt, B.S., Vaisar, T., Becker, L., 2010. PepC: proteomics software for identifying differentially expressed proteins based on spectral counting. *Bioinformatics* 26, 1574–5.
- Hendriks, A., Zeeman, G., 2009. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour. Technol.* 100, 10–8.
- Henniges, U., Kloser, E., Patel, A., Potthast, A., Kosma, P., Fischer, M., Fischer, K., Rosenau, T., 2007. Studies on DMSO-containing carbanilation mixtures: chemistry, oxidations and cellulose integrity. *Cellulose* 14, 497–511.
- Henrissat, B., Driguez, H., Viet, C., 1985. Synergism of cellulases from *Trichoderma reesei* in the degradation of cellulose. *Nat. Biotechnol.*
- Higashide, W., Li, Y., Yang, Y., Liao, J.C., 2011. Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. *Appl. Environ. Microbiol.* 77, 2727–33.
- Himmel, M.E., Bayer, E. a, 2009. Lignocellulose conversion to biofuels: current challenges, global perspectives. *Curr. Opin. Biotechnol.* 20, 316–7.
- Hongoh, Y., Deevong, P., Inoue, T., Moriya, S., Trakunleamsai, S., Ohkuma, M., Vongkaluang, C., Noparatnaraporn, N., Kudo, T., 2005. Intra- and interspecific comparisons of bacterial diversity and community structure support coevolution of gut microbiota and termite host. *Appl. Environ. Microbiol.* 71, 6590–9.
- Hongoh, Y., Sharma, V.K., Prakash, T., Noda, S., Taylor, T.D., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M., Ohkuma, M., 2008. Complete genome of the uncultured Termite Group 1 bacteria in a single host protist cell. *Proc. Natl. Acad. Sci. U. S. A.* 105, 5555–5560.
- Hungate, R.E., 1950. The anaerobic mesophilic cellulolytic bacteria. *Bacteriol. Rev.* 14, 1–49.
- Imam, S., Chen, Z., Roos, D.S., Pohlschröder, M., 2011. Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria. *PLoS One* 6, e28919.
- Izquierdo, J. a, Sizova, M. V, Lynd, L.R., 2010. Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl. Environ. Microbiol.* 76, 3545–53.
- Jackowiak, D., Frigon, J.C., Ribeiro, T., Pauss, A., Guiot, S., 2011. Enhancing solubilisation and methane production kinetic of switchgrass by microwave pretreatment. *Bioresour. Technol.* 102, 3535–40.
- Jami, E., Israel, A., Kotser, A., Mizrahi, I., 2013. Exploring the bovine rumen bacterial community from birth to adulthood. *ISME J.* 7, 1069–79.
- Jami, E., Mizrahi, I., 2012. Composition and similarity of bovine rumen microbiota across individual animals. *PLoS One* 7, e33306.
- Jeihanipour, A., Niklasson, C., Taherzadeh, M.J., 2011. Enhancement of solubilization rate of cellulose in anaerobic digestion and its drawbacks. *Process Biochem.* 46, 1509–1514.

- Jin, M., Balan, V., Gunawan, C., Dale, B.E., 2011. Consolidated bioprocessing (CBP) performance of *Clostridium phytofermentans* on AFEX-treated corn stover for ethanol production. *Biotechnol. Bioeng.* 108, 1290–7.
- Kaiser, J.-P., Hanselmann, K.W., 1982. Fermentative metabolism of substituted monoaromatic compounds by a bacterial community from anaerobic sediments. *Arch. Microbiol.* 133, 185–194.
- Kataeva, I.A., Iii, R.D.S., Shah, A., West, L.T., Li, X., Ljungdahl, L.G., 2002. The Fibronectin Type 3-Like Repeat from the *Clostridium thermocellum* Cellobiohydrolase CbhA Promotes Hydrolysis of Cellulose by Modifying Its Surface 68, 4292–4300.
- Kato, S., Haruta, S., Cui, Z.J., Ishii, M., Yokota, A., Igarashi, Y., 2004. *Clostridium straminisolvans* sp. nov., a moderately thermophilic, aerotolerant and cellulolytic bacterium isolated from a cellulose-degrading bacterial community. *Int. J. Syst. Evol. Microbiol.* 54, 2043–7.
- Kawakami, S., Kubota, K., Imachi, H., Yamaguchi, T., Harada, H., Ohashi, A., 2010. Detection of Single Copy Genes by Two-Pass Tyramide Signal Amplification Fluorescence in situ Hybridization (Two-Pass TSA-FISH) with Single Oligonucleotide Probes. *Microbes Environ.* 25, 15–21.
- Kim, O.-S., Cho, Y.-J., Lee, K., Yoon, S.-H., Kim, M., Na, H., Park, S.-C., Jeon, Y.S., Lee, J.-H., Yi, H., Won, S., Chun, J., 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* 62, 716–21.
- Koike, S., Pan, J., Kobayashi, Y., Tanaka, K., 2003. Kinetics of in sacco fiber-attachment of representative ruminal cellulolytic bacteria monitored by competitive PCR. *J. Dairy Sci.* 86, 1429–35.
- Kuhad, R.C., Singh, A., Eriksson, K.E., 1997. Microorganisms and enzymes involved in the degradation of plant fiber cell walls. *Adv. Biochem. Eng. Biotechnol.* 57, 45–125.
- Kumar, P., Barrett, D.M., Delwiche, M.J., Stroeve, P., 2009. Methods for Pretreatment of Lignocellulosic Biomass for Efficient Hydrolysis and Biofuel Production. *Ind. Eng. Chem. Res.* 48, 3713–3729.
- Lamed, R., Naimark, J., Morgenstern, E., Bayer, E. a, 1987. Specialized cell surface structures in cellulolytic bacteria. *J. Bacteriol.* 169, 3792–800.
- Laube, V.M., Martin, S.M., 1981. Conversion of Cellulose to Methane and Carbon Dioxide by Triculture of *Acetivibrio cellulolyticus*, *Desulfovibrio* sp., and *Methanosarcina barkeri*. *Appl. Environ. Microbiol.* 42, 413–20.
- Lee, S.S., Ha, J.K., Cheng, K., 2000. Relative contributions of bacteria, protozoa, and fungi to in vitro degradation of orchard grass cell walls and their interactions. *Appl. Environ. Microbiol.* 66, 3807–13.
- Leschine, S.B., 1995. Cellulose degradation in anaerobic environments. *Annu. Rev. Microbiol.* 49, 399–426.
- Li, T., Mazéas, L., Sghir, A., Leblon, G., Bouchez, T., 2009. Insights into networks of functional microbes catalysing methanization of cellulose under mesophilic conditions. *Environ. Microbiol.*

11, 889–904.

- Lindmark, J., Leksell, N., Schnürer, A., Thorin, E., 2012. Effects of mechanical pre-treatment on the biogas yield from ley crop silage. *Appl. Energy* 97, 498–502.
- Liu, H.W., Walter, H.K., Vogt, G.M., Vogt, H.S., Holbein, B.E., 2002. Steam pressure disruption of municipal solid waste enhances anaerobic digestion kinetics and biogas yield. *Biotechnol. Bioeng.* 77, 121–130.
- Loy, A., Lehner, A., Lee, N., Meier, H., Ernst, J., Schleifer, K.K.-H.K., Wagner, M., Adamczyk, J., 2002. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl.* 68, 5064–5081.
- Lü, F., Bize, A., Guillot, A., Monnet, V., Madigou, C., Chapleur, O., Mazéas, L., He, P., Bouchez, T., 2013. Metaproteomics of cellulose methanisation under thermophilic conditions reveals a surprisingly high proteolytic activity. *ISME J.*
- Ludwig, W., Schleifer, K., Whitman, W.B., 2007. Revised road map to the phylum Firmicutes.
- Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* 66, 506–77, table of contents.
- Maki, M., Leung, K.T., Qin, W., 2009. The prospects of cellulase-producing bacteria for the bioconversion of lignocellulosic biomass. *Int. J. Biol. Sci.* 5, 500–16.
- Mansfield, S., Mooney, C., Saddler, J., 1999. Substrate and Enzyme Characteristics that Limit Cellulose Hydrolysis. *Biotechnol. Prog.* 15, 804–816.
- Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., Schleifer, K.H., 1996. Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* 142 (Pt 5, 1097–106.
- Manz, W., Amann, R., Ludwig, W., Wagner, M., Schleifer, K.-H., 1992. Phylogenetic Oligodeoxynucleotide Probes for the Major Subclasses of Proteobacteria: Problems and Solutions. *Syst. Appl. Microbiol.* 15, 593–600.
- Martin, J.G.A., Nussey, D.H., Wilson, A.J., Réale, D., 2011. Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression models. *Methods Ecol. Evol.* 2, 362–374.
- Matsui, H., Ogata, K., Tajima, K., Nakamura, M., Nagamine, T., Aminov, R.I., Benno, Y., 2000. Phenotypic Characterization of Polysaccharidases Produced by Four *Prevotella* Type Strains. *Curr. Microbiol.* 41, 45–49.
- McAllister, T. a, Bae, H.D., Jones, G. a, Cheng, K.J., 1994. Microbial attachment and feed digestion in the rumen. *J. Anim. Sci.* 72, 3004–18.
- McBee, R.H., 1954. The characteristics of *Clostridium thermocellum*. *J. Bacteriol.* 67, 505–6.
- McCarter, J., Withers, G.S., 1994. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.*

- McLeod, A.I., 2011. Kendall rank correlation and Mann-Kendall trend test.
- McNeil, M., Darvill, A.G., Fry, S.C., Albersheim, P., 1984. Structure and function of the primary cell walls of plants. *Annu. Rev. Biochem.* 53, 625–63.
- McSweeney, C., Dulieu, A., 1994. Solubilization of lignin by the ruminal anaerobic fungus *Neocallimastix patriciarum*. *Appl.* 60, 2985–2989.
- Mechaly, A, Fierobe, H.P., Belaich, A., Belaich, J.P., Lamed, R., Shoham, Y., Bayer, E. a, 2001. Cohesin-dockerin interaction in cellulosome assembly: a single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition. *J. Biol. Chem.* 276, 9883–8.
- Medie, F.M., Davies, G.J., Drancourt, M., Henrissat, B., 2012. Genome analyses highlight the different biological roles of cellulases. *Nat. Rev. Microbiol.* 10, 227–234.
- Meyer, F., Paarmann, D., D’Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A, Stevens, R., Wilke, A, Wilkening, J., Edwards, R. a, 2008. The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9, 386.
- Michelland, R.J., Combes, S., Dejean, S., Fortun-Lamothe, L., Cauquil, L., 2009. StatFingerprints: a friendly graphical interface program for processing and analysis of microbial fingerprint profiles. *Mol. Ecol.* 9, 1359–63.
- Miron, J., Ben-Ghedalia, D., Morrison, M., 2001. Invited review: adhesion mechanisms of rumen cellulolytic bacteria. *J. Dairy Sci.* 84, 1294–309.
- Miron, J., Yokoyama, M., Lamed, R., 1989. Bacterial cell surface structures involved in lucerne cell wall degradation by pure cultures of cellulolytic rumen bacteria. *Appl. Microbiol.* 218–222.
- Moraïs, S., Barak, Y., Caspi, J., Hadar, Y., Lamed, R., Shoham, Y., Wilson, D.B., Bayer, E. a, 2010. Contribution of a xylan-binding module to the degradation of a complex cellulosic substrate by designer cellulosomes. *Appl. Environ. Microbiol.* 76, 3787–96.
- Moraru, C., Lam, P., Fuchs, B.M., Kuypers, M.M.M., Amann, R., 2010. GeneFISH--an in situ technique for linking gene presence and cell identity in environmental microorganisms. *Environ. Microbiol.* 12, 3057–73.
- Morgavi, D.P., Forano, E., Martin, C., Newbold, C.J., 2010. Microbial ecosystem and methanogenesis in ruminants. *Animal* 4, 1024–36.
- Morgavi, D.P., Kelly, W.J., Janssen, P.H., Attwood, G.T., 2012. Rumen microbial (meta)genomics and its application to ruminant production. *Animal* 1–18.
- Morrison, M., Miron, J., 2000. Adhesion to cellulose by *Ruminococcus albus*: a combination of cellulosomes and Pil-proteins? *FEMS Microbiol. Lett.* 185, 109–15.
- Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour. Technol.* 96, 673–86.
- Mosoni, P., Gaillard-Martinie, B., 2001. Characterization of a spontaneous adhesion-defective mutant

- of *Ruminococcus albus* strain 20. *Arch. Microbiol.* 176, 52–61.
- Mosoni, P., Martin, C., Forano, E., Morgavi, D.P., 2011. Long-term defaunation increases the abundance of cellulolytic ruminococci and methanogens but does not affect the bacterial and methanogen diversity in the rumen of sheep. *J. Anim. Sci.* 89, 783–791.
- Motte, J.-C., Escudié, R., Bernet, N., Delgenes, J.-P., Steyer, J.-P., Dumas, C., 2013. Dynamic effect of total solid content, low substrate/inoculum ratio and particle size on solid-state anaerobic digestion. *Bioresour. Technol.* 144, 141–8.
- Musat, N., Foster, R., Vagner, T., Adam, B., Kuypers, M.M.M., 2011. Detecting metabolic activities in single cells, with emphasis on nanoSIMS. *FEMS Microbiol. Rev.*
- Neilson, K. a, Ali, N. a, Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., van Sluyter, S.C., Haynes, P. a, 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11, 535–53.
- Ng, T.K., Weimer, T.K., Zeikus, J.G., 1977. Cellulolytic and physiological properties of *Clostridium thermocellum*. *Arch. Microbiol.* 114, 1–7.
- Nishiyama, T., Ueki, A., Kaku, N., Ueki, K., 2009. *Clostridium sufflavum* sp. nov., isolated from a methanogenic reactor treating cattle waste. *Int. J. Syst. Evol. Microbiol.* 59, 981–6.
- Ogeda, T.L., Silva, I.B., Fidale, L.C., El Seoud, O. a, Petri, D.F.S., 2012. Effect of cellulose physical characteristics, especially the water sorption value, on the efficiency of its hydrolysis catalyzed by free or immobilized cellulase. *J. Biotechnol.* 157, 246–52.
- Ohkuma, M., 2008. Symbioses of flagellates and prokaryotes in the gut of lower termites. *Trends Microbiol.* 16, 345–52.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H., Wagner, H., 2011. *Vegan: community ecology package*. R package version 1.17-8.
- Pagès, S., Bélaïch, A., Fierobe, H.P., Tardif, C., Gaudin, C., Bélaïch, J.P., 1999. Sequence analysis of scaffolding protein CipC and ORFXp, a new cohesin-containing protein in *Clostridium cellulolyticum*: comparison of various cohesin domains and subcellular localization of ORFXp. *J. Bacteriol.* 181, 1801–10.
- Pagès, S., Bélaïch, A., Bélaïch, J.P., Morag, E., Lamed, R., Shoham, Y., Bayer, E. a, 1997. Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. *Proteins* 29, 517–27.
- Palmqvist, E., Hahn-Hägerdal, B., 2000. Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. *Bioresour. Technol.* 74, 25–33.
- Park, S., Baker, J.O., Himmel, M.E., Parilla, P. a, Johnson, D.K., 2010. Cellulose crystallinity index: measurement techniques and their impact on interpreting cellulase performance. *Biotechnol. Biofuels* 3, 10.
- Patel, G.B., Khan, A.W., Agnew, B.J., Colvin, J.R., 1980. Isolation and Characterization of an Anaerobic,

- Cellulolytic Microorganism, *Acetivibrio cellulolyticus* gen. nov., sp. nov. *Int. J. Syst. Bacteriol.* 30, 179–185.
- Pereyra, L.P., Hiibel, S.R., Prieto Riquelme, M. V, Reardon, K.F., Pruden, A, 2010. Detection and quantification of functional genes of cellulose- degrading, fermentative, and sulfate-reducing bacteria and methanogenic archaea. *Appl. Environ. Microbiol.* 76, 2192–202.
- Pernthaler, A., Pernthaler, J., Amann, R., 2002. Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* 68, 3094–101.
- Petitdemange, E., Caillet, F., Giallo, J., Gaudin, C., 1984. *Clostridium cellulolyticum* sp. nov., a cellulolytic, mesophilic: species from decayed grass. *Int. J. Syst. Evol. Microbiol.* 34, 155.
- Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6, 7–11.
- Pohlschröder, M., Canale-Parola, E., Leschine, S.B., 1995. Ultrastructural diversity of the cellulase complexes of *Clostridium papyrosolvens* C7. *J. Bacteriol.* 177, 6625–9.
- Pommier, S., Llamas, A.M., Lefebvre, X., 2010. Analysis of the outcome of shredding pretreatment on the anaerobic biodegradability of paper and cardboard materials. *Bioresour. Technol.* 101, 463–8.
- Puri, V.P., 1984. Effect of crystallinity and degree of polymerization of cellulose on enzymatic saccharification. *Biotechnol. Bioeng.* 26, 1219–22.
- Qu, X., Mazéas, L., Vavilin, V. a, Epissard, J., Lemunier, M., Mouchel, J.-M., He, P., Bouchez, T., 2009a. Combined monitoring of changes in $\delta^{13}\text{C}_4$ and archaeal community structure during mesophilic methanization of municipal solid waste. *FEMS Microbiol. Ecol.* 68, 236–45.
- Qu, X., Vavilin, V. a, Mazéas, L., Lemunier, M., Duquennoi, C., He, P., Bouchez, T., 2009b. Anaerobic biodegradation of cellulosic material: batch experiments and modelling based on isotopic data and focusing on acetoclastic and non-acetoclastic methanogenesis. *Waste Manag.* 29, 1828–37.
- Qureshi, N., Annous, B. a, Ezeji, T.C., Karcher, P., Maddox, I.S., 2005. Biofilm reactors for industrial bioconversion processes: employing potential of enhanced reaction rates. *Microb. Cell Fact.* 4, 24.
- Rakotoarivonina, H., Jubelin, G., Hebraud, M., Gaillard-Martinie, B., Forano, E., Mosoni, P., 2002. Adhesion to cellulose of the Gram-positive bacterium *Ruminococcus albus* involves type IV pili. *Microbiology* 148, 1871–80.
- Raman, B., Pan, C., Hurst, G.B., Rodriguez, M., McKeown, C.K., Lankford, P.K., Samatova, N.F., Mielenz, J.R., 2009. Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS One* 4, e5271.
- Ranjard, L., Poly, F., Lata, J.C., Mougél, C., Thioulouse, J., Nazaret, S., 2001. Characterization of bacterial and fungal soil communities by automated ribosomal intergenic spacer analysis fingerprints: biological and methodological variability. *Appl. Environ. Microbiol.* 67, 4479–87.

- Rincon, M.T., Ding, S., Mccrae, S.I., Martin, C., Aurilia, V., Lamed, R., Shoham, Y., Bayer, E.A., Flint, H.J., Martin, J.C., 2003. Novel Organization and Divergent Dockerin Specificities in the Cellulosome System of *Ruminococcus flavefaciens*. *Novel Organization and Divergent Dockerin Specificities in the Cellulosome System of Ruminococcus flavefaciens*.
- Rivers, D.B., Emert, G.H., 1988. Factors affecting the enzymatic hydrolysis of municipal-solid-waste components. *Biotechnol. Bioeng.* 31, 278–81.
- Rivière, D., Desvignes, V., Pelletier, E., Chaussonnerie, S., Guermazi, S., Weissenbach, J., Li, T., Camacho, P., Sghir, A., 2009. Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge. *ISME J.* 3, 700–14.
- Roger, V., Fonty, G., Komisarczuk-Bony, S., Gouet, P., 1990. Effects of Physicochemical Factors on the Adhesion to Cellulose Avicel of the Ruminal Bacteria *Ruminococcus flavefaciens* and *Fibrobacter succinogenes* subsp. *succinogenes*. *Appl. Environ. Microbiol.* 56, 3081–7.
- Romano, R.T., Zhang, R., Teter, S., McGarvey, J. a, 2009. The effect of enzyme addition on anaerobic digestion of Jose Tall Wheat Grass. *Bioresour. Technol.* 100, 4564–71.
- Ross, P., Mayer, R., Benziman, M., 1991. Cellulose biosynthesis and function in bacteria. *Microbiol. Rev.*
- Rubin, E.M., 2008. Genomics of cellulosic biofuels. *Nature* 454, 841–5.
- Russell, J.B., 1985. Fermentation of cellodextrins by cellulolytic and noncellulolytic rumen bacteria. *Appl. Environ. Microbiol.* 49, 572–6.
- Russell, J.B., Muck, R.E., Weimer, P.J., 2009. Quantitative analysis of cellulose degradation and growth of cellulolytic bacteria in the rumen. *FEMS Microbiol. Ecol.* 67, 183–97.
- Rychlik, J.L., May, T., 2000. The effect of a methanogen, *Methanobrevibacter smithii*, on the growth rate, organic acid production, and specific ATP activity of three predominant ruminal cellulolytic bacteria. *Curr. Microbiol.* 40, 176–80.
- Sára, M., Sleytr, U., 2000. S-layer proteins. *J. Bacteriol.* 182.
- Sasaki, D., Morita, M., Sasaki, K., Watanabe, A., Ohmura, N., 2012. Acceleration of cellulose degradation and shift of product via methanogenic co-culture of a cellulolytic bacterium with a hydrogenotrophic methanogen. *J. Biosci. Bioeng.* 114, 435–9.
- Schlüter, A., Bekel, T., Diaz, N.N., Dondrup, M., Eichenlaub, R., Gartemann, K.-H., Krahn, I., Krause, L., Krömeke, H., Kruse, O., Mussnug, J.H., Neuweiger, H., Niehaus, K., Pühler, A., Runte, K.J., Szczepanowski, R., Tauch, A., Tilker, A., Viehöver, P., Goesmann, A., 2008. The metagenome of a biogas-producing microbial community of a production-scale biogas plant fermenter analysed by the 454-pyrosequencing technology. *J. Biotechnol.* 136, 77–90.
- Schrank, K., Choi, B., Grund, S., 1999. *Treponema brennaborensis* sp. nov., a novel spirochaete isolated from a dairy cow suffering from digital dermatitis. *Int. J. ...* 43–50.
- Schuurman, T., de Boer, R.F., Kooistra-Smid, A.M.D., van Zwet, A.A., 2004. Prospective Study of Use of PCR Amplification and Sequencing of 16S Ribosomal DNA from Cerebrospinal Fluid for

- Diagnosis of Bacterial Meningitis in a Clinical Setting. *J. Clin. Microbiol.* 42, 734–740.
- Schwarz, W.H., 2001. The cellulosome and cellulose degradation by anaerobic bacteria. *Appl. Microbiol. Biotechnol.* 56, 634–49.
- Sharma, S.K., Mishra, I.M., Sharma, M.P., Saini, J.S., 1988. Effect of particle size on biogas generation from biomass residues. *Biomass* 17, 251–263.
- Shinkai, T., Ueki, T., Kobayashi, Y., 2010. Detection and identification of rumen bacteria constituting a fibrolytic consortium dominated by *Fibrobacter succinogenes*. *Anim. Sci. J.* 81, 72–79.
- Shoham, Y., Lamed, R., Bayer, E. a, 1999. The cellulosome concept as an efficient microbial strategy for the degradation of insoluble polysaccharides. *Trends Microbiol.* 7, 275–81.
- Silva, G.G.D., Couturier, M., Berrin, J.-G., Buléon, A., Rouau, X., 2012. Effects of grinding processes on enzymatic degradation of wheat straw. *Bioresour. Technol.* 103, 192–200.
- Silva, G.G.D., Rouau, S.G.X., 2011. Successive centrifugal grinding and sieving of wheat straw. *Powder Technol.* 208, 266–270.
- Singh, R., Paul, D., Jain, R.K., 2006. Biofilms: implications in bioremediation. *Trends Microbiol.* 14, 389–97.
- Sleat, R., Mah, R.A., Robinson, R., 1984. Isolation and Characterization of an Anaerobic, Cellulolytic Bacterium, *Clostridium cellulovorans* sp. nov. *Appl. Environ. Microbiol.* 48, 88–93.
- Song, H., Clarke, W.P., Blackall, L.L., 2005. Concurrent microscopic observations and activity measurements of cellulose hydrolyzing and methanogenic populations during the batch anaerobic digestion of crystalline cellulose. *Biotechnol. Bioeng.* 91, 369–78.
- Sorokin, D.Y., Panteleeva, a N., Tourova, T.P., Kaparullina, E.N., Muyzer, G., 2011. *Natronoflexus pectinivorans* gen. nov. sp. nov., an obligately anaerobic and alkaliphilic fermentative member of Bacteroidetes from soda lakes. *Extremophiles* 15, 691–6.
- Stahl, D., 1991. Development and application of nucleic acid probes. *Nucleic acid Tech. Bact. Syst.*
- Stams, a J., 1994. Metabolic interactions between anaerobic bacteria in methanogenic environments. *Antonie Van Leeuwenhoek* 66, 271–94.
- Stams, a J., Van Dijk, J.B., Dijkema, C., Plugge, C.M., 1993. Growth of syntrophic propionate-oxidizing bacteria with fumarate in the absence of methanogenic bacteria. *Appl. Environ. Microbiol.* 59, 1114–9.
- Stanton, T., Canale-Parola, E., 1980. *Treponema bryantii* sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. *Arch. Microbiol.* 6.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., 2005. Gene set enrichment analysis : A knowledge-based approach for interpreting genome-wide.
- Sukharnikov, L.O., Alahuhta, M., Brunecky, R., Upadhyay, A., Himmel, M.E., Lunin, V. V, Zhulin, I.B., 2012. Sequence, structure, and evolution of cellulases in glycoside hydrolase family 48. *J. Biol. Chem.* 287, 41068–77.

- Sukharnikov, L.O., Cantwell, B.J., Podar, M., Zhulin, I.B., 2011. Cellulases: ambiguous nonhomologous enzymes in a genomic perspective. *Trends Biotechnol.* 29, 473–479.
- Taherzadeh, M.J., Karimi, K., 2008. Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: a review. *Int. J. Mol. Sci.* 9, 1621–51.
- Tam, R., Saier, M.H., 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* 57, 320–46.
- Tamaru, Y., Miyake, H., Kuroda, K., Ueda, M., Doi, R.H., 2010. Comparative genomics of the mesophilic cellulosome-producing *Clostridium cellulovorans* and its application to biofuel production via consolidated bioprocessing. *Environ. Technol.* 31, 889–903.
- Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E. V, Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., Rao, B.S., Smirnov, S., Sverdlov, A. V, Vasudevan, S., Wolf, Y.I., Yin, J.J., Natale, D. a, 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4, 41.
- Tatusov, R.L., Galperin, M.Y., Natale, D. a, Koonin, E. V, 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–6.
- Teghammar, A., Yngvesson, J., Lundin, M., Taherzadeh, M.J., Horváth, I.S., 2010. Pretreatment of paper tube residuals for improved biogas production. *Bioresour. Technol.* 101, 1206–12.
- Teunissen, M.J., Kets, E.P., Op den Camp, H.J., Huis in't Veld, J.H., Vogels, G.D., 1992. Effect of coculture of anaerobic fungi isolated from ruminants and non-ruminants with methanogenic bacteria on cellulolytic and xylanolytic enzyme activities. *Arch. Microbiol.* 157, 176–82.
- Tokura, M., Ohkuma, M., Kudo, T., 2000. Molecular phylogeny of methanogens associated with flagellated protists in the gut and with the gut epithelium of termites. *FEMS Microbiol. Ecol.* 33, 233–240.
- Trinci, A.P.J., Davies, D.R., Gull, K., Lawrence, M.I., Bonde Nielsen, B., Rickers, A., Theodorou, M.K., 1994. Anaerobic fungi in herbivorous animals. *Mycol. Res.* 98, 129–152.
- Tsavkelova, E. a., Netrusov, a. I., 2012. Biogas Production from Cellulose Containing Substrates : A Review. *Appl Biochem Microbiol* 48, 421–433.
- Vallenet, D., Belda, E., Calteau, A., Cruveiller, S., Engelen, S., Lajus, A., Le Fevre, F., Longin, C., Mornico, D., Roche, D., Rouy, Z., Salvignol, G., Scarpelli, C., Thil Smith, A.A., Weiman, M., Medigue, C., 2013. MicroScope--an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res.* 41, D636–D647.
- van der Lelie, D., Taghavi, S., McCorkle, S.M., Li, L.-L., Malfatti, S. a, Monteleone, D., Donohoe, B.S., Ding, S.-Y., Adney, W.S., Himmel, M.E., Tringe, S.G., 2012. The metagenome of an anaerobic microbial community decomposing poplar wood chips. *PLoS One* 7, e36740.
- Veith, N., Feldman-Salit, A., Cojocar, V., Henrich, S., Kummer, U., Wade, R.C., 2013. Organism-adapted specificity of the allosteric regulation of pyruvate kinase in lactic acid bacteria. *PLoS Comput. Biol.* 9, e1003159.

- Vidal, B.C., Dien, B.S., Ting, K.C., Singh, V., 2011. Influence of feedstock particle size on lignocellulose conversion--a review. *Appl. Biochem. Biotechnol.* 164, 1405–21.
- Vivekanand, V., Olsen, E.F., Eijsink, V.G.H., Horn, S.J., 2013. Effect of different steam explosion conditions on methane potential and enzymatic saccharification of birch. *Bioresour. Technol.* 127, 343–9.
- Wakelin, J.H., Virgin, H.S., Crystal, E., 1959. Development and Comparison of Two X-Ray Methods for Determining the Crystallinity of Cotton Cellulose. *J. Appl. Phys.* 30, 1654.
- Wang, L., Zhang, Y., Gao, P., 2008. A novel function for the cellulose binding module of cellobiohydrolase I. *Sci. China. C. Life Sci.* 51, 620–9.
- Wang, Z.-W., Lee, S.-H., Elkins, J.G., Li, Y., Hamilton-Brehm, S., Morrell-Falvey, J.L., 2013. Continuous live cell imaging of cellulose attachment by microbes under anaerobic and thermophilic conditions using confocal microscopy. *J. Environ. Sci.* 25, 849–856.
- Wang, Z.-W., Lee, S.-H., Elkins, J.G., Morrell-Falvey, J.L., 2011. Spatial and temporal dynamics of cellulose degradation and biofilm formation by *Caldicellulosiruptor obsidiansis* and *Clostridium thermocellum*. *AMB Express* 1, 30.
- Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., Cayouette, M., McHardy, A.C., Djordjevic, G., Aboushadi, N., Sorek, R., Tringe, S.G., Podar, M., Martin, H.G., Kunin, V., Dalevi, D., Madejska, J., Kirton, E., Platt, D., Szeto, E., Salamov, A., Barry, K., Mikhailova, N., Kyrpides, N.C., Matson, E.G., Ottesen, E. a, Zhang, X., Hernández, M., Murillo, C., Acosta, L.G., Rigoutsos, I., Tamayo, G., Green, B.D., Chang, C., Rubin, E.M., Mathur, E.J., Robertson, D.E., Hugenholtz, P., Leadbetter, J.R., 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–5.
- Watanabe, H., 2001. Animal cellulases. *Cell. Mol. Life Sci.* 58, 1167– 1178.
- Watnick, P., Kolter, R., 2000. Biofilm, city of microbes. *J. Bacteriol.* 182, 2675.
- Weimer, P.J., Lopez-Guisa, J.M., French, a D., 1990. Effect of cellulose fine structure on kinetics of its digestion by mixed ruminal microorganisms in vitro. *Appl. Environ. Microbiol.* 56, 2421–9.
- Weimer, P.J., Price, N.P.J., Kroukamp, O., Joubert, L.-M., Wolfaardt, G.M., Van Zyl, W.H., 2006. Studies of the extracellular glycocalyx of the anaerobic cellulolytic bacterium *Ruminococcus albus* 7. *Appl. Environ. Microbiol.* 72, 7559–66.
- Weimer, P.J., Russell, J.B., Muck, R.E., 2009. Lessons from the cow: what the ruminant animal can teach us about consolidated bioprocessing of cellulosic biomass. *Bioresour. Technol.* 100, 5323–31.
- Wells, J.E., Russell, J.B., Shi, Y., Weimer, P.J., 1995. Cellodextrin efflux by the cellulolytic ruminal bacterium *Fibrobacter succinogenes* and its potential role in the growth of nonadherent bacteria. *Appl. Environ. Microbiol.* 61, 1757–62.
- Wolfenden, R., Lu, X., Young, G., 1998. Spontaneous hydrolysis of glycosides. *J. Am. Chem. Soc.* 120, 6814–6815.

- Wu, X.L., Conrad, R., 2001. Functional and structural response of a cellulose-degrading methanogenic microbial community to multiple aeration stress at two different temperatures. *Environ. Microbiol.* 3, 355–62.
- Wyman, C.E., Dale, B.E., Elander, R.T., Holtzapple, M., Ladisch, M.R., Lee, Y.Y., 2005. Coordinated development of leading biomass pretreatment technologies. *Bioresour. Technol.* 96, 1959–66.
- Wyss, C., Choi, B.K., Schüpbach, P., Moter, A., Guggenheim, B., Göbel, U.B., 1999. *Treponema lecithinolyticum* sp. nov., a small saccharolytic spirochaete with phospholipase A and C activities associated with periodontal diseases. *Int. J. Syst. Bacteriol.* 49 Pt 4, 1329–39.
- Xiao, B., Sun, X., Sun, R., 2001. Chemical, structural, and thermal characterizations of alkali-soluble lignins and hemicelluloses, and cellulose from maize stems, rye straw, and rice straw. *Polym. Degrad. Stab.* 74, 307–319.
- Xiao, W., Clarkson, W.W., 1997. Acid solubilization of lignin and bioconversion of treated newsprint to methane 61–66.
- Xu, C., Huang, R., Teng, L., Wang, D., Hemme, C.L., Borovok, I., He, Q., Lamed, R., Bayer, E. a, Zhou, J., Xu, J., 2013. Structure and regulation of the cellulose degradome in *Clostridium cellulolyticum*. *Biotechnol. Biofuels* 6, 73.
- Xu, Q., Gao, W., Ding, S.-Y., Kenig, R., Shoham, Y., Bayer, E.A., Lamed, R., 2003. The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. *J. Bacteriol.* 185, 4548–57.
- Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–51.
- Yu, Y., Lee, C., Kim, J., Hwang, S., 2005. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnol. Bioeng.* 89, 670–9.
- Yuan, X., Cao, Y., Li, J., Wen, B., Zhu, W., Wang, X., Cui, Z., 2012. Effect of pretreatment by a microbial consortium on methane production of waste paper and cardboard. *Bioresour. Technol.* 118, 281–8.
- Yuan, X., Shi, X., Zhang, P., Wei, Y., Guo, R., Wang, L., 2011. Anaerobic biohydrogen production from wheat stalk by mixed microflora: kinetic model and particle size influence. *Bioresour. Technol.* 102, 9007–12.
- Zhao, B., Chen, S., 2012. *Alkalitalea saponilacus* gen. nov., sp. nov., an obligately anaerobic, alkaliphilic, xylanolytic bacterium from a meromictic soda lake. *Int. J. Syst. Evol. Microbiol.* 62, 2618–23.
- Zhao, X., Zhang, L., Liu, D., 2012. Biomass recalcitrance . Part I : the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose 465–482.
- Zhu, L., O’Dwyer, J.P., Chang, V.S., Granda, C.B., Holtzapple, M.T., 2008. Structural features affecting biomass enzymatic digestibility. *Bioresour. Technol.* 99, 3817–28.

Résumé

La méthanisation des composés lignocellulosiques présente un fort intérêt en raison de leur haut potentiel énergétique et de leur abondance, notamment dans les ordures ménagères résiduelles. Toutefois, leur complexité de structure et de composition rend ces matériaux difficilement dégradables en conditions anaérobies et l'utilisation de prétraitements est généralement requise afin d'améliorer leurs rendements de biodégradation. Outre l'effet de ces prétraitements sur la biodégradation de ces composés, la colonisation des lignocelluloses par les micro-organismes cellulolytiques est une étape clé pour l'efficacité de sa dégradation. Dans ce cadre, le travail de thèse a pour objectifs de mieux comprendre le déterminisme de la colonisation de déchets, d'établir le lien entre la colonisation des déchets lignocellulosiques et l'efficacité de leur dégradation et enfin de caractériser plus finement les mécanismes et interactions mises en jeu au sein de la biomasse. Afin de répondre à ces questions, une approche transversale a été développée, combinant des modèles de cultures de souches pures et des systèmes de méthanisation en laboratoire par des communautés complexes. Des approches intégratives ont été appliquées à l'étude de ces systèmes, couplant des analyses haut-débit (métagénomique/(méta)protéomique), un suivi physico-chimique de la biodégradation et des caractérisations physico-chimiques des composés lignocellulosiques étudiés. L'ensemble des résultats met en évidence le rôle des propriétés chimiques et micro- et macro-structurales des composés lignocellulosiques dans leur récalcitrance, leur performances de dégradation et la réponse du compartiment microbien. La réalisation de la première étude de protéomique totale et quantitative sur la souche pure cellulolytique *Clostridium cellulolyticum*, modèle des *Clostridia* cellulolytiques mésophiles, a permis de mettre en évidence que la vitesse maximale de biodégradation du mouchoir en papier est supérieure à celle du coton et que cette dégradation est associée à un profil métabolique particulier, à une colonisation plus rapide et plus étendue et à une modulation quantitative du système cellulosique. D'autre part, une étude sur un système plus réaliste pour l'étude de la méthanisation des déchets lignocellulosiques a confirmé la bonne concordance entre ce système et le système modèle utilisé et a également permis de mettre en évidence les effets substrats sur la structure des communautés microbiennes avec la dominance de la classe *Bacteroidia* en présence de mouchoir en papier et la forte proportion de la classe *Spirochaetes* en présence de coton. Enfin l'étude des effets de broyages très fins de la paille de blé et du carton plat ont mis en évidence les limites de ces prétraitements sur les performances de leur dégradation, avec l'effet positif modéré du broyage fin de la paille. Ils ont également montré la sensibilité des communautés microbiennes aux changements de surface du substrat, qui se manifeste par l'émergence de communautés parfois différentes en fonction du prétraitement mécanique appliqué. En conclusion, ce travail a permis de traiter sous un angle nouveau les questions liées à la récalcitrance des déchets lignocellulosiques en abordant à la fois les aspects structuraux, écologiques et fonctionnels. Ces résultats alimentent le corps de connaissances fondamentales sur les bioprocédés. Ils confirment que les matériaux lignocellulosiques sont particuliers parmi les déchets non-dangereux et qu'une exploitation plus large de leur potentiel énergétique nécessiterait la mise en œuvre de procédés adaptés.

Abstract

Lignocellulosic materials have a high energy potential and are abundant, especially in municipal solid waste and their methanization is a promising waste-to-energy bioprocess. However, owing to their highly complex and heterogeneous structure, they are recalcitrant to anaerobic conditions and the use of pre-treatments is usually required to improve their biodegradation yields. Besides, lignocellulose colonization by cellulolytic microorganisms is a key step for an efficient biodegradation. In this context, the PhD work aimed to better understand the factors affecting waste colonization, to establish the link between lignocellulosic waste colonization and its biodegradation efficiency and to characterize more precisely the mechanisms and interactions within the biomass. A transversal approach was developed, combining cultures of model pure strains and lab-scale methanization microcosms with a complex biomass. Integrated approaches were applied to these studies, combining high-throughput analyses (metagenomics/(meta) proteomics), physico-chemical monitoring of bioconversion and finally physico-chemical characterization of substrates. The main results highlight the important role of lignocellulosic materials chemical and micro- and macro-structural features for their recalcitrance, their biodegradation efficiency and the response of the microbial compartment. The first global quantitative proteomic study on the cellulolytic model *Clostridium cellulolyticum* was conducted. Results showed an increased biodegradation rate of the facial tissue compared to cotton. This enhanced biodegradation was associated to a particular metabolic profile, a faster and more extensive colonization and finally a quantitative modulation of the cellulosic system. On the other hand, study of lignocellulosic waste methanization confirmed the good agreement between this more realistic system and the above-described model system. It also provided new information about the effects of substrate on microbial community structure. Noticeably, *Bacteroidia* members predominated in the presence of tissue and a high proportion of *Spirochaetes* members was observed in the presence of cotton. Finally, study of the effects of wheat straw and cardboard dry grinding revealed the limitations of these pretreatments on biodegradation efficiency. Main key points were a moderate positive effect of wheat straw fine grinding, and the sensitivity of the microbial communities to substrate surface characteristics, as evidenced by the emergence of different microbial communities according to the applied mechanical pretreatment. In conclusion, this work brings new perspectives to the study of lignocellulosic waste recalcitrance by addressing both the structural, functional and ecological aspects. These results contribute to the core fundamental knowledge on bioprocesses. They confirm that the lignocellulosic materials are specific among non- hazardous waste and require the implementation of adapted processes.