# Microbial and metabolic diversity of anaerobic D-galacturonate fermentation

Valk, Laura

**DOI**
10.4233/uuid:b34efe06-28d4-4f4c-9fe1-9d2b1696f111

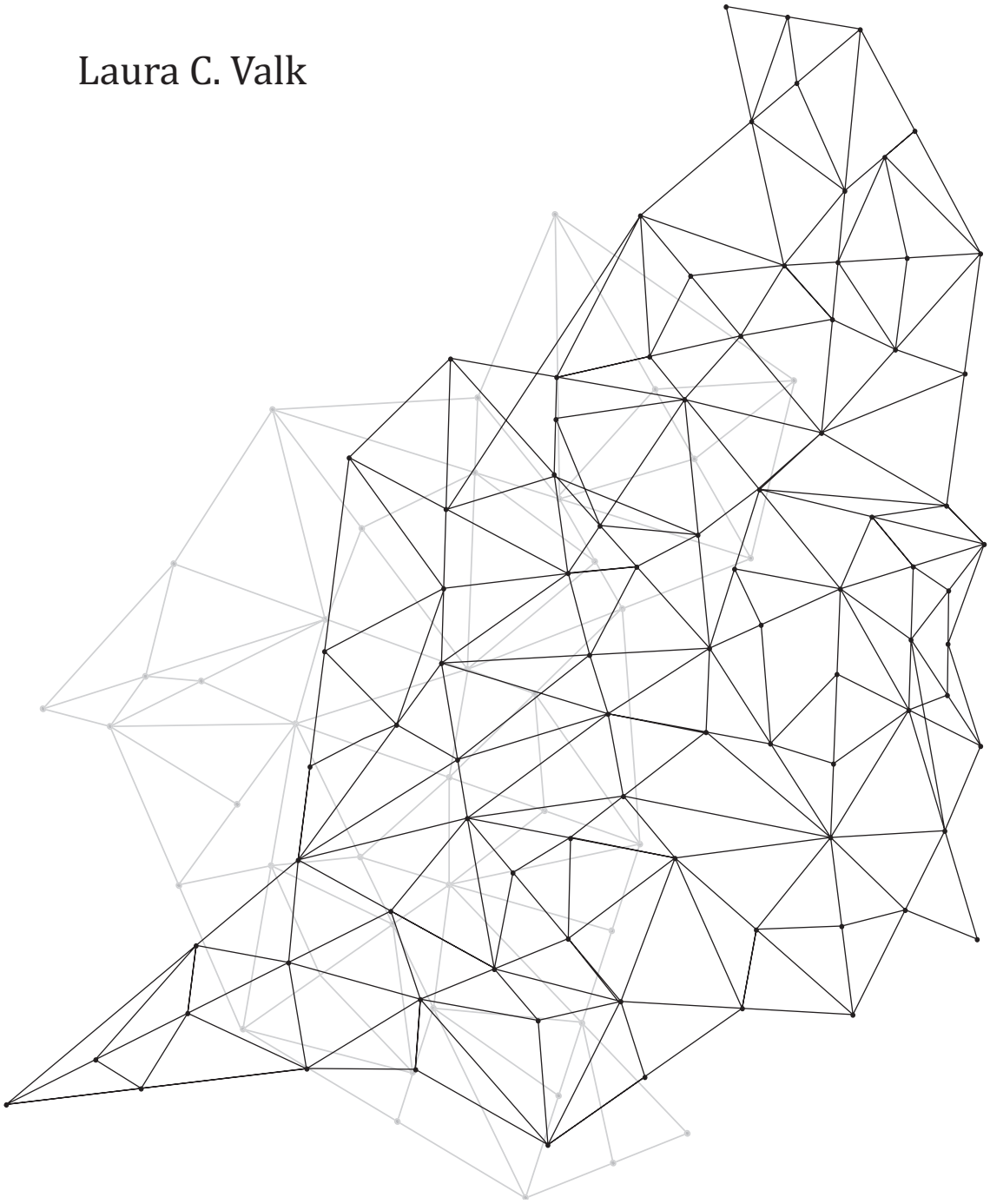
**Publication date**
2020

**Document Version**
Final published version

**Citation (APA)**
Valk, L. (2020). *Microbial and metabolic diversity of anaerobic D-galacturonate fermentation*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:b34efe06-28d4-4f4c-9fe1-9d2b1696f111

# Microbial and metabolic diversity of anaerobic D-galacturonate fermentation

Laura C. Valk



# Microbial and metabolic diversity of anaerobic D-galacturonate fermentation

## Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
**27 februari 2020 om 12:30 uur**

door

**Laura Christina VALK**

Ingenieur in Life Science and Technology
Technische Universiteit Delft, Nederland
geboren te Broek in Waterland, Nederland

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

| | |
|---|---|
| Rector magnificus | voorzitter |
| Prof. dr. J.T. Pronk | Technische Universiteit Delft, promotor |
| Prof. dr. dr. h.c. ir. M.C.M. van Loosdrecht | Technische Universiteit Delft, promotor |

Onafhankelijke leden:

| | |
|---|---|
| Prof. dr. ir. M.S.M. Jetten | Radboud Universiteit Nijmegen |
| Prof. dr. D. Machado de Souza | Wageningen University and Research |
| Prof. dr. L. Dijkhuizen | Rijksuniversiteit Groningen |
| Prof dr. F. Hollman | Technische Universiteit Delft |
| Prof. dr. P. A. S. Daran-Lapujade | Technische Universiteit Delft, reservelid |

Overig lid:

| | |
|---|---|
| Prof. dr. ir. A.J.A van Maris | KTH Royal Institute of Technology, Sweden |

The research presented in this thesis was performed at the Environmental Biotechnology and Industrial Microbiology sections, Biotechnology department, Applied Sciences faculty Delft University of Technology, The Netherlands.

This research was supported by the SIAM Gravitation Grant 024.002.002 from the Netherlands Organization for Scientific Research, within the Soehngen Institute of Anaerobic Microbiology.

ISBN 978-94-6366-241-3

An electronic version of this dissertation is available at

http://repository.tudelft.nl

# Table of contents

Voor mama

*To infinity and beyond*

Buzz Lightyear
Toy Story (Walt Disney, 1995)

# Summary

Over the past decade the demand for technology which can make our global economy more sustainable has increased exponentially. In the field of biotechnology, one of the many advancements is focussed on the conversion of agricultural waste streams towards commodity chemicals. Plant biomass can be an environmental and sustainable alternative for the use of petrochemical substrates, with a special interest towards processes which use plant biomass in waste streams rather than currently used feedstocks, such as corn and sugarcane molasses that are competing with food production.

In **Chapter 1** the sociological relevance of developing the underlying technology is discussed with respect to currently conventional petrochemical-based production of fuels and commodity chemicals that cause major environmental issues due to pollution and excess greenhouse gas production. Plant biomass streams are primarily abundant in hemicellulose and cellulose but often additionally contain a large fraction of pectin. Pectin is a complex polymer that can be degraded via commercially available enzymatic cocktails of pectic enzymes, currently already used in the beverage and clothing industry. The hydrolysis of pectin releases its major constituent, D-galacturonate, which can be further convert via distinct catabolic pathways. However due to the oxidized nature of D-galacturonate, compared to staple substrates such as hexose or pentose sugars, the known catabolic pathways require input of redox co-factors under fermentative conditions to produce industrially relevant products and the majority of the pathways also do not yield ATP through substrate level phosphorylation. Consequently, production of industrially relevant products is not possible under anaerobic conditions. Enrichment-based selection strategies have previously helped identify many novel and interesting microbial conversions for other substrates. Additionally, recent advancement of omics-strategies and whole genome sequencing techniques have enabled much deeper insights in the genomic potential of microbial populations. The goal of the PhD research project described in this thesis was to explore metabolic and microbial diversity of anaerobic D-galacturonate metabolism using a combination of enrichment-based selection strategies, and meta-omics analysis.

**Chapter 2** describes the effect of the oxidized D-galacturonate substrate on anaerobic microbial enrichment cultures and their catabolic product profiles. Two anaerobic chemostat enrichments were setup as replicates, with D-galacturonate as sole limiting carbon source. The inoculum used was derived from pectin-rich environments, cow rumen content and rotting orange peels, and the enrichment yielded stable microbial communities. The cultures were dominated by a novel *Lachnospiraceae* species, for which the name "*Candidatus* Galacturonibacter soehngenii" was proposed and a side population of a *Klebsiella oxytoca*. The dominant catabolic product was acetate, with formate and $H_2$ as co-products. However, the expected equimolar amount of acetate and formate and or

hydrogen was not observed. This suggested the presence of acetogenesis which caused part of the hydrogen and $CO_2$ produced in the fermentation to be converted into acetate via the Wood-Ljungdahl pathway (WLP). Indeed, metagenomic analysis of the enrichment cultures indicated that the genome of "*Candidatus* G. soehngenii" contained genes of the fermentative adapted Entner-Doudoroff pathway for D-galacturonate catabolism as well a partial Wood-Ljungdahl pathway. Concurrent expression of these pathways could lower the residual D-galacturonate concentration and enable a higher specific ATP production rate than would be possible if D-galacturonate was converted solely via the fermentative route.

In **Chapter 3** the acetogenic catabolism of *"Candidatus* Galacturonibacter soehngenii" is further explored. In this study, $NaH^{13}CO_3$ fed to chemostat-grown, galacturonate-limited enrichment cultures of *Ca.* G. soehngenii was shown to be incorporated into acetate. The carbon labelling experiment also showed preferential labelling of the carbonyl-group of acetate which is consistent with a WLP which derived the methyl group from formate produced by fermentation and carbonyl from extracellular $CO_2$. This analysis was corroborated by high enzymatic activities in the cell extract of *ca*. G. soehngenii with CO and high transcript levels of putative pyruvate-formate lyase genes and low transcript levels of a candidate formate dehydrogenase. Long-read nanopore sequencing techniques were deployed to reassemble the MAG of "*Ca*. G. soehngenii", which was also devoid of canonical CODH/ACS complex genes or their homologues. However, high CO-dehydrogenase activities were measured in cell extracts of Ca. G. soehngenii enrichment cultures, contradicting the absence of corresponding homologues in the MAG. A novel anaerobic Ni-CO dehydrogenase candidate was identified based on the highly conserved amino-acid motif associated with this protein-type. These results demonstrate operation of an acetogenic pathway, most probably as a yet unresolved variant of the Wood-Ljungdahl pathway, in anaerobic, D-galacturonate-limited cultures of "*Ca*. G. soehngenii".

As stated in **Chapter 1** in currently known microorganisms able to ferment D-galacturonate, the catabolism occurs via the galacturonate isomerase or adapted Entner-Doudoroff pathway. Redox-cofactor balancing in this isomerase pathway constrains the possible range of products generated from anaerobic D-galacturonate fermentation, with acetate as the predominant organic fermentation product. **Chapter 4** discusses anaerobic enrichment experiments which were performed on this carbon source were at pH 4, to explore the microbial and metabolic diversity on this substrate when the production of acetate was deterred, resulting in the enrichment of a *Lactobacillus* species. Subsequent isolation and whole genome analysis showed this organism to be a *Lactobacillus suebicus* strain. Characterisation of its physiology in bioreactors under chemostat and batch operation showed an approximate equimolar production of lactate and acetate from D-galacturonate. A combination of whole-genome sequence analysis, quantitative proteomics, enzyme activity assays in cell extracts and *in vitro* product identification demonstrated that D-galacturonate metabolism in D-galacturonate-grown *L. suebicus* occurs via a novel pathway. In this pathway, mannonate generated by the initial reactions

of the canonical isomerase pathway is converted to 6-phosphogluconate by the concerted action of two previously unknown biochemical reactions; a mannonate kinase and a 6-phosphomannonate 2-epimerase. 6-phospholgluconate is then further catabolised via known reactions of the phosphoketolase pathway. This novel pathway enables the production of ribulose-5-phosphate from D-galacturonate redox co-factor neutral, in contrast to the classical isomerase pathway for D-galacturonate catabolism. Indeed, this redox-cofactor neutral production of an intermediate of the pentose phosphate pathway could enable metabolic engineering of microbial cell factories for the production of added-value productions, e.g. ethanol or lactate, from pectin-rich feedstocks. Further research will be required to identify the structural genes which encode the key enzymes for this hybrid pathway. This study illustrates the potential of microbial enrichment cultivation to identify novel pathways for the conversion of environmentally and industrially relevant compounds.

Methods used in classical enzymology, such as whole cell protein fractionation with chromatography or gel-based techniques, might enable further exploration of the unconventional acetogenesis discussed in **Chapter 2** and **3** and identification of the novel genes discussed in **Chapter 4**. Additionally, other enrichment strategies than used in this thesis could enable the discovery of as of yet undiscovered D-galacturonate metabolisms. Ultimately these methods could allow further exploration of the microbial metabolic diversity on D-galacturonate fermentation. With enrichment-based strategies potentially providing more currently unknown pathways.

# Samenvatting

De vraag naar technologie die de economie wereldwijd kan verduurzamen is gedurende het afgelopen decennium exponentieel gestegen. De focus op omzetting van organische afvalstromen in industrieel relevante producten in de biotechnologie is hier een voorbeeld van. Plantenbiomassa kan een milieuvriendelijk en duurzaam alternatief zijn voor het gebruik van petrochemische producten. Hierbij is er speciale interesse voor het gebruik van plantenbiomassa afkomstig van reststromen, in plaats van het nu veel voorkomende gebruik van gewassen zoals mais en suikerrietmelasse, dat mogelijk ten koste zou kunnen gaan van de wereldvoedselproductie.

**Hoofdstuk 1** beschrijft de socio-ecologische relevantie van het ontwikkelen van de hiervoor benodigde technologieën, vanwege de grote problemen voor milieu en klimaat die veroorzaakt worden door de huidige productie van brandstof en essentiële chemicaliën vanuit petrochemische grondstoffen. Biomassa van planten kan gebruikt worden als milieuvriendelijk en duurzaam alternatief voor petrochemische substraten, waarbij het gebruik van plantaardige biomassa uit afvalreststromen uit de landbouw veruit de voorkeur heeft. Deze reststromen zijn niet alleen rijk aan hemicellulose en cellulose, maar ook aan pectine. Pectine is een complex polymeer, dat afgebroken kan worden met behulp van commercieel beschikbare enzymmengsels die al gebruikt worden in de vruchtensap- en kledingindustrie. Deze afbraak leidt tot het vrijkomen van het hoofdbestanddeel van het pectine-polymeer, D-galacturonzuur. Meerdere microbiële routes kunnen D-galacturonzuur omzetten naar andere producten. Echter, door de oxidatieve aard van D-galacturonzuur, in tegenstelling tot bijvoorbeeld substraten zoals hexose- of pentosesuiker, vinden deze omzettingen alleen plaats bij toevoeging van redox-equivalenten onder fermentatieve condities, daarnaast produceren ze vaak geen ATP via substraat level fosforylering. Derhalve is het niet mogelijk om de gewenste producten op industriële wijze onder anaerobe condities te produceren. Op verrijking gebaseerde selectiestrategieën hebben in het verleden geholpen bij het identificeren van vele nieuwe en interessante omzettingen. Daarnaast hebben recente 'meta-omics'-strategieën en het 'sequencen' van volledige genomen voor diepere inzichten gezorgd in het genetische potentieel van microbiële populaties. Het doel van het promotieonderzoek beschreven in dit proefschrift was het nader verkennen van de microbiële en metabole diversiteit van het anaerobe D-galacturonzuurmetabolisme met behulp van verrijkingsstrategieën en meta-omicsanalyse.

**Hoofstuk 2** beschrijft het effect van het geoxideerde substraat, D-galacturonzuur, op anaerobe verrijkingsculturen en de bijbehorende katabole productprofielen. Een anaerobe chemostaat verrijkingscultuur was in tweevoud opgezet, met D-galacturonzuur als enige koolstofbron. Het gebruikte entmateriaal was afkomstig van pectinerijke

milieus, namelijk de inhoud van een koeien pens en rottende sinaasappelschillen, en leverde na verrijking een stabiele microbiële gemeenschap op. De verrijkingsculturen werden gedomineerd door een nieuwe *Lachnospiraceae* soort, waarvoor de naam "*Candidatus* Galacturonibacter soehngenii" is gekozen. Ze bevatten ook een zijpopulatie die werd gedomineerd door *Klebsiella oxytoca*. Het dominante katabole product was acetaat met formiaat en waterstof ($H_2$) als nevenproducten. De gelijke molaire verhouding van acetaat op formiaat en $H_2$ werd echter niet waargenomen. Dit suggereerde de aanwezigheid van acetogenese, wat ervoor zorgt dat de in de fermentatie geproduceerde waterstof en kooldioxide worden omgezet in acetaat via de Wood-Ljungdahl route (WLP). Dit werd inderdaad bevestigd in de analyse van het metagenoom van de verrijkingscultuur, die de indicatie gaf dat in het genoom van "*Ca*. G. Soehngenii" genen zaten die codeerden voor zowel de aangepaste Entner-Doudoroff route voor galacturonzuurfermentatie als voor een deel van de WLP. Gelijktijdige expressie van deze twee routes zorgt zowel voor een lagere residuele concentratie van galacturonzuur als voor een hogere ATP-productiesnelheid dan mogelijk zou zijn geweest als D-galacturonzuur alleen was gefermenteerd.

In **hoofdstuk 3** wordt het acetogenese katabolisme van de "*Candidatus* Galacturonibacter soehngenii" verder geëxploreerd. In deze studie werd $NaH^{13}CO_3$ gevoed aan een in een chemostaat opgekweekte galacturonzuur-gelimiteerde verrijkingscultuur van "*Ca*. G. Soehngenii", waarbij gelabeld acetaat werd aangetoond. Daarnaast toonde het labelingsexperiment ook aan dat de carbonylgroep van het acetaat preferentieel gelabeld werd. Dit kan verklaard worden doordat binnen de WLP route een methylgroep van formiaat wordt geproduceerd via de fermentatie route, en een carbonylgroup van extracellulair $CO_2$. Deze analyse werd bevestigd door de hoge enzymatische activiteiten in het celextract met koolmonoxide en een hoog transcriptieniveau van mogelijke pyruvaat-formiaat-lyasegenen en lage transcriptieniveaus van een kandidaat-formiaatdehydrogenase. 'Long-read nanopore sequencing'-technieken zijn ingezet om het vanuit het metagenoom geassembleerde genoom (MAG) van "*Ca*. G. Soehngenii" te her-assembleren. Ook dit genoom was verstoken van het canonieke CODH/ACS-complex. Echter, de hoge CO-dehydrogenase activiteiten die gemeten waren in het cel vrije extract was in tegenspraak met de afwezigheid van corresponderende homologen. Een nieuwe anaerobe Ni-CO dehydrogenase kandidaat werd geïdentificeerd op basis van een sterk geconserveerd aminozuur-motief geassocieerd met dit eiwit-type. Deze resultaten demonstreren de activiteit van een acetogenese route, hoogst waarschijnlijk een tot nu toe nog onopgehelderde variant van de Wood-Ljungdahl route, in een anaerobe D-galacturonzuur-gelimiteerde cultuur van "*Ca*. G. Soehngenii".

Bij de tot op heden bekende micro-organismen die D-galacturonzuur kunnen fermenteren, zoals aangegeven in **hoofdstuk 1**, gebeurt dit altijd via de D-galacturonzuur-isomeraseroute, ook wel de aangepaste Entner-Doudoroff route genoemd. Het in evenwicht houden van de redox cofactor in deze isomerase route beperkt de omvang van mogelijke producten die gegenereerd kunnen worden via D-galacturonzuurfermentatie.

Hierbij is acetaat het voornaamste organische fermentatie product. **Hoofstuk 4** bediscussieert de anaërobe verrijkingen op D-galacturonzuur bij lage pH. Deze waren opgezet om de microbiële en metabole diversiteit op dit substraat te onderzoeken en resulteerde in een verrijking van een melkzuurbacterie, een *Lactobacillus* soort. Opeenvolgende isolatie en analyse van de sequentie van het hele genoom liet zien dat dit een *Lactobacillus suebicus* stam was; karakterisering van de fysiologie bij chemostaat- en batchkweken liet een equimolaire productie van lactaat en acetaat vanuit D-galacturonzuur zien. Een combinatie van volledige genoomanalyse, kwantitatieve proteomics, enzymactiviteits-assays in cel extracten en in vitro product identificatie demonstreerde dat het D-galacturonzuur metabolisme in de *L. suebicus* stam tot op heden onbekend was. In deze route wordt het uit de initiële omzettingen van de canonieke isomerase-route gegenereerde mannonaat omgezet naar 6-fosfogluconaat door de gecoördineerde actie van twee tot op heden onbekende biochemische reacties: een met het enzym mannonaatkinase en een met 6-fosfomannonaat-2-epimerase. Vervolgens wordt 6-fosfogluconaat verder gekataboliseerd via de bekende reacties van de fosfoketolase-route. Deze nieuwe route maakt de redox-neutrale productie van ribulose-5-fosfaat vanuit D-galacturonzuur mogelijk, in tegenstelling tot de klassieke isomerase route voor D-galacturonzuur katabolisme. Deze redox-neutrale productie van een intermediair van de pentosefosfaat-route is hoogst interessant voor 'metabolic engineering' toepassingen voor het maken van microbiële celfabrieken voor producten met meerwaarde, e.g. ethanol of lactaat, vanuit pectinerijke grondstoffen. Vervolgonderzoek zal nodig zijn om de structurele genen, die coderen voor de sleutelenzymen van deze hybride route, te identificeren. Deze studie illustreert het potentieel van microbiële verrijkingen voor het identificeren van nieuwe routes voor de omzettingen van milieuvriendelijke en duurzame, industrieel relevante producten.

Methoden gebruikt in de klassieke enzymologie, zoals fractionering van alle eiwitten van de cel met chromatografie of op gel gebaseerde technieken, zijn ook erg geschikt voor verdere exploratie van de onconventionele acetogenese bediscussieerd in **hoofdstuk 2** en **3**. Deze methoden zouden ook toepasbaar zijn voor de identificatie van de nieuwe genen bediscussieerd in **hoofdstuk 4**. Daarnaast zal het over een lange termijn verrijken van culturen in gecontroleerde bioreactoren kunnen zorgen voor een beter begrip van microbiële processen in de veranderende natuurlijke omgeving, b.v. als gevolg van klimaatverandering. Verdere exploraties van de microbiële en metabole diversiteit van D-galacturonzuurfermentatie met behulp van verrijkingen onder zorgvuldig geselecteerde omstandigheden zullen naar verwachting leiden tot de ontdekking van nog meer tot op heden onbekende metabolische routes.

# Chapter 1 | General introduction

# 1 Towards a bio-refinery by using agricultural waste streams as substrates for industrial production of (bulk) chemicals

Crude oil is the still the main raw material for production of transport fuels and commodity products in the chemical industry. However, using crude oil comes at a heavy cost in terms of the environmental damage that occurs during mining and refining of the petrochemical products. Moreover, release of carbon dioxide produced from fossil resources during their utilization by our societies has become a critical point of concern on a global scale (1, 2). To prevent escalation of the climate crisis requires transition to a truly circular sustainable economy. For such a transition, introduction of alternative, renewable raw materials is of paramount importance (2, 3).

The current atmospheric accumulation of the greenhouse gas $CO_2$ is primarily caused by the hugely different time constants of the conversion of fossil resources into $CO_2$ and those of their formation through fossilisation of plants and micro-algal lipids (Figure 1A; (4)). While the process of capturing atmospheric $CO_2$ in organic matter and subsequent sedimentation and conversion to reserves of natural gas, oil and coal took millions of years, mankind is releasing large fractions of these reserves as $CO_2$ at a fast and increasing rate. The resulting imbalance in the carbon cycle (2, 5) can, in principle, be avoided by directly converting plant biomass into usable substrates for the production of commodity products, thus bypassing their conversion to fossil reserves (Figure 1B). Micro-organisms are capable to convert carbohydrates into many of the products that are currently produced by the petrochemical industry (3, 6–8). These carbohydrates can be released from plant biomass by enzymatic hydrolysis of plant polymers (e.g. corn starch) or sugar oligomers (e.g. cane sugar) and used as substrates for the production of commodity chemicals (9). Already in the early 20$^{th}$ century, A.J. Kluyver (10) predicted such renewable carbohydrates would become important raw materials for the production of chemicals.

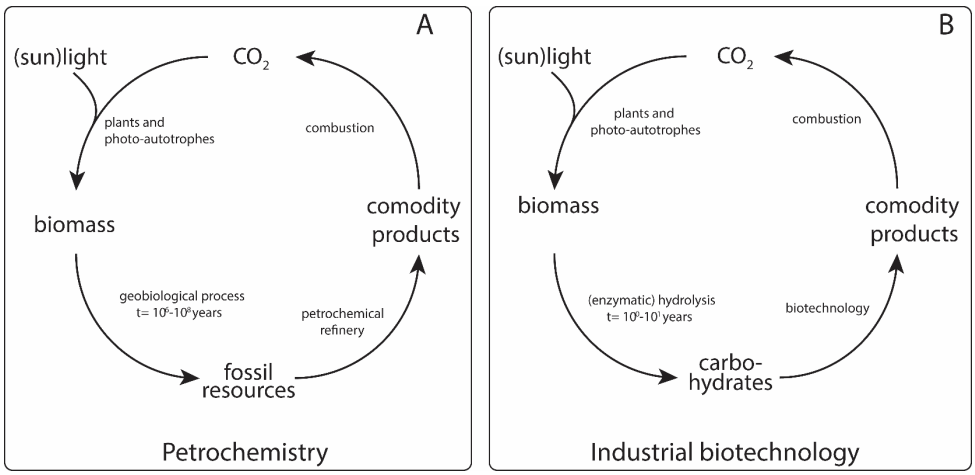



**Figure 1 |** The carbon cycle of the material streams of the petro-chemistry (left) and industrial biotechnology (right). *Adapted from* (4).

In order to be cost-competitive with the petrochemical industry and to prevent competition with global food supply chains (e.g. by requiring valuable resources such as water, fertilizer and/or arable land), carbohydrates for production of transport fuels and chemicals should preferably be derived from non-edible plant biomass (11–13). A solution that has been intensively explored is the use of large carbohydrate streams derived from parts of plants such as stems, leaves and other agricultural residues, which are generally rich in lignocellulose and hemicellulose polymers (14–16). Cellulose and hemicellulose can be degraded into their sugar monomers D-glucose, D-xylose and L-arabinose via enzymatic hydrolysis, steam combustion and/or mechanic shearing (17–20). The monomers can subsequently be used by microorganisms which, often after genetic modification, convert them to commodity chemicals (21–25).

## Investigation into pectin, an abundant but less intensively studied feedstock

For economically viable utilization of agricultural waste-streams, full utilization of all available carbohydrates is essential (11, 19). Besides cellulose and hemicellulose, plant biomass contains pectin as an abundant polymer. Consisting of a group of four complex polymers (Figure 2), pectin is abundantly present in the plant cell wall, middle lamellae and soft plant tissues such as fruit peels (26). Unlike cellulose and hemicellulose, the common denominator between the four pectin-types is their high abundance of a specific monomer, D-galacturonate, which is not found in cellulose or hemicellulose. D-galacturonate is the main constituent of the backbone of pectin polymers and accounts for up to 70 % of the pectin-monomers (26).

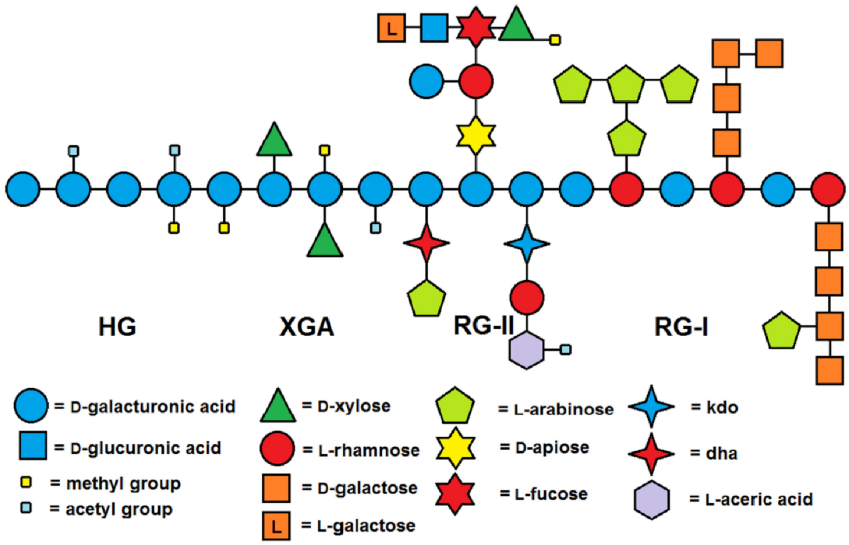



**Figure 2 |** Pectin structure with homogalacturonan (HG), xylogalacturonan (XGA), rhamnogalacturonan II (RG-II) and rhamnogalacturonan I (RG-I). Figure from (27).

In order to release D-galacturonate monomers, pectin can be hydrolysed chemically by boiling it in 2.5 % sulfuric acid (28) or, alternatively, by using defined enzyme cocktails. Collectively, pectin-degrading enzymes are called pectic enzymes or pectinases (29–32). Cocktails of these enzymes can be formulated to break down the complex structure of the specific pectin polymers. Four pectin-polymer types are distinguished based on distinctive structural characteristics: homogalacturonan (HG), xylogalacturonan (XGA), rhamnogalacturonan I (RG-I) and rhamnogalacturonan II (RG-II). These four types can be divided into two sub-groups based on their side groups; smooth pectin (HG) and hairy pectin (XGA, RG-I and RG-II). Smooth pectin is the most commonly occurring type of pectin and represents approximately 65 % of all pectic polymers (26). In smooth pectin, the homogalacturonan is α-1,4-linked and partially methyl-esterified at the C-6 carbonyl group and may be acetylated at the O-2 and O-3 position (33). Smooth pectin galacturonan can span uninterrupted for up to approximately 100 galacturonate residues between the other pectin-polymer types (34). The pectic backbone of XGA and RG-II also consists of methylated and acetylated α-1,4-linked-D-galacturonate moieties, which are additionally linked to side groups of predominantly L-rhamnose and other neutral sugars, such as L-arabinose and D-galactose. The backbone of RG-I consists of alternating α-1,4-linked-D-galacturonate and α-1,2-linked-L-rhamnose (Figure 2, (26, 27)).

Pectic enzymes and formulations are commercially available and used in the food and beverage industry for e.g. clarification of fruit juice or coffee and tea fermentation and in the clothing industry for retting and gumming of fibre crops (35, 36). Depending on the

type of pectin that needs to be degraded, pectinase formulations either contain polygalacturonases and rhamnogalacturonases, or polygalacturonan and rhamnogalacturonan hydrolases. These hydrolases are exo- and endo-acting enzymes, with the endo-acting enzymes acting on the non-reducing end of the polymer and the endo-acting enzymes acting on the α-1,4-bonds between the monomers (31, 37, 38). The other subgroup of enzymes degrading the backbone of pectic polymers are the pectate and rhamnogalacturonan lyases, which use a β-elimination mechanism to cleave the HG and RG-I backbone and generate unsaturated, non-reducing ends (31). As smooth HG, as well as the hairy XGA and RGA-II, are often acetylated and methylated (Figure 2), the pectinases formulations also contain pectin methyl and acetyl esterases required for removing these groups (39, 40).

**Table 1 |** Composition of cell-wall fractions of common pectin-rich agricultural waste streams (*adapted from* (41)). N.D. is not determined

| **Component** | **Residue (%w/w)** | | |
|---|---|---|---|
| | **Apple pomace** | **Orange peel** | **Sugar beet pulp** |
| **Glucose** | 25.2 - 33.3 | 23.7 | 21.6 - 26.5 |
| **Galactose** | 3.0 - 7.0 | 8.2 | 4.2 – 4.9 |
| **Arabinose** | 5.1 - 14.3 | 14.2 | 16.3 – 20.1 |
| **Xylose** | 5.8 - 6.6 | < 5 | 1.4 - 1.6 |
| **Rhamnose** | 0.3 – 1.5 | < 2 | 1.0 - 2.25 |
| **Galacturonate** | 18.7 - 28.2 | 26.0 | 18.4 -23.0 |
| **Lignin** | N.D. | 3.0 | 1.0 - 2.0 |
| **Protein** | 9 - 11 | 5.5-6.7 | 3.6 – 8.0 |
| **Ash** | 1.5 - 2.0 | 3.9-4.1 | 4.4 – 12.0 |

Hydrolysis of pectins yields mixtures of monomers that predominantly consists of D-glucose, L-arabinose and D-galacturonate (Table 1) and only small amounts of the recalcitrant polymer lignin (26, 41, 42). Major pectin-rich agricultural residues include sugar beet pulp, citrus peel (predominantly from oranges) and apple pomace. These three major streams of pectin-containing waste streams predominantly originate from three continents. Citrus peel, a side product from the production of orange juice concentrates, is mainly produced in large volumes in the USA and in Brazil. While citrus peel waste can be used as gelling agent in the food industry, demand for this application is much lower than the availability (41, 43, 44), resulting in a large, essentially unused waste stream. Sugar-beet pulp and apple pomace are predominantly produced in Europe. These pectin-rich agricultural residues are currently mostly dried and sold as low-value cattle feed. The energy costs for drying leads to very low profit margins of this enterprise (32). In 2016, approximately 35 million tonnes of citrus peel, 28 million tonnes of sugar beet pulp and 2 million tonnes of apple pomace were produced in the world (45).

1

# Known microbial pathways for galacturonate conversion

## D-galacturonate transport

D-Galacturonate can be catabolized by many organisms across the tree of life. Two prokaryotic pathways and one fungal pathway are currently known (32, 43, 46). Mechanisms of D-galacturonate transport across microbial membranes remain underexplored, with only one prokaryotic and two fungal D-galacturonate transporters known (47–50). The bacterial hexuronate transporter (ExuT, TC 2.A.1.14.2) was identified in *Escherichia coli* (50, 51), *Erwinia chrysanthemi* (52) and *Erwinia carotovora* (49) and it is the only known bacterial transporter able to transport D-galacturonate. ExuT is a symporter-type transporter specific for hexuronates and it belongs to the anion:cation symporter (ACS) family of the major facilitator superfamily (MFS) of transporters (53). The gene encoding for this transporter was identified almost ten years after the first demonstration of its functionality (54) and (putative) *exuT* genes have since been identified in a large variety of prokaryotes (55).

Passive transport of D-galacturonate was shown to be possible in *Saccharomyces cerevisiae*, which cannot convert D-galacturonate, when the extracellular pH was below the $pK_a$ of D-galacturonate (3.5, (56)) and no transporter protein was identified (48, 57). Two years later, a bona fide D-galacturonate transporter (GAT-1) was identified in *Neurospora crassa*, and functionally expressed into a *Saccharomyces cerevisiae* strain (58). A second fungal D-galacturonate transporter, GatA of *Aspergillus niger*, was only identified in 2018. When expressed in *Saccharomyces cerevisiae,* GatA was approximately 50 times more active than Gat1 and showed no inhibition in the presence of glucose (47). These characteristics make GatA an interesting candidate for metabolic engineering strategies aimed at enabling D-galacturonate based product formation in *S. cerevisiae*.

# Catabolic pathways for galacturonate metabolism in prokaryotes

## The isomerase or adapted Entner-Doudoroff pathway

In many well-studied bacteria, such as *Escherichia coli*, *Bacillus subtilis, Thermotoga maritima* and *Lactobacillus brevis*, D-galacturonate is converted via the isomerase pathway (also known as adapted Entner-Doudoroff pathway; Figure 3A). Before the research described in Chapter 4 of this thesis, this pathway was the only one to support fermentative growth on D-galacturonate under fully anaerobic conditions (4, 32, 43).

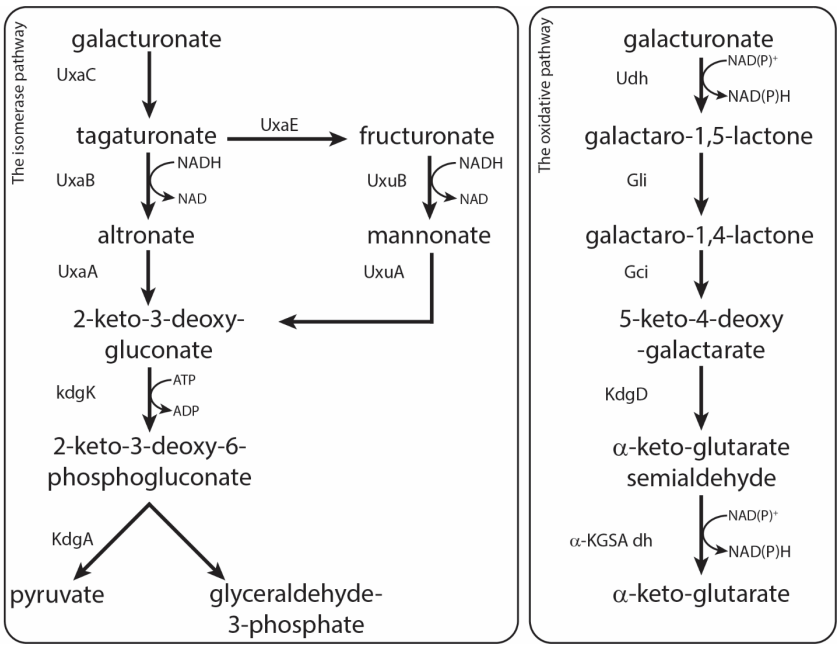



**Figure 3 |** Known bacterial pathways for D-galacturonate catabolism. A. The isomerase pathway, with UxaC, uronate isomerase; UxaB, tagaturonate reductase; UxaA, altronate hydratase; UxaE, tagaturonate 4-epimerase; UxuB, fructuronate reductase; UxuA, mannonate dehydratase; KdgK, 2-keto-3-deoxygluconate kinase; KdgA, 2-keto-3-deoxy-6-phosphogluconate aldolase. B. The oxidative pathway, with Udh, galacturonate dehydrogenase; Gli, galactarolactone isomerase; Gci, galactrolactone cycloisomerase; KdgD, 5-keto-4-deoxy-glucarate dehydratase; α-KGSA dh, 2-5- dioxypentanate dehydrogenase. *Adapted from* (43).

The isomerase pathway was first elucidated in *Escherichia coli* (59, 60). It starts with the isomerization of D-galacturonate to D-tagaturonate by uronate isomerase (UxaC, EC 5.3.1.12 (59)). D-tagaturonate is subsequently reduced to D-altronate with NAD(P)H by tagaturonate reductase (UxaB, EC 1.1.1.58 (61)). D-Altronate its then dehydrated to 2-keto-3-deoxygluconate with altronate dehydratase (UxaA, EC 4.2.1.7 (62)). An alternative version of this route was recently discovered and involves conversion D-tagaturonate to D-fructuronate with a tagaturonate 4-epimerase (UxaE, EC 5.1.2.7 (63)), followed by reduction of D-fructuronate to D-mannonate by an NAD(P)H-dependent fructuronate reductase (UxuB, EC 1.1.1.57 (61)). D-Mannonate is then dehydrated to 2-keto-3-deoxygluconate by a mannonate dehydratase (UxuA, EC 4.2.1.8 (62)). Although the conversions catalysed by UxaBC and UxuBC are highly similar, these enzymes are clearly different and predominantly associated with D-galacturonate and D-glucuronate fermentation, respectively (61, 62). From 2-keto-3-deoxygluconate, D-galacturonate metabolism via the isomerase pathway feeds into reactions of the canonical Entner-Doudoroff pathway for sugar metabolism (hence the alternative name 'adapted Entner-Doudoroff pathway'). In these reactions, 2-keto-3-deoxygluconate is first phosphorylated to 2-keto-3-deoxy-6-gluconate by 2-keto-3-deoxygluconate kinase (KdgK, EC 2.7.1.45

1

(64), the signature intermediate of the Entner-Doudoroff pathway (65), in a reaction that costs one ATP. Finally, 2-keto-3-deoxy-6-gluconate is split into two C3-molecules, pyruvate and glyceraldehyde-3-phosphate, by 2-keto-3-deoxy-6-gluconate aldolase (KdgA, EC 4.1.2.14 (66)). Glyceralde-3-phosphate is subsequently converted to pyruvate via the reactions of the lower part of the Embden-Meyerhof pathway of glycolysis, producing two ATP and one NADH. If no ATP costs are incurred for D-galacturonate uptake, the isomerase pathway converts one mole of D-galacturonate to two moles while yielding one mole of ATP. As will be discussed below, the net redox-cofactor neutrality of the isomerase pathway has a strong impact on the possible range of fermentation products that can be formed by anaerobic bacteria that employ it.

### The oxidative pathway

Not all prokaryotes able to catabolise D-galacturonate harbour the isomerase pathway. Many respiratory bacteria metabolize D-galacturonate via the oxidative pathway (Figure 3B), which was first observed in a *Pseudomonas* species but is best understood from research on the plant pathogen *Agrobacterium tumefaciens* (67). In the oxidative pathway, D-galacturonate is first oxidized with $NAD^+$ to D-galactaro-1,5-lactone by D-galacturonate dehydrogenase (Udh, EC 1.1.1.203 (68)). D-Galactaro-1,5-lactone is then converted to α-keto-glutarate semi-aldehyde by two consecutive isomerization reactions, catalysed by D-galactaro-1,5-lactone isomerase (Gli, EC 5.4.1.4 (69)) and D-galactaro-1,5-lactone cyclo-isomerase (Gci, EC 5.5.1.27 (70)). The last two steps in this pathway convert α-keto-glutarate semi-aldehyde to α-ketoglutarate and are catalysed by 5-dehydro-4-deoxy-glucarate dehydratase (KdgD, EC 4.2.1.41 (71)) and $NAD^+$-dependent 2,5-dioxypentanate dehydrogenase (α-KGSA dh, EC 1.2.1.26 (71)). The oxidative pathway results into the overall conversion of one mole D-galacturonate into one mole each of α-ketoglutarate, an intermediate of the citric acid cycle (72), and carbon dioxide and is coupled to the generation of two moles of NAD(P)H.

### Fungal D-galacturonate pathway

The fungal pathway of D-galacturonate metabolism was first identified in the mould *Trichoderma reesei* (73), and subsequently also in the well-studied fungus *Aspergillus niger* (46) and in *Botrytis cinerea* (74). The fungal pathway requires two moles of NAD(P)H to convert one mole of D-galacturonate into one mole each of pyruvate and glycerol (Figure 4). D-Galacturonate is first reduced to galactonate by galacturonate reductase (GaaA, EC 1.1.1.365 (73)). Galactonate is then dehydrated by galactonate dehydratase (GaaB, EC 4.2.1.146 (73)) to yield 2-keto-3-deoxygalactonate. This characteristic intermediate of the fungal pathway for D-galacturonate metabolism is then split into two C3-compounds, pyruvate and L-glyceraldehyde, by 2-keto-3-deoxygalactonate aldolase (GaaC, EC 4.1.2.54 (75)), followed by the reduction of L-glyceraldehyde to glycerol by glyceraldehyde reductase (GaaD, EC 1.1.1.372 (76)). Just like the bacterial isomerase pathway, the fungal pathway involves a split of a C6-

intermediate and yields two C3-molecules as products. However, in contrast to the isomerase pathway, the fungal pathway is not coupled to ATP formation by substrate-level phosphorylation (Figures 3 and 4).

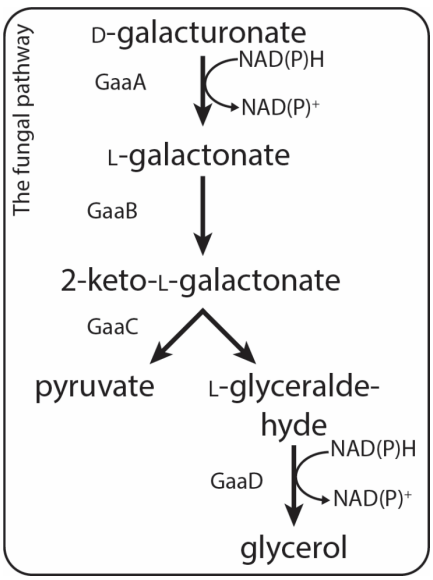



**Figure 4 |** The fungal pathway for galacturonate catabolism. With GaaA, galacturonate reductase; GaaB, galactonate dehydratase; GaaC, 2-keto-3-deoxy galactonate aldolase; GaaD, glyceraldehyde reductase. *Adapted from* (43).

## Challenges in anaerobic, fermentative product formation from galacturonate

Recent studies have shown how L-galactarate can be produced from D-galacturonate in aerobic cultures of engineered strains of the filamentous fungi *Trichoderma reesei* and *Aspergillus niger*. In these engineered strains, the gene encoding L-galactonate dehydratase (EC 4.2.1.146; *lgd1* and *gaab* in *T. reesi* and *A. niger*, respectively) were deleted, leading to accumulation of the intermediate L-galactonate (Figure 5, Table 2). L-galactonate has the potential to be used as a chelator, in skincare products and as a precursor for vitamin C (43, 77, 78). Recently, a study showed, which expressed the fungal D-galacturonate pathway in a *Saccharomyces cerevisiae* strain, that the consumption and fermentation of D-galacturonate improved when the uronic acid was co-consumed with D-xylose and L-arabinose (79). The oxidative pathway, engineered into an *Escherichia coli* strain, was used for the formation of 1,4-butanediol from D-galacturonate (Table 2;(80)).

Conversion of plant biomass hydrolysates to biotechnological bulk products such as ethanol, lactate and isobutanol should be performed under anaerobic conditions, since use of oxygen as external electron acceptor goes at the expense of product yield and requires expensive aeration of bioreactors (81–83). This requirement does not pose fundamental problems for conversion of sugars via glycolysis and/or pentose-phosphate pathway, which provides ATP via substrate phosphorylation and, moreover, generates reduced redox cofactors that can be used to reduce pyruvate, the product of glycolysis, into fermentation products.

Of the three pathways for D-galacturonate metabolism discussed above, the oxidative and fungal pathways do not allow for anaerobic, fermentative growth due to the absence of substrate-level phosphorylation. While the isomerase pathway (Figure 3) does enable many bacteria to ferment D-galacturonate under anaerobic conditions, its configuration severely constrains the possible range of fermentation products. The perfect redox-cofactor balance during conversion, via the isomerase pathway, of D-galacturonate into pyruvate implies that any reactions beyond pyruvate should also be redox-cofactor balanced. In practice, acetate, combined with formate, hydrogen and/or $CO_2$, is typically the predominant product of anaerobic D-galacturonate fermentation (84–86). This severely constrained product range precludes the use of known wild-type D-galacturonate fermenting micro-organisms in industrial processes for the anaerobic, fermentative production of added-value products. One of the solutions is to metabolically engineer micro-organisms to anaerobically produce industrial-interesting bulk products.

Several metabolic engineering studies have focused on enabling ethanol production by species of the *Enterobacteriaceae* family who already expressed the isomerase pathway for D-galacturonate metabolism (Table 2, (44, 87, 88)). However, these studies only resulted in a low ethanol yield on D-galacturonate metabolism (0.12 to 0.19 $g_p\ g_s^{-1}$), while acetate remained the predominant fermentation product (44, 87, 88).

Based on the currently described catabolic pathways for D-galacturonate metabolism, there do not appear to be feasible strategies for efficient, anaerobic microbial conversion of pectin-rich agricultural residues into large-volume fermentation products. Tackling this problem will therefore require the discovery of novel pathways or the design and construction of synthetic pathways that result in redox-cofactor balance and enable substrate-level phosphorylation.

**Table 2 |** Metabolically engineered organisms capable of catabolizing D-galacturonate, either as sole substrate or as co-substrate. The Table indicates the catabolic pathways employed in different studies along with the product, the product yield ($g_p\ g_s^{-1}$) and the oxygen status of the cultures. * Yeast extract.

| Organism | Substrate | Pathway | Main Product | Product yield ($g_p\ g_s^{-1}$) | Anaerobic | Reference |
|---|---|---|---|---|---|---|
| *Escherichia coli* KO11 | D-Galacturonate | Isomerase | Ethanol | 0.19 | Yes | (88) |
| *Erwinia chrysanthemi* EC16 | D-Galacturonate | Isomerase | Ethanol | 0.12 | Yes | (44) |
| *Erwinia carotovora* SR28 | D-Galacturonate | Isomerase | Ethanol | 0.16 | Yes | (44) |
| *Erwinia chrysanthemi E16* | D-Galacturonate | Isomerase | Ethanol | 0.16 | Yes | (87) |
| *Klebsiella oxytoca* P2 | D-Galacturonate | Isomerase | Ethanol | 0.12 | Yes | (87) |
| *Escherichia coli* KO11 | D-Galacturonate | Isomerase | Ethanol | 0.19 | Yes | (87) |
| *Escherichia coli* BDO7 | D-Galacturonate and glucose | Oxidative | 1,4 butanediol | 0.33 | No | (80) |
| *Trichoderma reesei* | D-Galacturonate | Fungal | L-galactonate | 0.43 | No | (77) |
| *Aspergillus niger* | D-Galacturonate | Fungal | L- galactonate | 0.59 | No | (77) |
| *Aspergillus niger* | D-Galacturonate and D-xylose | Fungal | L-ascorbate | 0.04 | No | (89) |
| *Trichoderma reesi* | D-Galacturonate, lactose and YE* | Fungal | L-galactarate | 0.99 | No | (78) |

1

### The D-galacturonate decarboxylase pathway – an interesting theoretical possibility

Direct decarboxylation of D-galacturonate could lead to the production of L-arabinose, a pentose sugar that can be fermented by many anaerobic microorganisms. This mechanism has been proposed to play a role in microbial D-galacturonate metabolism since the late 19th century (90, 91). Chemical decarboxylation of D-galacturonate under acidic conditions showed the production of furfural or reductic acid (92–95). Zweifel and Deuel, (1956) showed that, under mildly acidic conditions, metal-catalysed decarboxylation of D-galacturonate to L-arabinose occurred. This observation reinvigorated speculation that enzyme-catalysed D-galacturonate decarboxylation might occur in nature and play a role in pectin degradation (97–100). Activity of an enzyme that catalyses a closely related reaction, UDP-galacturonate decarboxylase, has been reported to occur in *Ampullariella digitata* (EC 4.1.1.67) (99), but no sequence information or other follow-up research on this observation has been published.

Discovery of a D-galacturonate decarboxylase would be of immediate applied interest, since L-arabinose can be readily fermented to ethanol, an important biofuel, by engineered strains of the yeast *Saccharomyces cerevisiae* (20, 24). Theoretically, combined expression of a D-galacturonate transporter and a D-galacturonate decarboxylase in such engineered yeast strains should enable efficient conversion of hydrolysates of pectin-containing agricultural residues to ethanol. However, microorganisms that metabolize D-galacturonate via a decarboxylase pathway have hitherto not been found in nature (101).

## Exploring microbial biodiversity by enrichment cultivation

In 1934, the Dutch biologist Baas Becking, referring to microbial ecology, formulated his now famous statement “Alles is overal, maar het milieu selecteert” (“Everything is everywhere, but the environment selects” (102)). This principle lies at the basis of selective enrichment cultivation as a tool for the exploration of the bewildering diversity of microbial metabolism.

In natural environments, competition between micro-organisms for resources is strongly affected by chemical and physical conditions. For example, specific combinations of the presence and concentrations of nutrients, products and inhibitors with specific ranges of pH and temperature can strongly affect competitiveness of microbes that contain specific metabolic pathways over others (103–105)). In addition, dynamics of these parameters, e.g. as a result of circadian cycles, can impose additional selective pressures (106, 107).

In the laboratory, cultivation conditions designed to select for specific types of metabolism or other traits of interests can easily be designed, implemented and maintained over a large number of generations of selective growth. In practice, creating a selective environment can be accomplished by altering cultivation parameters e.g.

substrate, pH, temperature, anaerobicity, dilution rate or cultivation regime (108–110). After inoculation with an environmental sample that harbours a large variety of microbial species, the organisms or metabolic types that are most competitive during prolonged cultivation will become enriched in the resulting microbial communities (105, 111–113).

Studies on microbial wastewater treatment have repeatedly demonstrated how organisms with highly interesting, novel metabolic capabilities can be selected from diverse environments. A spectacular example from research in the TU Delft's Department of Biotechnology resulted in the discovery of the anaerobic ammonium oxidizing (anammox) bacteria (114) and subsequent enrichment of the responsible organism (115). More recently, a *Geobacter* species, enriched by continuous cultivation, was shown to apply dissimilatory nitrate reduction to ammonium (DNRA) as energy-conserving pathway (116). Independently, two research groups used enrichment cultivation to demonstrate the complete oxidation of ammonium to nitrate by 'comammox' bacteria belonging to the genus *Nitrospira* (117, 118). This discovery finally proved that a process already predicted by Winogradsky (1890), occurred in nature.

The use of dynamic cultivation regimes offers additional possibilities to enrich microorganisms with interesting metabolic capabilities. The power of this approach is illustrated by the successful use of a 'feast-famine' feeding regime to enrich bacteria that convert acetate into polyhydroxyalkanoates (PHA), which are microbial polymers that can be used as bioplastics (106). Enrichment cultivation can be used to isolate microorganisms that are tolerant to industrially relevant process conditions. Using enrichment under acidic conditions, the fungus *Monascus ruber* was identified as a novel potential microbial platform for lactic acid production (120).

The population composition and dynamics of enrichment cultures can be assessed based on the abundance of 16S-rRNA gene sequences, either via denaturing gel electrophoreses (DGGE), via 16S-rRNA amplicon sequencing or fluorescence *in situ* hybridization (121–124). Based on phylogeny and metabolism of the closest cultured relative, these analyses allow for a first prediction of the physiological role of a species in the enrichment culture (104). The recent advancement of omics-strategies and whole genome sequencing techniques have enabled much deeper insights in the genomic potential of microbial populations (125–127). Although often used for genome-centric shot-gun microbial ecology approaches, whole-genome sequencing data can additionally be harnessed to couple microbial physiology of enrichment cultures to the genetic potential of individual species in mixed populations. This is illustrated by research on 'comammox' bacteria, which demonstrated presence of a novel conversion based on a combination of microbial physiological and metagenomic data (117, 118). In some cases, metagenomic approaches can reduce the impact or even avoid the need for traditional culture-based isolation strategies, as they can directly provide genomic information of microorganism without isolation them from their natural environment (128, 129).

The examples mentioned above illustrate that enrichment-based strategies combined with metagenomic analysis are an interesting and successful strategy to broaden our knowledge on the metabolic diversity of anaerobic micro-organisms. This strategy was therefore employed in the research described in this thesis to explore hitherto undiscovered pathways for anaerobic fermentation of D-galacturonate.

## Scope and outline of this thesis

Microbial conversion of pectin-rich agricultural waste streams for the efficient, anaerobic production of major fermentation products is currently precluded by a lack of suitable catabolic pathways for conversion of D-galacturonate. The goal of the research described in this thesis was to explore metabolic and microbial diversity of anaerobic D-galacturonate metabolism. The research approach encompassed a combination of enrichment cultivation, followed by a thorough genomic and physiological characterization of the enriched microbial populations.

After describing the relevance of the use of pectin-rich agricultural waste streams for the microbial production of industrial relevant products, **Chapter 1** discusses the currently know pathways for microbial catabolism of the D-galacturonate, the major monomer of pectin. Limitations related to carbon and redox conversion in known pathways for D-galacturonate catabolism are linked to the limited range of fermentation products that, even in engineered micro-organisms, can be made from this substrate. The Chapter concludes with a brief introduction to enrichment cultivation as a tool for exploring the diversity of microbial metabolism. This approach plays a central role in the following Chapters.

**Chapter 2** describes the use of anaerobic, mixed-population chemostat cultures grown under D-galacturonate limitation for selection of D-galacturonate-fermenting microorganisms. The resulting population was subject to 16S-rRNA gene amplicon sequencing as well as whole-genome sequencing to identify dominant microorganisms and explore their metabolisms. The observed product profile did not match the stoichiometry expected from the classical isomerase pathway for D-galacturonate metabolism but suggested that, instead, this pathway was combined with heterotrophic acetogenesis. A first test of this hypothesis was based on analysis of a metagenomic assembled genome (MAG) of the dominant organism from metagenome sequence data. This analysis raised interesting questions about the pathway responsible for acetogenesis.

**Chapter 3** describes a detailed analysis of the proposed acetogenic catabolism of D-galacturonate of the dominant organism identified in Chapter 2. In particular, $^{13}C$-labelling experiments, in which $NaH^{13}CO_3$ was fed to D-galacturonate-limited chemostat

enrichment cultures to rigorously test if acetogenesis via a Wood-Ljungdahl-type pathway was indeed responsible for the unexpected product stoichiometries observed in Chapter 2. In addition, long-read DNA sequencing and meta-transcriptome analysis were employed to test a tentative conclusion that the dominant organism harboured an atypical variant of the canonical Wood-Ljungdahl pathway for acetogenesis, due to the absence of the CODH/ACS complex.

**Chapter 4** describes experiments that were designed to explore the effect of enrichment in anaerobic, galacturonate-grown batch cultures under mildly acidic conditions. The rationale of this experiment was that acetate, the predominant product of the conventional isomerase pathway for D-galacturonate, is toxic to microorganisms at low pH and that selection under acidic conditions might therefore favour formation of different fermentation products. The product profile of pure, D-galacturonate-grown anaerobic cultures of a dominant organism isolated from the enrichment cultures was studied in detail in bioreactor batch and chemostat experiments and was shown to differ substantially from that of previously studied D-galacturonate-fermenting bacteria. A combination of genome analysis, proteomics and *in vitro* enzyme-activity analyses and product identification were used to elucidate the responsible, novel pathway for D-galacturonate catabolism.

1

# References


1. O'Rourke D, Connolly S. 2003. Just Oil ? The Distribution of Environmental and Social Impacts of Oil Production and Consumption. Annu Rev Environ Resour 28:587–617.
2. Edenhofer O, Pichs-Madruga R, Sokona Y, Kadner S, Minx JC, Brunner S, Agrawala S, Baiocchi G, Bashmakov IA, Blanco G, Broome J, Bruckner T, Bustamante M, Clarke L, Grand MC, Creutzig F, Cruz-Núñez X, Dhakal S, Dubash NK, Eickemeier P, Farahani E, Fischedick M, Fleurbaey M, Gerlagh R, Gómez-Echeverri L, Gupta S, Harnisch J, Jiang K, Jotzo F, Kartha S, Klasen S, Kolstad C, Krey V, Kunreuther H, Lucon O, Masera O, Mulugetta Y, Norgaard RB, A. Patt, Ravindranath NH, Riahi K, Roy J, Sagar A, Schaeffer R, Schlömer S, Seto KC, Seyboth K, Sims R, Smith P, Somanathan E, Stavins R, Stechow C von, Sterner T, Sugiyama T, Suh S, Ürge-Vorsatz D, Urama K, Venables A, Victor DG, Weber E, Zhou D, Zou J, Zwickel T. 2014. Technical Summary, p. . *In* Edenhofer, O, Pichs-Madruga, R, Sokona, Y, Farahani, E, Kadner, S, Seyboth, K, Adler, A, Baum, I, Brunner, S, Eickemeier, P, Kriemann, B, Savolainen, J, Schlömer, S, Stechow, C von, Zwickel, T, Minx, JC (eds.), Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, USA.
3. Hermann BG, Blok K, Patel MK. 2007. Producing bio-based bulk chemicals using industrial biotechnology saves energy and combats climate change. Environ Sci Technol 41:7915–7921.
4. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT. 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 90:391–418.
5. Bigg GR, Jickells TD, Liss PS, Osborn TJ. 2003. The role of the oceans in climate. Int J Climatol 23:1127–1159.
6. Keasling JD. 2007. Synthetic biology for synthetic chemistry. ACS Chem Biol 3:64–76.
7. Ekas H, Deaner M, Alper HS. 2019. Recent advancements in fungal-derived fuel and chemical production and commercialization. Curr Opin Biotechnol 57:1–9.
8. Becker J, Wittmann C. 2015. Advanced biotechnology: Metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products. Angew Chemie - Int Ed 54:3328–3350.
9. Benoit I, Culleton H, Zhou M, DiFalco M, Aguilar-Osorio G, Battaglia E, Bouzid O, Brouwer CPJM, El-Bushari HBO, Coutinho PM, Gruben BS, Hildén KS, Houbraken J, Barboza LAJ, Levasseur A, Majoor E, Mäkelä MR, Narang H-M, Trejo-Aguilar B, van den Brink J, vanKuyk P a, Wiebenga A, McKie V, McCleary B, Tsang A, Henrissat B, de Vries RP. 2015. Closely related fungi employ diverse enzymatic strategies to degrade plant biomass. Biotechnol Biofuels 8:107.
10. Kamp A, La Revière J, Verhoeven W. 1959. Albert Jan Kluyver his life and his work; biographical memoranda, selected papers, bibliography and addenda. Noord-Holland Pub. Co, Amsterdam.
11. Tao L, Aden A. 2009. The economics of current and future biofuels. Vitr Cell Dev

Biol - Plant 45:199–217.
12. Tilman D, Socolow R, Foley JA, Hill J, Larson E, Lynd L, Pacala S, Reilly J, Searchinger T, Somerville C, Williams R. 2009. Beneficial biofuels - The food, energy, and environment trilemma. Science 325:270–271.
13. Klein-Marcuschamer D, Blanch HW. 2009. Renewable fuels from Biomass: Technical hurdles and economic assessment of biological routes. Am Inst Chem Eng J 7:2689–2701.
14. Liao JC, Mi L, Pontrelli S, Luo S. 2016. Fuelling the future : microbial engineering for the production of sustainable biofuels. Nat Rev Microbiol 14:288–304.
15. Fargione J, Hill J, Tillman D, Polasky S, Hawthorne P. 2008. Leand clearing and biofuel carbon dept. Science (80- ) 319:1235–1238.
16. Bothast RJ, Schlicher MA. 2005. Biotechnological processes for conversion of corn into ethanol. Appl Microbiol Biotechnol 67:19–25.
17. Kumar P, Barrett DM, Delwiche MJ, Stroeve P. 2009. Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production. Ind Eng Chem Res 48:3713–3729.
18. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS. 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66:506–577.
19. Agler MT, Wrenn BA, Zinder SH, Angenent LT. 2011. Waste to bioproduct conversion with undefined mixed cultures: the carboxylate platform. Trends Biotechnol 29:70–78.
20. Jansen MLA, Bracher JM, Papapetridis I, Verhoeven MD, de Bruijn H, de Waal PP, van Maris AJA, Klaassen P, Pronk JT. 2017. *Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. FEMS Yeast Res 17:1–20.
21. Bothast RJ, Saha BC, Vance Flosenzier A, Ingram LO. 1994. Fermentation of L-arabinose, D-xylose and D-glucose by ethanolegenic recombinant *Klebsiella oxytoca* strain P2. Biotechnol Lett 16:401–406.
22. Singh N, Mathur AS, Tuli DK, Gupta RP, Barrow CJ, Puri M. 2017. Biotechnology for Biofuels Cellulosic ethanol production via consolidated bioprocessing by a novel thermophilic anaerobic bacterium isolated from a Himalayan hot spring. Biotechnol Biofuels 1–18.
23. Verhoeven MD, de Valk SC, Daran JMG, van Maris AJA, Pronk JT. 2018. Fermentation of glucose-xylose-arabinose mixtures by a synthetic consortium of single-sugar-fermenting *Saccharomyces cerevisiae* strains. FEMS Yeast Res 18:1–12.
24. Wisselink HW, Toirkens MJ, Berriel MDRF, Winkler AA, Van Dijken JP, Pronk JT, Van Maris AJA. 2007. Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. Appl Environ Microbiol 73:4881–4891.
25. Holtzapple MT, Granda CB. 2009. Carboxylate platform: The MixAlco process part 1: Comparison of three biomass conversion platforms. Appl Biochem Biotechnol 156:95–106.
26. Mohnen D. 2008. Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277.
27. Kuivanen J. 2015. Metabolic engineering of the fungal D-galacturonate pathway. VTT Technical Research Centre of Finland.
28. Link KP, Nedden R. 1931. Preparation of D-galacturonic acid. J Biol Chem 94:307–

314.
29. Ehrlich F. 1932. Über die Pektolaase, ein neuaugefundenes Pektinferment. II. Biochem Zeitung 251:204–222.
30. Ehrlich F. 1932. Über die Pektolase, ein neuaugefundenes Pektinferment. I. Biochem Zeitung 250:525–534.
31. de Vries RP, Visser J. 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev 65:497–522.
32. Richard P, Hilditch S. 2009. D-Galacturonic acid catabolism in microorganisms and its biotechnological relevance. Appl Microbiol Biotechnol 82:597–604.
33. Willats WGT, Orfila C, Limberg G, Buchholt HC, Van Alebeek GJWM, Voragen AGJ, Marcus SE, Christensen TMIE, Mikkelsen JD, Murray BS, Knox JP. 2001. Modulation of the degree and pattern of methyl-esterification of pectic homogalacturonan in plant cell walls: Implications for pectin methyl esterase action, matrix properties, and cell adhesion. J Biol Chem 276:19404–19413.
34. Yapo BM, Lerouge P, Thibault JF, Ralet MC. 2007. Pectins from citrus peel cell walls contain homogalacturonans homogenous with respect to molar mass, rhamnogalacturonan I and rhamnogalacturonan II. Carbohydr Polym 69:426–435.
35. Rebello S, Anju M, Aneesh EM, Sindhu R, Binod P, Pandey A. 2017. Recent advancements in the production and application of microbial pectinases: an overview. Rev Environ Sci Biotechnol 16:381–394.
36. Kashyap DR, Vohra PK, Chopra S, Tewari R. 2001. Applications of pectinases in the commercial sector: A review. Bioresour Technol 77:215–227.
37. Culleton H, Mckie V, De Vries RP. 2013. Physiological and molecular aspects of degradation of plant polysaccharides by fungi: What have we learned from Aspergillus? Biotechnol J 8:884–894.
38. Guillon F, Thibault JF. 1989. Enzymatic hydrolysis of the "hairy" fragments of sugar-beet pectins. Carbohydr Res 190:97–108.
39. Searle-van Leeuwen MJF, Vincken JP, Beldman G. 1996. Acetyl esterases of *Aspergillus niger:* purification and mode of action on pectinsProc. Int. Symp. on Progress in Biotechnology 14: Pectins and pectinases. Elsevier, Amsterdam.
40. Khanh NQ, Ruttkowski E, Leidinger K, Albrecht H, Gottschalk M. 1991. Characterization and expression of a genomic pectin methyl esterase-encoding gene in *Aspergillus niger*. Gene 106:71–77.
41. Grohmann K, Bothast RJ. 1994. Pectin-rich residues generated by processing of citrus fruits, apples, and sugar beets - enzymatic hydrolysis and biological conversion to value-added products. ACS Symp Ser Enzym Convers Biomass Fuels Prod 566:372–390.
42. Micard V, Renard CMGC, Thibault JF. 1996. Enzymatic saccharification of sugar-beet pulp. Enzyme Microb Technol 19:162–170.
43. Kuivanen J, Biz A, Richard P. 2019. Microbial hexuronate catabolism in biotechnology. AMB Express 9:1–11.
44. Grohmann K, Manthey JA, Cameron RG, Buslig BS. 1998. Fermentation of galacturonic acid and pectin-rich materials to ethanol by genetically modified strains of *Erwinia*. Biotechnol Lett 20:195–200.
45. Food and Agriculture Organization of the United Nations. 2016. Crops, Production quantities of sugar beet, citrus fruits and apples in the World. FAOSTAT. http://www.fao.org/faostat/en/#data/QC

46. Martens-Uzunova ES, Schaap PJ. 2008. An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45:1449–1457.
47. Protzko RJ, Latimer LN, Martinho Z, de Reus E, Seibert T, Benz JP, Dueber JE. 2018. Engineering *Saccharomyces cerevisiae* for co-utilization of D-galacturonic acid and D-glucose from citrus peel waste. Nat Commun 9:5059 (2018).
48. Souffriau B, den Abt T, Thevelein JM. 2012. Evidence for rapid uptake of D-galacturonic acid in the yeast *Saccharomyces cerevisiae* by a channel-type transport system. FEBS Lett 586:2494–2499.
49. Hugouvieux-Cotte-Pattat N, Quesneau Y, Robert-baudouy J. 1983. Aldohexuronate transport system in *Erwinia carotovora*. J Bacteriol 154:663–668.
50. Nemoz G, Robert Baudouy J, Stoeber F. 1976. Physiological and genetic regulation of the aldohexuronate transport system in *Escherichia coli*. J Bacteriol 127:706–718.
51. Abendano JJ, Kepes A. 1973. Sensitization of D-glucuronic acid transport system of *Escherichia coli* to protein group reagents in presence of substrate or absence of energy source. Biochem Biophys Res Commun 54:1342–1346.
52. San Francisco MJD, Keenan RW. 1993. Uptake of galacturonic acid in *Erwinia chrysanthemi* EC16. J Bacteriol 175:4263–4265.
53. Pao SS, Paulsen IT, Saier Jr HM. 1998. Major Facilitator Superfamily. Microbiol Mol Biol Rev 62:1–34.
54. Mata-Gilsinger M, Ritzenthaler P. 1983. Physical mapping of the *exuT* and *uxaC* operators by use of *exu* plasmids and generation of deletion mutants *in vitro*. J Bacteriol 155:973–982.
55. Huisjes EH, Luttik MAH, Almering MJH, Bisschops MMM, Dang DHN, Kleerebezem M, Siezen R, Maris AJA Van, Pronk JT. 2012. Toward pectin fermentation by *Saccharomyces cerevisiae* : Expression of the first two steps of a bacterial pathway for D-galacturonate metabolism. J Biotechnol 162:303–310.
56. Kohn R, Kovác P. 1978. Dissociation constants of D-galacturonic and D-glucuronic acid and their O-methyl derivatives. Chem Zvesti 32:478–485.
57. Huisjes EH. 2013. Towards fermentation of galacturonic acid-containing feedstocks with Saccharomyces cerevisiae.
58. Benz JP, Protzko RJ, Andrich JM, Bauer S, Dueber JE, Somerville CR. 2014. Identification and characterization of a galacturonic acid transporter from *Neurospora crassa* and its application for *Saccharomyces cerevisiae* fermentation processes. Biotechnol Biofuels 7:20.
59. Ashwell G, Wahba AJ, Hickman J. 1960. Uronic acid metabolism in bacteria I Purification and properties of uronic acid isomerase in *Escherichia coli*. J Biol Chem 235:1559–1565.
60. Kilgore WW, Starr M. P. 1959. Catabolism of galacturonic and glucuronic acids by *Erwinia carotovora*. J Biol Chem 234:2227–2235.
61. Hickman J, Ashwell G. 1960. Uronic acid metabolism in bacteria II Purification and properties of D-altronic acid and D-mannonic acid dehydrogenases in *Escherichia coli*. J Biol Chem 235:1566–1570.
62. Smiley JD, Ashwell G. 1960. Uronic acid metabolism in bacteria III purification and properties of D-altronic acid and D-mannonic acid dehydrases in *Escherichia coli*. J Biol Chem 235:1571–1575.
63. Rodionova IA, Scott DA, Grishin N V., Osterman AL, Rodionov DA. 2012.

Tagaturonate-fructuronate epimerase UxaE, a novel enzyme in the hexuronate catabolic network in *Thermotoga maritima*. Environ Microbiol 14:2920–2934.
64. Cynkin MA, Ashwell G. 1960. Uronic acid metabolism in Bacteria IV Purification and properties of 2-keto-3-deoxy-D-gluconokinase in *Eschericia coli*. J Biol Chem 235:1576–1579.
65. Peekhaus N, Conway T. 1998. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J Bacteriol 180:3495–3502.
66. Kovachevich R, Wood WA. 1955. Carbohydrate metabolism by *Pseudomonas fluorescens* IV purification and properties of 2-keto-3-deoxy-6-phospho-gluconate aldolase. J Biol Chem 213:757–767.
67. Chang YF, Feingold DS. 1970. D-Glucaric acid and galactaric acid catabolism by *Agrobacterium tumefaciens*. J Bacteriol 102:85–96.
68. Boer H, Maaheimo H, Koivula A, Penttilä M, Richard P. 2010. Identification in *Agrobacterium tumefaciens* of the D-galacturonic acid dehydrogenase gene. Appl Microbiol Biotechnol 86:901–909.
69. Bouvier JT, Groninger-Poe FP, Vetting M, Almo SC, Gerlt JA. 2014. Galactaro δ-lactone isomerase: Lactone isomerization by a member of the amidohydrolase superfamily. Biochemistry 53:614–616.
70. Andberg M, Maaheimo H, Boer H, Penttilä M, Koivula A, Richard P. 2012. Characterization of a novel *Agrobacterium tumefaciens* galactarolactone cycloisomerase enzyme for direct conversion of D-galactarolactone to 3-deoxy-2-keto-L-threo-hexarate. J Biol Chem 287:17662–17671.
71. Aghaie A, Lechaplais C, Sirven P, Tricot S, Besnard-Gonnet M, Muselet D, De Berardinis V, Kreimeyer A, Gyapay G, Salanoubat M, Perret A. 2008. New insights into the alternative D-glucarate degradation pathway. J Biol Chem 283:15638–15646.
72. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. Molecular Biology of the Cell, 4th ed. Garland Science, New York.
73. Kuorelahti S, Kalkkinen N, Penttilä M, Londesborough J, Richard P. 2005. Identification in the mold *Hypocrea jecorina* of the first fungal D-galacturonic acid reductase. Biochemistry 44:11234–11240.
74. Zhang L, Thiewes H, van Kan J a L. 2011. The D-galacturonic acid catabolic pathway in *Botrytis cinerea*. Fungal Genet Biol 48:990–997.
75. Hilditch S, Berghäll S, Kalkkinen N, Penttilä M, Richard P. 2007. The missing link in the fungal D-galacturonate pathway: Identification of the L-threo-3-deoxy-hexulosonate aldolase. J Biol Chem 282:26195–26201.
76. Liepins J, Kuorelahti S, Penttilä M, Richard P. 2006. Enzymes for the NADPH-dependent reduction of dihydroxyacetone and D-glyceraldehyde and L-glyceraldehyde in the mould *Hypocrea jecorina*. FEBS J 273:4229–4235.
77. Kuivanen J, Mojzita D, Wang Y, Hilditch S, Penttilä M, Richard P, Wiebe MG. 2012. Engineering filamentous fungi for conversion of D-galacturonic acid to L-galactonic acid. Appl Environ Microbiol 78:8676–8683.
78. Barth D, Wiebe MG. 2017. Enhancing fungal production of galactaric acid. Appl Microbiol Biotechnol 101:4033–4040.
79. Jeong D, Ye S, Park H, Kim SR. 2020. Simultaneous fermentation of galacturonic acid and five-carbon sugars by engineered *Saccharomyces cerevisiae*. Bioresour Technol 295:122259.
80. Tai YS, Xiong M, Jambunathan P, Wang J, Wang J, Stapleton C, Zhang K. 2016.

Engineering nonphosphorylative metabolism to generate lignocellulose-derived products. Nat Chem Biol 12:247–253.
81. Weusthuis R a., Lamot I, van der Oost J, Sanders JPM. 2011. Microbial production of bulk chemicals: Development of anaerobic processes. Trends Biotechnol 29:153–158.
82. Cueto-Rojas HF, van Maris AJA, Wahl SA, Heijnen JJ. 2015. Thermodynamics-based design of microbial cell factories for anaerobic product formation. Trends Biotechnol 33:534–546.
83. Porro D, Branduardi P, Sauer M, Mattanovich D. 2014. Old obstacles and new horizons for microbial chemical production. Curr Opin Biotechnol 30:101–106.
84. Kraght AJ, Starr MP. 1952. Fermentation of galacturonic acid and glucose by a strain of Erwinia carotovora. J Bacteriol 64:259–264.
85. Dehority BA. 1969. Pectin-fermenting bacteria isolated from the bovine rumen. J Bacteriol 99:189–196.
86. Ng H, Vaughn RH. 1963. *Clostridium rubrum* sp. n. and other pectinolytic *Clostridia* from soil. J Bacteriol 85:1104–1113.
87. Doran JB, Cripe J, Sutton M, Foster B. 2000. Fermentations of pectin-rich biomass with recombinant bacteria to produce fuel ethanol. Appl Biochem Biotechnol 84–86:141–153.
88. Grohmann K, Baldwin EA, Buslig BS. 1994. Fermentation of galacturonic acid and other sugars in orange peel hydrolysates by the ethanologenic strain of *Escherichia coli*. Biotechnol Lett 16:281–286.
89. Kuivanen J, Penttilä M, Richard P. 2015. Metabolic engineering of the fungal D-galacturonate pathway for L-ascorbic acid production. Microb Cell Fact 14:1–9.
90. Ruff O. 1898. Über die Verwandlung der D-gluconsäure in D-arabinose. Berichte der Dtsch Chem Gesellschaft 31:1573–1577.
91. McKinnis RB. 1928. An investigation of the hypothetical combined pentose and the so-called free pentose with inferences on the composition of pectin. J Am Chem Soc 50:1911–1915.
92. Conrad CM. 1931. The decarboxylation of D-galacturonic acid with special reference to the hypothetical formation of L-arabinose. J Am Chem Soc 53:2282–2287.
93. Feather MS, Harris JF. 1966. Relationships between Some Uronic Acids and Their Decarboxylation Products. J Org Chem 31:4018–4021.
94. Popoff T, Theander O. 1972. Formation of aromatic compounds from carbohydrates. Part 1. Reaction of D-glucuronic acid, D-galacturonic acid, D-xylose, and L-arabinose in slightly acidic, aqueous solution. Carbohydr Res 22:135–149.
95. Stutz E, Deuel H. 1956. Über die Bildung von 5-formylbrenzschleimsäure aus D-Galakturonsäure. Helv Chim Acta 39:2126–2130.
96. Zweifel G, Deuel H. 1956. Über die Decarboxylierung von Galakturonsäure mit Schwermetallionen. Helv Chim Acta 77:662–667.
97. Hirst EL. 1942. Recent progress in the chemistry of the pectic materials and plant gums. J Chem Soc 70–78.
98. Gulland JM. 1943. Some aspects of the chemistry of nucleotides. J Chem Soc 208–217.
99. Fan DF, Feingold DS. 1972. Biosynthesis of uridine diphosphate D-xylose V. UDP-D-glucuronate and UDP-D-galacturonate carboxy-lyase of *Ampullariella digitata*. Arch Biochem Biophys 148:576–580.

100. Marounek M, Dušková D. 1999. Metabolism of pectin in rumen bacteria *Butyrivibrio fibrisolvens* and *Prevotella ruminicola*. Lett Appl Microbiol 29:429–433.
101. Cohen SS. 1948. Adaptive enzyme formation in the study of uronic acid utilization by the K-12 strain of *Escherichia coli*. J Biol Chem 607–620.
102. Baas-Becking LGM. 1934. Geologie of inleiding tot de milieukunde.
103. Stouten GR, Hogendoorn C, Douwenga S, Kilias ES, Muyzer G, Kleerebezem R. 2019. Temperature as competitive strategy determining factor in pulse-fed aerobic bioreactors. ISME J.
104. Temudo MF, Muyzer G, Kleerebezem R, van Loosdrecht MCM. 2008. Diversity of microbial communities in open mixed culture fermentations: Impact of the pH and carbon source. Appl Microbiol Biotechnol 80:1121–1130.
105. Schlegel HG, Jannasch HW. 1967. Enrichment cultures. Annu Rev Microbiol 21:49–70.
106. Johnson K, Jiang Y, Kleerebezem R, Muyzer G, van Loosdrecht MCM. 2009. Enrichment of a mixed bacterial culture with a high polyhydroxyalkanoate storage capacity. Biomacromolecules 10:670–676.
107. Gilbert JA, Steele JA, Caporaso JG, Steinbruck L, Reeder J, Temperton B, Huse S, McHardy a C, Knight R, Joint I, Somerfield P, Fuhrman J a, Field D. 2012. Defining seasonal marine microbial community dynamics. ISME J 6:298–308.
108. Kleerebezem R, van Loosdrecht MCM. 2007. Mixed culture biotechnology for bioenergy production. Curr Opin Biotechnol 18:207–212.
109. Zoetemeyer RJ, Arnoldy P, Cohen A, Boelhouwer C. 1982. Influence of temperature on the anaerobic acidification of glucose in a mixed culture forming part of a two-stage digestion process. Water Res 16:313–321.
110. Temudo MF, Kleerebezem R, van Loosdrecht M. 2007. Influence of the pH on (open) mixed culture fermentation of glucose : A chemostat study. Biotechnol Bioeng 98:69–79.
111. Bijerinck MW. 1901. Enrichment culture studies with urea bacteria. Centalblatt f Bacteriol Part II:33–61.
112. van Niel CB. 1949. The “Delft School” and the rise of general microbiology. Bacteriol Rev 13:161–174.
113. Großkopf T, Soyer OS. 2016. Microbial diversity arising from thermodynamic constraints. ISME J 10:2725–2733.
114. Van De Graaf AA, De Bruijn P, Robertson LA, Jetten MSM, Kuenen JG. 1996. Autotrophic growth of anaerobic ammonium-oxidizing micro-organisms in a fluidized bed reactor. Microbiology 142:2187–2196.
115. Strous M, Heijnen JJ, Kuenen JG, Jetten MSM. 1998. The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. Appl Microbiol Biotechnol 50:589–596.
116. van den Berg EM, van Dongen U, Abbas B, van Loosdrecht MCM. 2015. Enrichment of DNRA bacteria in a continuous culture. ISME J 9:2153–2161.
117. Daims H, Elena V, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard RH, Bergen M Von, Rattei T. 2015. Complete nitrification by Nitrospira bacteria. Nature 528:504–509.
118. van Kessel MAHJ, Speth DR, Albertsen M, Nielsen PH, Op den Camp HJM, Kartal B, Jetten MSM, Lücker S. 2015. Complete nitrification by a single microorganism. Nature 528:555–559.

119. Winogradsky S. 1890. Recherches sur les organismes de la nitrification. Ann l'Institut Pasteur 4:213–231.
120. Weusthuis RA, Mars AE, Springer J, Wolbert EJ, van der Wal H, de Vrije TG, Levisson M, Leprince A, Houweling-Tan GB, Moers AP, Hendriks SN, Mendes O, Griekspoor Y, Werten MW, Schaap PJ, van der Oost J, Eggink G. 2017. *Monascus ruber* as cell factory for lactic acid production at low pH. Metab Eng 42:66–73.
121. Muyzer G. 1999. DGGE / TGGE a method for identifying genes from natural ecosystems. Curr Opin Microbiol 2:317–323.
122. Ward DM, Weller R, Bateson MM. 1990. 16S rRNA sequences reveal numerous uncultured inhabitants in a well-studied natural community. Nature 345:63–65.
123. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 56:1919–1925.
124. Muyzer G, Teske A, Wirsen CO, Jannasch HW. 1995. Phylogenetic relationships of Thiomicrospira species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Arch Microbiol 164:165–172.
125. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. Nat Biotechnol 31:533–541.
126. Figueroa IA, Barnum TP, Somasekhar PY, Carlström CI, Engelbrektson AL, Coates JD. 2018. Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural $CO_2$ fixation pathway. Proc Natl Acad Sci 115:E92–E101.
127. Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev V V., Rubin EM, Rokhsar DS, Banfield JF. 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428:37–43.
128. Albertsen M, McIlroy SJ, Stokholm-Bjerregaard M, Karst SM, Nielsen PH. 2016. "*Candidatus* Propionivibrio aalborgensis": A novel glycogen accumulating organism abundant in full-scale enhanced biological phosphorus removal plants. Front Microbiol 7:1–17.
129. Hao L, McIlroy SJ, Kirkegaard RH, Karst SM, Fernando WEY, Aslan H, Meyer RL, Albertsen M, Nielsen PH, Dueholm MS. 2018. Novel prosthecate bacteria from the candidate phylum *Acetothermia*. ISME J 12:2225–2237.

# Chapter 2 | Galacturonate metabolism in anaerobic chemostat enrichment cultures: combined fermentation and acetogenesis by the dominant sp. nov. "*Candidatus* Galacturonibacter soehngenii"

Laura C. Valk, Jeroen Frank, Pilar de la Torre-Cortes, Max van 't Hof, Antonius J.A van Maris, Jack T. Pronk and Mark C.M. van Loosdrecht

Published in Applied Environmental Microbiology, 84:e01370-18 (2018), https://doi.org/10.1128/AEM.01370-18

## Abstract

Agricultural residues such as sugar beet pulp and citrus peel are rich in pectin, which contains galacturonic acid as a main monomer. Pectin-rich residues are underexploited as feedstocks for production of bulk chemicals or biofuels. Anaerobic, fermentative conversion of D-galacturonate in anaerobic chemostat enrichment cultures provides valuable information towards valorisation of these pectin-rich feedstocks. Replicate anaerobic chemostat enrichments, with d-galacturonate as sole limiting carbon source and inoculum from cow rumen content and rotting orange peels, yielded stable microbial communities, which were dominated by a novel *Lachnospiraceae* species, for which the name "*Candidatus* Galacturonibacter soehngenii" was proposed. Acetate was the dominant catabolic product, with formate and $H_2$ as co-products. The observed molar ratio of acetate and the combined amount of $H_2$ and formate deviated significantly from 1, which suggested that some of the hydrogen and $CO_2$ formed during D-galacturonate fermentation was converted into acetate via the Wood-Ljungdahl acetogenesis pathway. Indeed, metagenomic analysis of the enrichment cultures indicated that the genome of "*Candidatus* G. soehngenii" encoded enzymes of the adapted Entner-Doudoroff pathway for D-galacturonate metabolism as well as enzymes of the Wood-Ljungdahl pathway. Simultaneous operation of these pathways may provide a selective advantage under D-galacturonate-limited conditions by enabling a higher specific ATP production rate and lower residual D-galacturonate concentration than would be possible with a strictly fermentative metabolism of this carbon and energy source.

### Importance

This study on D-galacturonate metabolism by open, mixed culture enrichments under anaerobic, D-galacturonate-limited chemostat conditions shows a stable and efficient fermentation of D-galacturonate into acetate as the dominant organic fermentation product. This fermentation stoichiometry and population analyses provide a valuable baseline for interpretation of the conversion of pectin-rich agricultural feedstocks by mixed microbial cultures. Moreover, the results of this study provide a reference for studies on microbial metabolism of D-galacturonate under different cultivation regimes.

# Introduction



Conversion of agricultural and food-processing residues by open, mixed microbial communities is highly relevant for treatment of waste streams and increasingly also investigated as a strategy for production platform chemicals from low-value feedstocks. As an alternative to complete anaerobic conversion to methane and carbon dioxide, fermentative metabolism of these residues can yield more valuable products, including organic acids, solvents and transport fuels. Production of such compounds in non-aseptic open cultures has the potential to enable much lower capital and processing costs than currently achievable with pure-culture processes (1–3).

Continuously operated, non-aseptic cultures enrich for microbial consortia that thrive under the imposed conditions (1, 4). By carefully designing and optimizing reactor configurations and cultivation regimes, microbial populations can be enriched that specifically and reproducibly yield desired product profiles (5). Parameters such as culture pH, biomass retention time and temperature, either in steady-state or dynamic cultivation regimes, are among the key parameters influencing population composition and product profiles (6, 7). An additional decisive parameter in determining product formation by enrichment cultures is feedstock composition. Cell wall polymers and storage carbohydrates represent the majority of the fermentable compounds in most plant-derived waste streams. Hydrolysis of these polymers by extracellular microbial enzymes generates monomers, which are the actual substrates for anaerobic fermentation. Model studies with enrichment cultures on individual monomeric substrates can generate valuable information on the metabolic pathways and product profiles that may be encountered under industrially relevant conditions.

Previous studies on anaerobic fermentation of plant biomass monomers by continuous anaerobic enrichment cultures mainly focused on hexose and pentose sugars released from the cellulose and hemicellulose fractions of plant biomass (8–10). However, depending on the feedstock, other monomers can represent a significant or even the major fraction of the fermentable compounds in biomass hydrolysates. In particular, sugar beet pulp, citrus peel and apple pomace are rich in pectin, a polymer that contains D-galacturonic acid as a main constituent (11–14). Large-volume citrus peel waste streams originate from producing orange juice concentrates, which mainly occurs in the USA and Brazil. Sugar beet pulp and apple pomace are predominantly produced in Europe as a waste stream of sugar beet and apple processing, respectively (15, 16). These pectin-rich agricultural residues are currently mostly dried and sold as cattle feed, but the cost of the drying process reduces profitability of this application. In 2016 approximately 35 million tonnes of citrus peel, 28 million tonnes of sugar beet pulp and 2 million tonnes of apple pomace were produced in the world which, by dry weight, contain 17 %, 14 % and 16 % galacturonate, respectively (12–14, 17). These large-volume residues represent a potential and currently underexploited feedstock for production of bulk chemicals or biofuels.

Intensively studied sugar substrates such as glucose, xylose or arabinose can be fermented to commodity products, including ethanol and lactate, at near-theoretical yields. However, the currently documented biochemical pathways for galacturonate metabolism do not enable high carbon and electron yields of these fermentation products (18–20). A major constraint in known pathways for dissimilation of galacturonate, which is more oxidized than glucose, involve cleavage of a $C_6$-intermediate into two $C_3$-moieties, eventually leading to the redox-neutral generation of two moles of pyruvate from one mole of galacturonate (Figure 1) (21–26). As a consequence, metabolic reactions beyond pyruvate need to be redox balanced, for example by conversion of pyruvate into equimolar amounts of acetyl-CoA and formate (Figure 1). Consistent with this observation, engineered microbial strains exhibit low carbon yields on galacturonate for compounds that are more reduced than pyruvate, such as ethanol and lactate (11, 19, 20, 27).

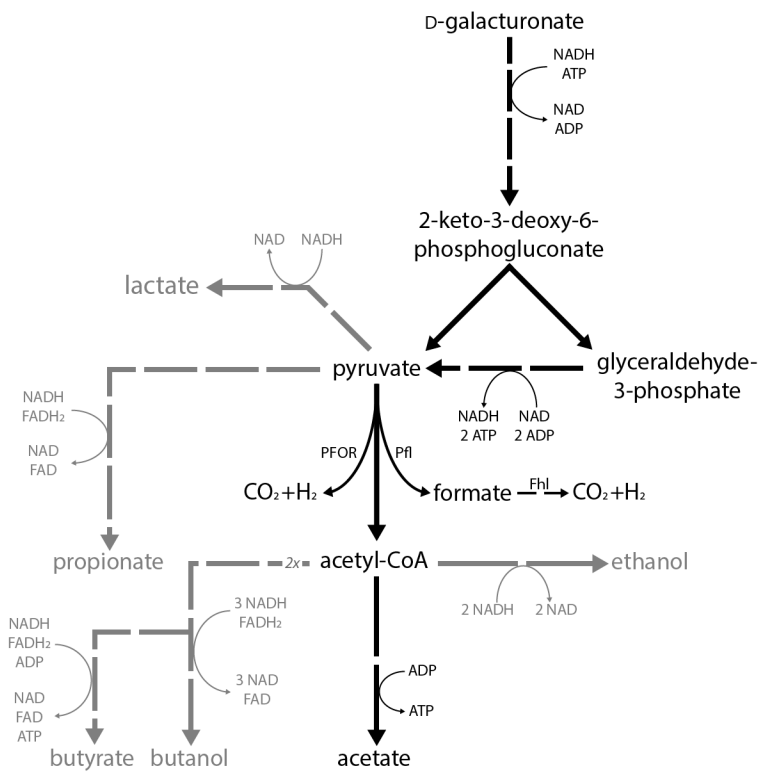



**Figure 5 |** Simplified scheme of pathways for fermentation of D-galacturonate to acetate and formate or to acetate, hydrogen and carbon dioxide. The black lines indicate conversions which are possible and grey lines indicate conversions which are not possible when fermenting D-galacturonate based on redox equivalents produced in the adapted Entner-Doudoroff pathway. The dashed lines represent multiple step conversions, with the number of dashes independent of the number of conversions. Abbreviations indicate the following enzyme activities: Pfl, pyruvate-formate lyase (EC 2.3.1.54); PFOR, pyruvate: ferredoxin oxidoreductase (EC 1.2.7.1) and Fhl, formate hydrogen lyase (EC 1.17.1.9 together with EC 1.12.1.2). Adapted from González-Cabaleiro et al. (26).

Despite its ecological and industrial significance as a plant-derived carbon source, fermentative conversion of galacturonate by anaerobic chemostat enrichment cultures has not been studied. The aim of this study was to characterize product profiles and microbial population of open, mixed-culture chemostat cultures fed with D-galacturonate as sole limiting substrate and compare these results to cultures fed with the much better studied neutral sugar glucose (8).

## Results

### Enrichment of galacturonate-fermenting microbial cultures in anaerobic chemostats.

To study composition and physiology of anaerobic microbial enrichment cultures on D-galacturonate, duplicate chemostat cultures were grown on a synthetic medium in which this pectin monomer was the sole carbon and energy substrate. Both bioreactors were started by adding anaerobic shake flask pre-cultures on the same medium, which had been inoculated with rotting orange peel and cow rumen content. The bioreactors were operated as batch cultures until all galacturonate was consumed and then switched to continuous operation at a dilution rate of 0.13 $h^{-1}$. This dilution rate and other parameters were chosen to enable direct comparison with previous experiments in glucose-limited chemostat cultures (8). Subsequently, product concentrations and carbon dioxide production were monitored over a period of six weeks. Constant rates of $CO_2$ production and of metabolite concentrations in the chemostats indicated that a 'metabolic steady state' was reached after two weeks of continuous cultivation (approximately 63 generations of the overall microbial population, Figure 2). In the steady-state cultures, the D-galacturonate concentration remained below the detection limit of 0.1 mM, consistent with the medium composition, which was designed to make D-galacturonate the growth-limiting nutrient. After one month of steady-state operation (approximately 135 generations), five samples were taken over a period of 5 days for further physiological and metagenomics analysis.

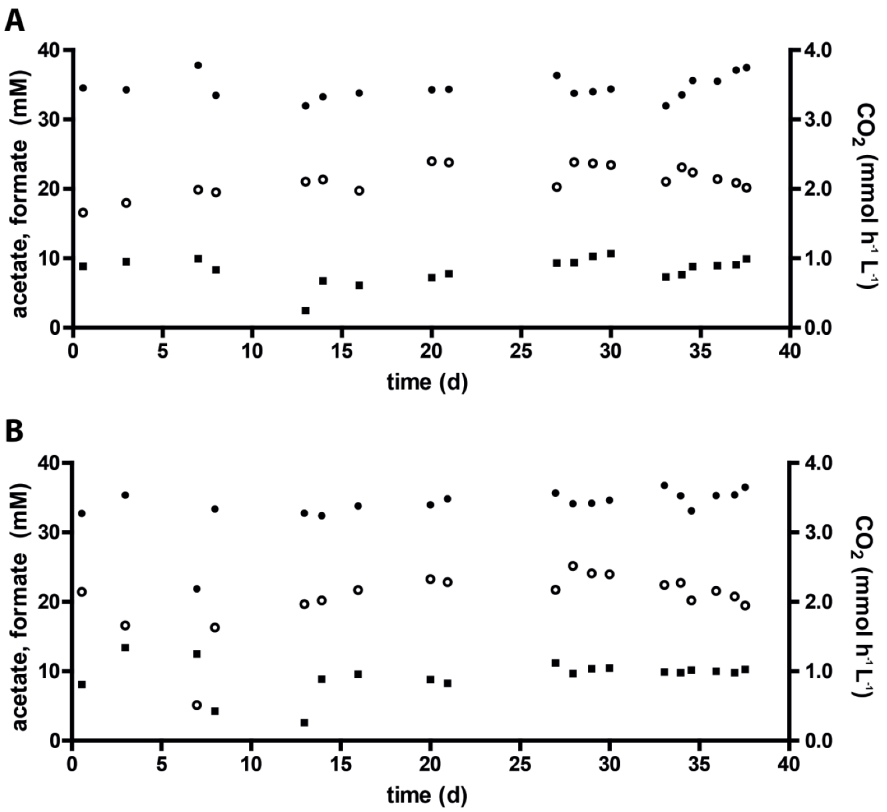



**Figure 6 |** Concentrations of catabolic products during anaerobic enrichment experiments in galacturonate-limited chemostat cultures of bioreactor 1 (panel A) and bioreactor 2 (panel B) with continuous cultivation starting at day 1. Symbols: acetate (black circles), and formate (open circles) concentrations in mM and $CO_2$ production (black squares) in mmol $h^{-1}$ $L^{-1}$. Both bioreactors were operated at a dilution rate of 0.13 $h^{-1}$, at pH 8 and at 30 °C.

## Product profiles in chemostat enrichment cultures suggest involvement of acetogenesis in anaerobic galacturonate metabolism.

To investigate galacturonate metabolism, fermentation products were analysed in culture supernatants and off-gas of the steady-state enrichment cultures. In both reactors, acetate was the main catabolic product of D-galacturonate fermentation, with additional production of $H_2$ and formate (Table 1). Average recoveries of carbon (92 %) and electrons (95 %) in the steady-state cultures indicated that major fermentation products were accounted for (Supplementary Table S1).

**Table 3 |** Product yields of anaerobic chemostat enrichment cultures on galacturonate. Replicate enrichment cultures (Bioreactor 1 and 2) were grown on synthetic medium with galacturonate as limiting substrate, at pH 8, at 30 °C and at a dilution rate of 0.13 $h^{-1}$. For comparison, the product yields of glucose-limited enrichment cultures grown by Temudo et al. under identical conditions are given (8). Yields of biomass and products on galacturonate or glucose are expressed as Cmol $Cmol^{-1}$ unless stated otherwise. Data represent average ± standard deviation of five sequential measurements during steady state.

| | Galacturonate | | Glucose |
|---|---|---|---|
| | **Bioreactor 1** | **Bioreactor 2** | **(8)** |
| **Biomass** | 0.16 ± 0.01 | 0.16 ± 0.02 | 0.21 ± 0.01 |
| **Succinate** | < 0.01 | 0.02 ± 0.01 | - |
| **Acetate** | 0.53 ± 0.02 | 0.53 ± 0.02 | 0.16 ± 0.02 |
| **Butyrate** | < 0.01 | < 0.01 | 0.30 ± 0.04 |
| **Ethanol** | < 0.01 | < 0.01 | 0.08 ± 0.01 |
| **Formate** | 0.16 ± 0.01 | 0.16 ± 0.01 | 0.22 ± 0.01 |
| **$H_2$ (mol $Cmol_s^{-1}$)** | 0.01 ± 0.00 | 0.02 ± 0.00 | 0.05 ± 0.01 |
| **$CO_2$** | 0.05 ± 0.00 | 0.06 ± 0.00 | - |
| **$H_2$ + Formate (mol $Cmol_s^{-1}$)** | 0.17 ± 0.02 | 0.18 ± 0.01 | 0.27 ± 0.01 |
| **Acetyl-CoA derivatives (mol $Cmol_s^{-1}$)[a]** | 0.27 ± 0.02 | 0.27 ± 0.02 | 0.27 ± 0.01 |

a. The sum of the acetyl-CoA derivatives: acetate, butyrate and ethanol as mol acetyl-CoA per Cmol of substrate.

In contrast to previously analysed anaerobic chemostat enrichment experiments on glucose, which were grown under the same conditions (8) as used in the present study, the galacturonate-grown enrichment cultures did not produce detectable amounts of reduced fermentation products such as butyrate or ethanol (Table 1). Qualitatively, formation of acetate, formate, $CO_2$ and $H_2$ as major fermentation products was consistent with previously described pathways for bacterial galacturonate metabolism (Figure 1). However, in contrast to the product distribution expected from such pathways and to previous observations on glucose-grown chemostat enrichment cultures (8), the sum of the acetyl-CoA-derived products significantly differed from the combined formate and $H_2$ concentrations. In bioreactors 1 and 2, the combined yields of hydrogen and formate on D-galacturonate (mol $mol^{-1}$) were 37 ± 2 % and 37 ± 1 % lower, respectively, then the corresponding yields of acetate (Table 1). The observed metabolite profiles would be consistent with a contribution of acetogenesis via the Wood-Ljungdahl (WL) pathway, which enables conversion of fermentatively produced hydrogen (or alternatively, reduced ferredoxin derived from a pyruvate:ferredoxin oxidoreductase, EC 1.2.7.1) and carbon dioxide to acetate (Equation 1). Combination of acetogenesis with fermentative conversion of D-galacturonate into acetate via a classical modified Entner-Doudoroff (ED) route (Equation 2, Figure 1) allows for conversion of D-galacturonate to acetate without net production of hydrogen (Equation 3, Figure 3).

(1) $2\ H_2 + 1\ CO_2 \rightarrow 0.5\ C_2H_4O_2 + 1\ H_2O$

(2) $C_6H_{10}O_7 + 1\ H_2O \rightarrow 2\ C_2H_4O_2 + 2\ H_2 + 2\ CO_2$

(3) $C_6H_{10}O_7 \rightarrow 2.5\ C_2H_4O_2 + 1\ CO_2$

Based on the assumption that D-galacturonate catabolism in the enrichment cultures indeed reflected simultaneous D-galacturonate fermentation via an ED pathway and acetogenesis via the WL pathway, distribution of carbon and electrons over both pathways can be easily estimated from measured rates of D-galacturonate consumption and of hydrogen, formate and acetate production in the steady-state cultures (three equations with two unknowns (28)). Based on these calculations (Supplementary calculations S2) the average specific rates of acetate formation from pyruvate generated in the ED pathway ($q_{acetate,ED}$) and of acetogenesis via the WL pathway ($q_{acetate,WL}$) were estimated to be 6.9 ± 0.4 mmol (g biomass)$^{-1}$ h$^{-1}$ and 1.7 ± 0.2 mmol (g biomass)$^{-1}$ h$^{-1}$, respectively. In such a scenario, 16 ± 3 % of the produced hydrogen and formate generated from D-galacturonate would have been used for acetogenesis (Figure 3).

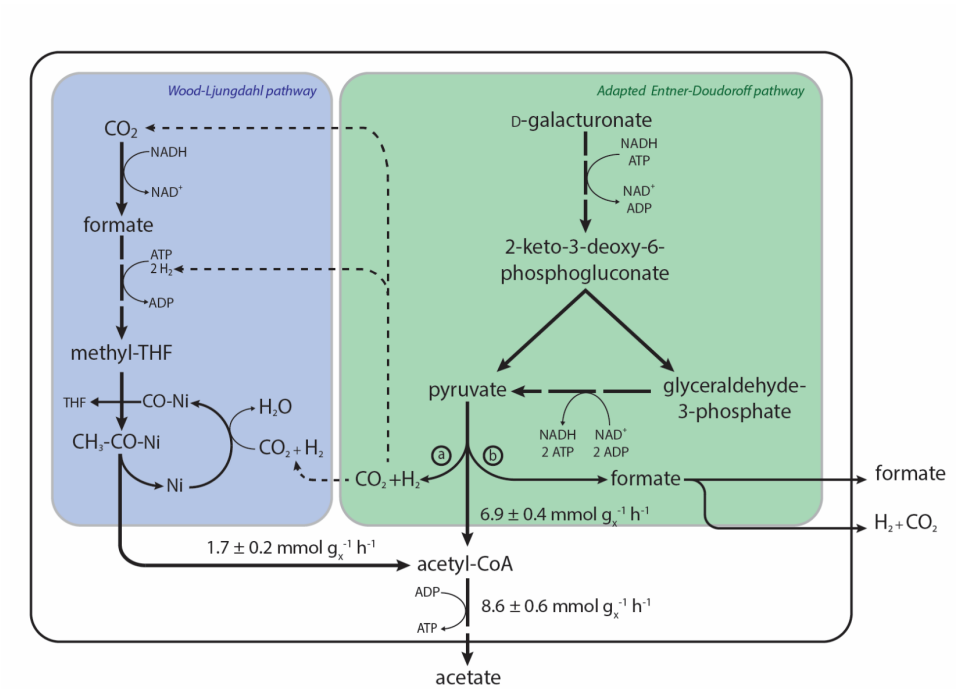



**Figure 7 |** Metabolic model of D-galacturonate catabolism in anaerobic chemostat enrichment cultures. Metabolic fluxes were calculated by mass balancing based on a simplified Wood-Ljungdahl pathway converting four $H_2$ and two $CO_2$ into one acetate (42) and D-galacturonate metabolism via an adapted Entner-Doudoroff pathway for galacturonate catabolism yielding two molecules of pyruvate (21–25). Redox equivalents have been represented as [H] and could both be derived from $H_2$ as well as NADH. Biomass (x)-specific production rates (mmol (gx)$^{-1}$ h$^{-1}$), represent average ± standard deviation of five sequential measurements during steady state of both biological duplicate bioreactors.

**Population analysis of anaerobic D-galacturonate-fermenting enrichment cultures.**

To study microbial population dynamics during enrichment and to assess if the hypothesis outlined above was supported by the presence of acetogenic micro-organisms, 16S-rRNA gene amplicon sequencing was performed on samples from both bioreactors. The inoculum, used to inoculate both bioreactors, contained species related to the genera C*lostridium* sensu stricto and *Lachnoclostridium* as well as to uncultured *Lachnospiraceae* (Figure 4). After complete consumption of D-galacturonate in the batch phase preceding the continuous culture, bioreactor 1 was dominated by bacteria related to *Raoultella* and *Yersinia*, with the latter showing a 99 % identity to *Yersinia massiliensis* in the SILVA SSU database release 128. At the same sampling time, Bioreactor 2 was dominated by bacteria related to *Klebsiella* and *Enterobacter*. Despite this difference in population composition, product profiles were similar for both bioreactors at this time point, with acetate as the dominant organic catabolic product and formate as the second major product (Figure 2).

Although the microbial community compositions were different at the start of the continuous enrichment, similar populations developed during prolonged continuous operation of the two bioreactors. In both reactors a bacterium related to *Lachnotalea* became dominant, with a side population of a *Klebsiella*-related bacterium (Figure 4). The steady-state communities in the two replicates were highly similar, indicating that the chosen cultivation conditions generated a selective advantage for these species and their conversions.

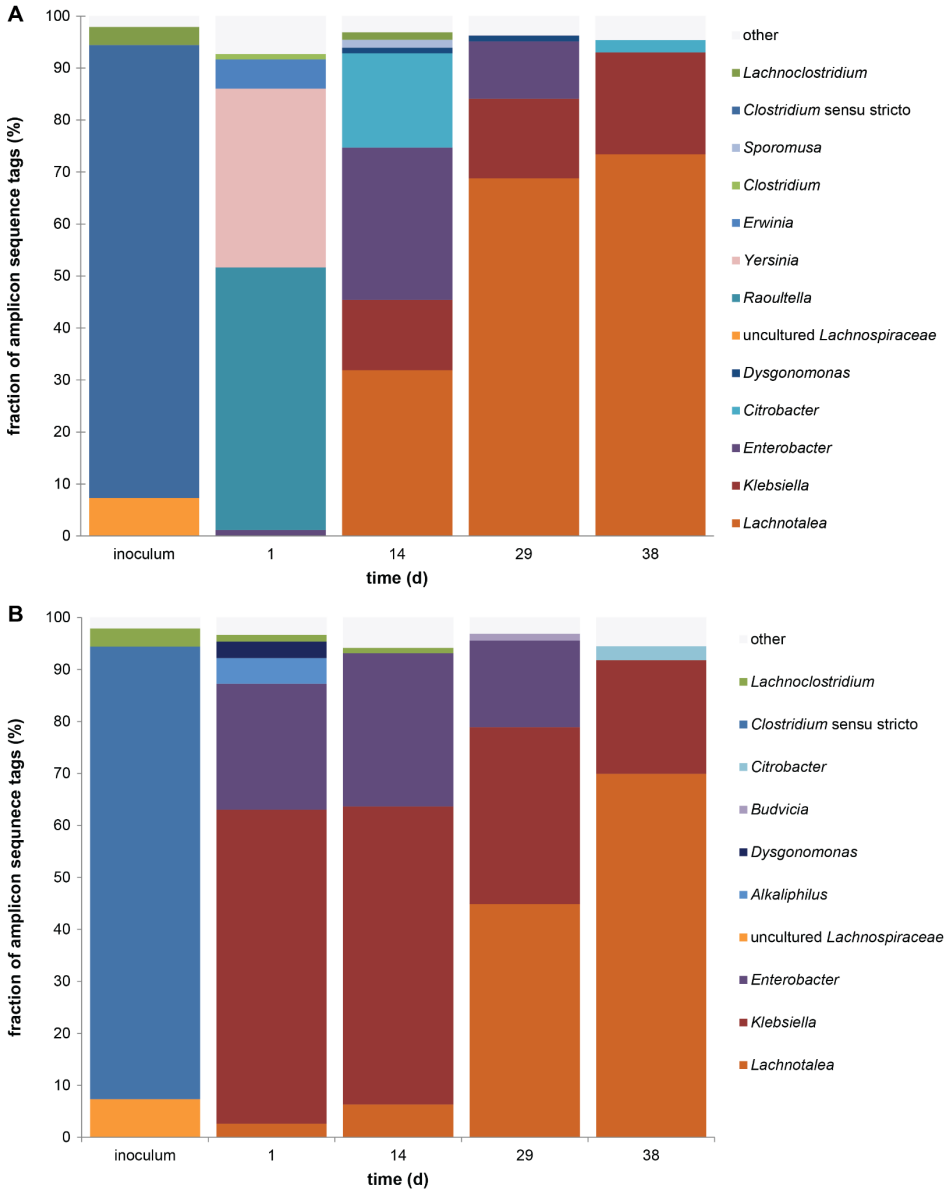



**Figure 8 |** Population dynamics in anaerobic chemostat enrichment cultures on galacturonate. Genus-level population compositions were derived from the 16S-rRNA gene amplicon sequencing data which make up ≥ 1 % of the amplicon sequences. A) population dynamics in bioreactor 1, B) population dynamics in bioreactor 2. Both bioreactors were operated at a dilution rate of 0.13 $h^{-1}$, at pH 8 and at 30 °C.

**Community distribution determination by fluorescence *in situ* hybridization (FISH) analysis.**

To further quantify the abundance of dominant community members during the steady state measurements, FISH analysis (29, 30) was performed (Figure 5). A general probe (EUB338) was used to stain all bacterial cells while, based on the results of the 16S-rRNA gene analysis, a probe specific for the *Lachnospiraceae* family, (Lac435) and a probe for the *Enterobacteriaceae* family (ENT) were used to quantify bacteria related to *Lachnotalea* and *Klebsiella*, respectively. Microorganisms belonging to the *Lachnospiraceae* were found to be abundant in both reactors, with another abundant population of bacteria belonging to the *Enterobacteriaceae* (Figure 4 and 5). These observations correlated with the identification of a *Lachnotale*- and a *Klebsiella*-related bacterium, respectively, in the 16S-rRNA gene analysis.

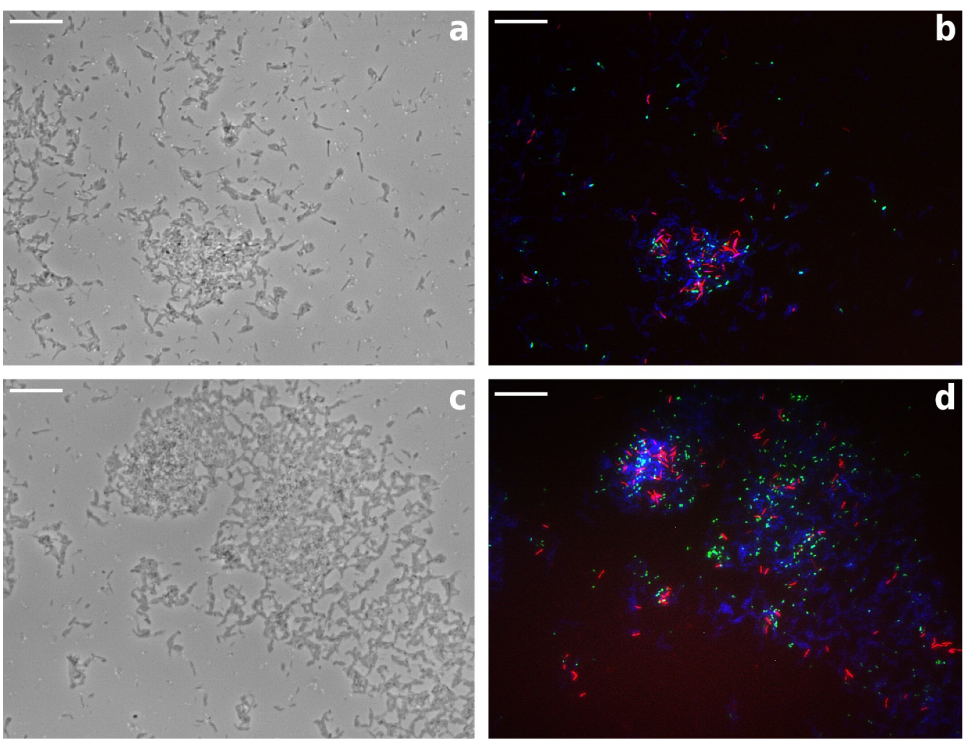



**Figure 9 |** Fluorescence in situ hybridization (FISH) of anaerobic chemostat enrichment cultures. Culture samples taken at day 37 in two parallel enrichment cultures were stained with a general probe for bacteria (EUB338, blue), a probe against members of the Lachnospiraceae family (Lac435, red) and a probe against members of the Enterobacteriaceae family (ENT, green). Panels A and B indicate unstained and stained cells, respectively, in bioreactor 1, while Panels C and D show the same information for bioreactor 2. Scale bars in all panels represent 20 µm. Both bioreactors were operated at a dilution rate of 0.13 $h^{-1}$, at pH 8 and at 30 °C.

**Metagenomic analysis of pathways and species in the enrichment cultures.**

To investigate whether the enriched *Lachnotalea*-related bacterium harboured genes encoding enzymes of the WL pathway and for DNA-based identification of the dominant micro-organisms, a metagenomic analysis was performed on the steady-state cultures. DNA of the entire community was extracted from both steady-state enrichment cultures and sequenced. In total 14.9 million paired-end reads remained post trimming, 7.9 Gbp of sequenced bases were obtained for each of the enrichments. After *de novo* assembly and metagenome binning based sequence composition and differential coverage, draft genomes of the dominant micro-organisms were assembled (31–35), checked for completeness and contamination and annotated with the RAST server (36–39).

The genome with the highest coverage (577 ± 94 fold, completeness 98 %) in both enrichments was identified to be a member of the *Lachnospiraceae* by analysis of the small subunit ribosomal RNA gene of the constructed genome. An identity of 93% was found within the SILVA SSU database release 132 with the closest cultured relative *Lachnotalea glycerini*, indicating that the dominant microorganism belongs to a novel genus within the *Lachnospiraceae* family (40, 41). We propose the name "*Candidatus* Galacturonibacter soehngenii" for this novel microbe. The assembled and annotated genome of this organism harboured all structural genes for the enzymes of the canonic ED pathways for D-galacturonate catabolism, as well as multiple signature genes encoding enzymes of the WL pathway for acetogenesis, including *fhs*, which encodes formate-tetrahydrofolate ligase (EC 6.3.4.3) and is specific for the WL pathway (42–44) (Table 2). However, as shown in Table 2, the genes *cooS,* encoding carbon-monoxide dehydrogenase (EC 1.2.7.4), and *acsBCD*, which encodes the CO-methylating acetyl-CoA synthase complex (EC 2.1.3.169), were not identified.

**Table 4 |** Genes of the Wood-Ljungdahl and adapted Entner-Doudoroff pathways identified in the genome of "Candidatus G. soehngenii". The name, EC number, genetic nomenclature of the genes with the highest annotation found and their corresponding expected value (E value) are shown, with '-' indicating no homologues were identified.

| Gene name | EC number | Gene | E value |
|---|---|---|---|
| ***Wood-Ljungdahl pathway*** | | | |
| **Hydrogen dehydrogenase** | 1.12.1.2 | *hoxF* | 6.00 $e^{-33}$ |
| **Formate dehydrogenase** | 1.17.1.9 | *fdhA* | 6.00 $e^{-33}$ |
| **Formate–tetrahydrofolate ligase** | 6.3.4.3 | *fhs* | 0.00 |
| **Methenyl-tetrahydrofolate cyclohydrolase/Methylene – tetrahydrofolate dehydrogenase** | 3.5.4.9 and 1.5.1.5 | *foldD* | 4.00 $e^{-117}$ |
| **Methyl–tetrahydrofolate reductase** | 1.5.1.20 | *metF* | 0.00 |
| **5-Methyl-tetrahydrofolate:corrinoid/iron-sulphur protein Co-methyltransferase** | 2.1.1.258 | *acsE* | 6.00 $e^{-32}$ |
| **CO-methylating acetyl-CoA synthase** | 2.3.1.169 | *acsBCD* | - |
| **Carbon-monoxide dehydrogenase** | 1.2.7.4 | *cooS* | - |
| ***Adapted Entner-Doudoroff pathway*** | | | |
| **Uronate isomerase** | 5.3.1.12 | *uxaC* | 0.00 |
| **Tagaturonate reductase** | 1.1.1.58 | *uxaB* | 0.00 |
| **Altronate dehydratase** | 4.2.1.7 | *uxaA* | 0.00 |
| **2-Dehydro-3-deoxygluconokinase** | 2.7.1.45 | *kdgK* | 2.00 $e^{-135}$ |
| **2-Dehydro-3-deoxyphosphogluconate aldolase** | 4.1.2.14 | *kdgA* | 2.00 $e^{-135}$ |
| **Pyruvate formate-lyase** | 2.3.1.54 | *pfl* | 2.00 $e^{-91}$ |
| **Pyruvate:ferredoxin oxidoreductase** | 1.2.7.1 | *PFOR* | 0.00 |
| **Phosphate acetyltransferase** | 2.3.1.8 | *pta* | 2.00 $e^{-142}$ |
| **Acetate kinase** | 2.7.2.1 | *ackA* | 4.0 $e^{-163}$ |

A complete set of WL pathway genes was identified in the metagenomic dataset within the genome of a species closely related to a *Sporomusa* species (99 % identity found in SILVA SSU database release 132), within a genus known for harbouring homoacetogenic bacteria (45). However, the genome of the *Sporomusa* species had a coverage of only 9 ± 2 fold, and a completeness of 80%. This coverage corresponds to approximately 1 ± 0 % of the total coverage. The other genome with a high coverage (166 ± 12 fold, completeness 100 %), was closely related to *Klebsiella oxytoca* (100 % identity of the 16S-rRNA gene in

the SILVA SSU database release 132). This observation was consistent with the 16S-rRNA gene amplicon sequencing data, which showed a side population of a bacterium closely related to *Klebsiella* species. The *K. oxytoca* genome did not harbour any WL pathway genes but, in accordance with literature (46), did contain all genes for the adapted ED pathway for D-galacturonate metabolism.



## Discussion

Identification of acetate as dominant catabolic product of anaerobic chemostat enrichment cultures on D-galacturonate (Table 1) is consistent with early studies on the metabolism of pectin-fermenting bacteria (47–49). The narrow product range of the currently described D-galacturonate-fermenting organisms is likely to be a consequence of the redox-neutral conversion of D-galacturonate to pyruvate via a modified Entner-Doudoroff (ED) pathway (Figure 1, (18)). In contrast, glucose fermentation, in which oxidative glycolytic pathways generate reduced cofactors that are subsequently reoxidized by reduction of pyruvate and/or acetylphosphate, generates a wide diversity of fermentation products. In enrichment cultures on glucose, performed under conditions identical to those used in the present study, a 2:1 molar ratio of butyrate to formate or hydrogen in such cultures was consistent with a classical Embden-Meyerhof glycolysis and subsequent reduction of pyruvate (Table 1, (8, 50)).

When D-galacturonate is metabolised to pyruvate via a modified ED pathway, subsequent redox-neutral conversion of pyruvate into acetate and formate can occur via pyruvate formate-lyase (EC 1.2.7.1), phosphate acetyltransferase (EC 2.3.1.8), and acetate kinase (EC 2.7.2.1) ($C_6H_{10}O_7 \rightarrow 2\ C_2H_4O_2 + 2\ HCOOH$), with a potential further conversion of formate to hydrogen and $CO_2$ by formate hydrogen-lyase (EC 1.12.1.2 and EC 1.17.1.9; Figure 1, (26, 50, 51)). However, since the 1:1 molar ratio of acetate to formate or hydrogen expected in such a scenario was not observed, we hypothesized that, in the chemostat enrichment cultures, acetogenesis via the Wood-Ljungdahl (WL) pathway converted some of the hydrogen and $CO_2$ produced during conversion of pyruvate to acetate.

The enrichment cultures contained different species suspected to be capable of acetogenesis. Based on sequence coverage estimate, a member of the *Sporomusa* genus, which is well-known to harbour autotrophic acetogens (45), made up less than 1 % of the population in the enrichment cultures. If this bacterium was solely responsible for the inferred rates of acetogenesis in the cultures, its specific activity would have to exceed previously reported rates of acetogenesis (52, 53) by at least one order of magnitude. This makes it likely that another bacterium was responsible for the majority of the acetogenesis.

The dominance of the "*Candidatus* G. soehngenii" that was clearly shown in both the 16S-rRNA gene (Figure 4) and metagenome analysis was not matched by the FISH analysis (Figure 5). This discrepancy was most probably due to the reported difficulty of staining Gram-positive micro-organisms (54, 55). We therefore assume that "*Candidatus* G. soehngenii", while underrepresented in the FISH analysis, was indeed dominant and responsible for the main conversions observed in both bioreactors.

After the onset of continuous cultivation, "*Candidatus* G. soehngenii" gradually replaced other microorganisms (Figure 3). The genome of this organism, which became dominant at the end of duplicate chemostat enrichment experiments, harboured most genes of the WL pathway. However, we did not identify genes with a high sequence similarity to known *cooS* and *acsBCD* genes encoding for the respectively carbon-monoxide dehydrogenase (EC 1.2.7.4) and CO-methylating acetyl-CoA synthase complex subunits (EC 2.1.3.169 (42)). Genomes of *Lachnospiraceae* species remain underexplored and contain many non-annotated genes and the possibility of novel functional homologues of the Wood-Ljungdahl pathway might be interesting for future research (56). Alternatively, the 'missing' *cooS* and *acsBCD* genes may be an artefact of the metagenome assembly and analysis. All traditional methods were used in an attempt to isolate "*Candidatus G.* soehngenii" (for a detailed list of isolation methods see Supplementary data Table S2), but none was successful. We postulate the "*Candidatus* G. soehngenii" is very well-adapted to growth at limiting substrate concentrations (K-strategist) as occurring in a chemostat operated bioreactor (57). Alternative isolation procedures not relying on batch growth might be needed for successful isolation. Although confirmation will require isolation of a pure culture, we expect the "*Candidatus* G. soehngenii" to be capable of simultaneous fermentation of D-galacturonate and acetogenesis.

Microbial competition is often interpreted in terms of the affinity for a single growth-limiting nutrient ($q_s^{max}$ / $K_s$) (58). Under nutrient-limited conditions, simultaneous utilization of mixed substrates enables growth at lower concentrations of each of the substrates than would be possible when growth is limited by a single substrate (59, 60). In this way, re-consumption of hydrogen could enable lower residual galacturonate concentrations in the chemostat cultures, thereby conferring a selective advantage to "*Candidatus* G. soehngenii". Additionally, the net ATP yield of D-galacturonate fermentation via an ED-type pathway amounts to 3 mol ATP (mol galacturonate)$^{-1}$ or 1.5 mol ATP (mol acetate)$^{-1}$, while extra formation of acetate via the WL pathway is estimated to yield an additional 0.33 mol ATP (mol acetate)$^{-1}$, assuming chemiosmotic energy conservation due to the generation of a transmembrane electrochemical ion gradient (61). With a constant growth rate under chemostat operations the total ATP flux, $q_{ATP}$, required for growth will also remain constant. When the ATP is supplied by both the adapted ED pathway as well as the WL pathway, the relative $q_{ATP}$ derived from the ED pathway can decrease, without lowering the overall $q_{ATP}$. Lowering the flux through the ED pathway decreases the galacturonate biomass specific uptake rate ($q_s$) and in

accordance to the substrate uptake-kinetics, shown in Equation 4, this will decrease the galacturonate concentration ($C_s$) in the bioreactor (62).



(4) $$q_s = q_{s,max} \cdot \frac{C_s}{C_s + K_s}$$

The very slow replacement of non-acetogenic bacteria in the chemostat enrichment cultures is in line with the relatively small impact of acetogenesis on the kinetics and energetics of D-galacturonate metabolism. Moreover, the dilution rate of the chemostat cultures was at the upper end of specific growth rates reported for typical heterotrophic homoacetogens on (semi-) defined media (63–65). Additionally, sparging of the reactors with nitrogen gas may have stripped $H_2$ formed by D-galacturonate fermentation. Estimated hydrogen partial pressures in the reactors were $2 - 5 \cdot 10^{-3}$ atm, resulting in a Gibbs free energy change of -17.3 ± 24.0 kJ $mol^{-1}$ for the WL pathway (for calculations see Supplementary data calculations S3, (66, 67)). With the Gibbs free energy change close to the minimal driving force for catabolic reactions, this low *in situ* hydrogen partial pressure may have limited the rate of $H_2$ oxidation in acetogenesis. This interpretation is consistent with the dominance of fast-growing non-acetogenic bacteria (e.g. *Klebsiella* and *Clostridium* species) during the batch phase that preceded the chemostat enrichment.

Although involvement of the WL pathway has been demonstrated in anaerobic chemostat enrichment cultures on glucose at a low dilution rate (0.05 $h^{-1}$, (66, 68)), such an involvement was not observed at higher dilution rates (with operational conditions similar to those used in this study (8)). This observation suggests that potentially lowering the residual glucose concentration by simultaneous acetogenesis did not provide a selective advantage in anaerobic enrichment cultures fed with glucose. More research is required to elucidate this difference in usage of the Wood-Ljungdahl pathway in glucose- and galacturonate-limited cultures. Although "*Candidatus* G. soehngenii", which became dominant in the D-galacturonate-limited enrichment cultures in this study, differed from the *Clostridium quinii* that dominated similar cultures grown on glucose as the carbon source (69), both are members of the *Clostridiales*. This observation illustrates the importance of these organisms in fermentative conversion of carbohydrates and related compounds in carbon-limited, anaerobic environments (69).

## Material and Methods

**Reactor Operation.** A continuous stirred reactor of 1.2 L capacity (0.5 L working volume) was used (mechanical stirring 300 rpm). Water was recirculated to maintain a constant temperature at 30 °C. The reactor was sparged with nitrogen gas at a flow rate of 120 mL $min^{-1}$, to maintain anaerobic conditions. The pH was controlled at pH 8 ± 0.1 by automatic titration (ADI 1030 Bio controller) with a 1 M NaOH solution. The dilution rate was 0.13 $h^{-1}$ and the working volume was kept constant by peristaltic effluent pumps (Masterflex, Cole-Parmer, Vernon Hills, USA) coupled to electrical level sensors.

**Steady State Characterization.** To characterize the product spectrum, the reactor was run in continuous mode until a stable product composition and biomass concentration was established. Steady state was reached after two weeks (63 generations). On-line analysis of system stability was achieved by online monitoring of the gas productivities, in the form of $CO_2$, and base addition. On-line gas detection of $CO_2$ was done with a Rosemount Analytical NGA 2000 MLT 1 Multicomponent analyser (infrared detector). Data acquisition was achieved with SCADA software (Sartorius BBI systems MFCS/win 2.1). When these rate measurements were stable or varied within a limited range (± 10 %) without a trend showing increase or decrease, a set of samples was taken during the subsequent cycles. Concentrations of soluble organic fermentation products, hydrogen produced and biomass concentration in the reactor volume were determined.

**Inoculum**. The reactor was inoculated (1 % v/v) with open mixed pre-cultures started with a mixture of two different sources. The first inoculum was obtained from the content of rumen offal of a grass-fed cow (Est, The Netherlands). The second inoculum was a sample from organic waste containing a high amount of citrus peels (Orgaworld Nederland B.V., The Netherlands). Shake flasks were inoculated in duplicate with compost (2 % v/v) and rumen extract (2 % v/v) and the initial pH was adjusted to 8 with a 2 M KOH solution. The shake flasks were cultivated in an anaerobic chamber (Bactron III, Shel Lab, Cornelius, USA) at 30 °C with a gas composition of 89 % $N_2$, 6 % $CO_2$ and 5 % $H_2$.

**Medium.** The cultivation medium contained the following (in g $L^{-1}$): D-galacturonate 4.0; $NH_4Cl$ 1.34; $KH_2PO_4$ 0.78; $Na_2SO_4.10H_2O$ 0.130; $MgCl_2.6H_2O$ 0.120; $FeSO_4.7H_2O$ 0.0031; $CaCl_2$ 0.0006; $H_3BO_4$ 0.0001; $Na_2MoO_4.2H_2O$ 0.0001; $ZnSO_4.7H_2O$ 0.0032; $CoCl_2.H_2O$ 0.0006; $CuCl_2.2H_2O$ 0.0022; $MnCl_2.4H_2O$ 0.0025; $NiCl_2.6H_2O$ 0.0005; EDTA 0.20. The D-galacturonate and mineral solutions were prepared and fed separately. The D-galacturonate solution was sterilized by filtration (0.2 µm Mediakap Plus, Spectrum Laboratories, Rancho Dominguez, USA) and the mineral solution was autoclaved for 20 min at 121°C. 3 mL Pluronic®PE 6100 antifoam (BASF, Ludwigshafen, Germany) was added per 40 L mineral solution to avoid excessive foaming.

**Analytical methods of substrate and extracellular metabolites.** To determine substrate and extracellular metabolite concentration, reactor sample supernatant was

obtained by centrifugation of culture samples (Biofuge Pico, Heraeus, Hanau, Germany) and subsequent filtration (0.2 µm Millex®-HV, Millipore-Merck, Darmstadt, Germany). Concentrations of galacturonate and extracellular metabolites were analysed using an Agilent 1100 Affinity HPLC machine (Agilent Technologies, Amstelveen, The Netherlands) with an Aminex HPX-87H ion-exchange column (BioRad, Hercules, USA) operated at 60 °C with a mobile phase of 5 mM $H_2SO_4$ and a flow rate of 0.6 mL $min^{-1}$. Measurements of $H_2$ and $CO_2$ during steady state analysis were performed off-line using gas-bags (3.8 L, Tedlar, Saint Gobain, France) and analysed using a mass spectrometer (Prima BT MS, Thermo scientific, USA).

**Biomass.** Culture dry weight was measured by filtering 20 mL of culture broth over pre-dried and pre-weighed membrane filters (0.2 µm Supor®-200, Pall corporation, New York, USA), which were then washed with demineralized water, dried in a microwave oven (Easytronic M591, Whirlpool, Michigan, USA) and weighed again. Carbon balances and electron balances were established based on the number of carbon atoms and electrons per mole, while a standard biomass composition of $CH_{1.8}O_{0.5}N_{0.2}$ was assumed (70).

**Microbial community stability analysis.** The bacterial DNA for the community analysis over time was obtained by extracting 2 mL broth and pelleting the biomass before storing it at -80 °C for further analysis. The DNA was extracted using the UltraClean DNA isolation kit (MOBIO laboratories Inc., USA), following the manufactures instructions. The 16S-rRNA genes of day 1 and 38 of bioreactor 1 and day 38 of bioreactor 2 were analysed using amplicon sequencing with an Illumina HiSeq (Novogene Bioinformatics Technology Co., Ltd), for this the primers 341F (5'- CCTAYGGGRBGCASCAG - 3') and 805R ( 5' – GGACTACNNGGGTATCTAAT - 3') were used generating 250 bp paired end reads. Sequences were analysed according to van den Berg et al. (71). Representative sequence for each operational taxonomic unit were submitted for BLAST analysis under default conditions. The extracted genomic DNA of the inoculum, day 9 and 14 of bioreactor 1 and day 1, 9 and 14 of bioreactor 2 were subsequently used for a two-step PCR reaction targeting the 16S-rRNA gene of most bacteria and archaea. For this the primers, U515F (5' – GTGYCAGCMGCCGCGGTA - 3') and U1071R (5'- GARCTGRCGRCRRCCATGCA- 3') were used according to Wang et al. (72). The first amplification was performed to enrich for 16S-rRNA genes, using a quantitative PCR assay with 2x iQ SYBR Green Supermix (Bio-rad, CA, USA), 500 nM primers each and 1 - 50 ng of genomic DNA added per well to a final volume of 20 µl. The thermocycles consisted of; first denaturation of 95 °C for 5 min and 20 cycles of 95 °C for 30 s, 50 °C for 40 s, 72 °C for 40 s and a final extension of 72 °C for 7 min. 454-adapters (Roche) and MID tags at the U515F primer were added to the produced PCR products. The second amplification was similar to the first one, with the exception of the use of Taq PCR Master Mix (Qiagen Inc, CA, USA), the program was run for 15 cycles and the template (the product from step one) was diluted ten times. After the second amplification, the PCR products were pooled equimolar and purified over an agarose gel using a GeneJET Gel Extraction Kit (Thermo Fisher Scientific, The Netherlands). The

resulting library was sent for 454 sequencing and run in 1/8 lane with titanium chemistry by Macrogen Inc. (Seoul, Korea). After analysis the reads library was imported into CLC genomics workbench v7.5.1 (CLC Bio, Aarhus, DK) and (quality, limit = 0.05) trimmed to a minimum of 200 bp and average of 284 bp and de-multiplexed. A build-it SILVA 123.1 SSURef Nr99 taxonomic database was used for BLASTn analysis on the reads under default conditions. The top result was imported into an excel spreadsheet and used to determine taxonomic affiliation and species abundance. Genome sequence data of the amplicon sequences of the 16S-rRNA gene analysis have been deposited at the NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/) with the corresponding accession numbers MG982381 - MG982451.

**Fluorescence *in situ* hybridization (FISH).** Fluorescence *in situ* hybridization was performed as described by (73), using a hybridization buffer containing 25% (v/v) formamide. Probes were synthesized and 5' labelled with either the FLUOS dye or with one of the sulfoindocyanine dyes Cy3 and Cy5 (Thermo Hybaid Interaciva, Ulm, Germany) (Table 3). The general probe EUB338 labelled with Cy5 was used to indicate all eubacterial species in the sample. Slides were observed with an epifluorescence microscope (Axioplan 2, Zeiss, Sliedrecht, The Netherlands), and images were acquired with a Zeiss MRM camera and compiled with the Zeiss microscopy image acquisition software (AxioVision version 4.7, Zeiss). The probes were excitated at 550 nm for Cy3, 494 for FLUOS, and 649 nm for Cy5, and images were captured at an emission of respectively 570 nm, 516 nm, and 670 nm.

**Table 5 |** FISH probes used in this study with the dyes, sequences, and specificity. With W indicating A or T.

| Probe | Dye | Sequence 5' - 3' | Specificity | Reference |
|---|---|---|---|---|
| **ENT** | Fluos | CCCCCWCTTTGGTCTTGC | *Enterobacteriaceae* except *Proteus* spp. | (82) |
| **Lac435** | CY3 | TCTTCCCTGCTGATAGA | *Lachnospiraceae* family | (83) |
| **EUB338** | CY5 | GCTGCCTCCCGTAGGAGT | Most bacteria | (84, 85) |

**Metagenomics.** 250 mL of the culture broth was centrifuged for 10 min at 4696 x g (Sorvall Legend X1R centrifuge, ThermoFisher Scientific) and the supernatant was discarded and the pellet was washed with TE (pH 8) buffer and stored at -20 °C. DNA for library construction was extracted with the Genomic DNA kit in combination with Genomic-tips 100/G (Qiagen Inc, CA, USA) according to the protocol except for the addition of 2.6 mg $mL^{-1}$ zymolyase (20T, Amsbio, UK) and with 4 mg $mL^{-1}$ lysozyme. Cells were subsequently lysed by application of 2.9 kbar pressure with a French Press (Constant Systems Ltd, UK). After purification with the genomic tip kit the starting amount of DNA was quantified with Qubit dsDNA HS assay kit (ThermoFisher Scientific) and the quality of DNA was assessed with NanoDrop 2000. Sequencing was performed in-house on an Illumina MiSeq Sequencer (Illumina, San Diego, CA, USA), with MiSeq Reagent Kit v3 with 2 x 300 bp read length and the DNA libraries were prepared using the Nextera XT

DNA sample preparation kit (Illumina). Quality-trimming, adapter removal and contaminant-filtering of paired-end sequencing reads was performed using BBDUK (BBTOOLS version 37.17)[1]. Trimmed reads for all samples were co-assembled using metaSPAdes v3.10.1 (74) at default settings. MetaSPAdes iteratively assembled the metagenome using kmer size 21, 33, 55, 77, 99 and 127. Reads were mapped back to the metagenome for each sample separately using Burrows-Wheeler Aligner 0.7.15 (BWA), employing the "mem" algorithm (75). The generated sequence mapping files were handled and converted as needed using SAMtools 2.1 (76). Metagenome binning was performed using five different binning algorithms: BinSanity v0.2.5.9 (32), COCACOLA (33), CONCOCT (34), MaxBin 2.0 2.2.3 (35) and MetaBAT 2 2.10.2 (31). The five bin sets were supplied to DAS Tool 1.0 (77) for consensus binning to obtain the final optimized bins. The quality of the generated bins was assessed using CheckM 1.0.7 (36). The bins with the highest coverage were annotated by the RAST server for further analysis (37). For the identification of the dominant individual bins the complete 16S-rRNA genes were submitted to the SILVA database, release 132 (41). Genome sequence data of the metagenomic assembly of the galacturonate fermenting enrichment culture metagenome, biosample SAMN08337232, has been deposited at the NCBI genomic database (www.ncbi.nlm.nih.gov/genome/) with the corresponding BioProject ID PRJNA429181.

[1] BBMap - Bushnell B. - sourceforge.net/projects/bbmap/ (unpublished)

# Acknowledgements

We thank Marcel van den Broek for assisting with the bioinformatics, Dimitry Sorokin for assistance with the isolation, Ben Abbas for help with the 16S-rRNA gene amplicon sequencing, Eveline van den Berg for critical reading of the manuscript, Maarten Verhoeven, Ioannis Papapetridis and Jasmine Bracher for fruitful discussions and Orgaworld Nederland B.V. and the butcher in Est for supplying the inoculum. This research was supported by the SIAM Gravitation Grant 024.002.002, the Netherlands Organization for Scientific Research.



## Author contribution

LV, ML, AM and JP designed the experimental setup. LV, ML and JP wrote the manuscript. MH assisted with testing the experimental setup of the bioreactors. LV did all the wet lab experiments, analysed the 16S-rRNA gene amplicon sequencing, extracted the DNA for the metagenomic analysis, annotated and analysed the metagenomic data. PC conducted the library preparation and sequencing of the enrichment cultures and JF performed the bioinformatic analysis of the metagenomic dataset. All have read and approved the final manuscript.

# Supplementary Material



**Table S6 |** The carbon and electron balance of the duplicate enrichment cultures with their corresponding recoveries. The galacturonate was consumed, which is indicated by a minus sign, at pH 8, a temperature of 30 °C and a dilution rate of 0.13 $h^{-1}$. The standard deviation was derived from the duplicate reactors, which each was derived from five sequential measurements.

| Compound | Carbon balance (mCmol $h^{-1}$) | Electron balance (mmol electrons $h^{-1}$) |
|---|---|---|
| **Galacturonate** | -8.5 ± 0.0 | -28.3 ± 0.0 |
| **Acetate** | 4.5 ± 0.1 | 17.9 ± 0.8 |
| **Formate** | 1.3 ± 0.1 | 2.7 ± 0.1 |
| **Succinate** | 0.2 ± 0.1 | 0.1 ± 0.2 |
| **Ethanol** | 0.1 ± 0.1 | 0.2 ± 0.4 |
| **$CO_2$** | 0.5 ± 0.0 | - |
| **$H_2$** | - | 0.1 ± 0.0 |
| **Biomass** | 1.4 ± 0.2 | 5.8 ± 0.3 |
| ***Recovery (%)*** | *92 ± 2* | *95 ± 2* |

## Calculations S2 | Determination of the attribution of the Wood-Ljungdahl pathway to the overall acetate production rate.

The overall substrates and products of the adapted Entner-Doudoroff pathway is shown in equation S2.a, with formate being shown as $H_2$ and $CO_2$ to simplify the calculations. The overall substrates and products of the Wood-Ljungdahl pathway are shown in equation S2.b.

(S2.a) - 1 Galacturonate -1 $H_2O$ + 2 $H_2$ + 2 $CO_2$ + 2 Acetate
(S2.b) - 4 $H_2$ – 2 $CO_2$ + 1 Acetate +2 $H_2O$

It was hypothesized that both pathways were active in "*Candidatus* Galacturonibacter soehngenii" and as such the total acetate production rate in mmol $h^{-1}$ could be described as a function of S2.a and S2.b. The overall galacturonate consumption rate, as well as the $H_2$, $CO_2$ production rate were known, as shown Table S1. So, the acetate production rate per pathway could be calculated as there are two equations and one unknown. The catabolic galacturonate consumption rate could be calculated by subtracting the amount of moles galacturonate going towards biosynthesis, which resulted in 6.7 ± 0.2 mCmol $h^{-1}$ galacturonate used for catabolism, equivalent to -1.1 ± 0.0 mmol $h^{-1}$ galacturonate. Galacturonate was solely consumed in reaction S2.a, making the galacturonate catabolic consumption rate the rate of this pathway. The overall hydrogen production rate which was measured, was derived from the production rate of equation S2.a and the consumption rate of S2.b, so with the rate of S2.a and the overall hydrogen production rate (1.5 ± 0.0 mmol $h^{-1}$) known, the rate of S2.b was determined to be 0.2 ± 0.0 mmol $h^{-1}$. Based on these calculations the resulting average specific rates of acetate formation from the adapted Entner-Doudoroff pathway was determined as

6.9 ± 0.4 mmol (g biomass)$^{-1}$ h$^{-1}$ and the average specific rates of acetate formation from the Wood-Ljungdahl pathway was determined as 1.7 ± 0.2 mmol (g biomass)$^{-1}$ h$^{-1}$.

**Table S2 |** Results of the isolation attempts with the main substrate, mode of cultivation and additives specified. The basis of the media was the salts media of (8) at pH 8 ± 0.1 with 2 mM substrate, vitamins, basic Se/W solution, 10 mM $NaHCO_3$ and 50 mg L$^{-1}$ yeast extract supplemented according to (78).

| Substrate | Mode of cultivation | Additives |
|---|---|---|
| **Galacturonate, xylose** | Liquid serial dilutions | 0.2 g L$^{-1}$ Cysteine or 0.2 mM $Na_2S$ |
| **Galacturonate, xylose** | Liquid serial dilutions | Cysteine or $Na_2S$ + 20 – 100 mg L$^{-1}$ kanamycin or streptomycin |
| **Galacturonate and $H_2$** | Liquid serial dilutions | 50 mg L$^{-1}$ kanamycin or streptomycin and SPLK[1] or RF[2] |
| **$H_2$ and $NaHCO_3$** | Liquid serial dilutions | Cysteine |
| **Galacturonate and $H_2$** | Liquid serial dilutions | Sulfide and cell-sorting[3] |
| **Pyruvate or ethanol or lactate** | Liquid serial dilutions | Sulfide |
| **Galacturonate** | Plating by agar-shake | Sulfide |
| **Galacturonate and $H_2$** | Plating by agar-shake | Sulfide |

1 | filter-sterilized supernatant from the reactor or *Klebsiella* cultures, 1 %(v/v)
2 | Rumen fluid, filter sterilized 0.1 %(v/v).
3 | several centrifugation steps in 2 mL sterile Eppendorf tubes at 3.7 x g (Heraeus Biofuge pico, Thermo Scientific) for 1-5 min, with recovering supernatant.

Extensive attempts in both liquid dilution series as well as agar plate isolations have been conducted, by the use of a large range of substrates, the addition of vitamins, yeast extract, rumen fluid and/or spend reactor broth from the chemostat enrichment culture. Additionally, kanamycin and/or streptomycin and cell sorting were used to supress the growth of *Klebsiella oxytoca,* see Table S3. All this did not result in a pure culture of "*Candidatus* G. soehngenii", based on microscope observations and sequencing results (*data not shown*).

## Calculations S3 | Determination of the dissolved hydrogen partial pressure ($pH_{2,\ liquid}$) in the bioreactors in chemostat operations and the corresponding Gibbs free-energy change of the Wood-Ljungdahl pathway under operational conditions.

The dissolved hydrogen partial pressure in the liquid can be calculated, with the use of the standard mass transfer theory and the measured hydrogen partial pressure in the off - gas. The mass transfer of hydrogen can be depicted as shown in equation S3.1, assuming a pseudo – steady state between the hydrogen production rate and the transfer rate.

$$(S3.1)\ H_2TR = k_L a_{H_2} * {H_{H_2}}^{-1} * (p_{H_2,liquid} - p_{H_2,gas})$$

In this equation the $H_2TR$ is the transfer rate of hydrogen (in mol m$^{-3}$ h$^{-1}$), $k_La_{H2}$ is the hydrogen mass transfer coefficient (in h$^{-1}$), $H_{H2}$ is the Henry coefficient of hydrogen (1.27

atm $dm^{-3}$ $mol^{-1}$, (79)) and $p_{H2,\ liquid}$ and $p_{H2,\ gas}$ the hydrogen partial pressure in the liquid and the off-gas respectively (in atm). The $k_La_{H2}$ used to approximate the hydrogen partial pressure in the liquid phase was derived from earlier $k_La_{H2}$ measurements conducted in similar bioreactors under similar operational conditions (66). The values used and the calculated $pH_{2,gas}$ and resulting $pH_{2,liquid}$ value are shown in Table S3.

**Table S3 |** The values used to approximate the hydrogen partial pressure in the liquid in the bioreactors with pH 8, a temperature of 30 °C and a dilution rate of 0.13 $h^{-1}$. The standard deviation was derived from the duplicate reactors, which each was derived from five sequential measurements

| **Hydrogen transfer rate (mmol $dm^{-3}$ $h^{-1}$)** | **$k_La_{H2}$( $h^{-1}$)*** | **Henry coefficient (atm $dm^{-3}$ $mol^{-1}$)** | **$pH_{2,gas}$ (atm)** | **$pH_{2,\ liquid}$ (atm)** |
|---|---|---|---|---|
| **0.13 ± 0.0** | 26.3 | 1.27 | 1.41 ± 0.0 · $10^{-5}$ | 3.92 ± 1.5 · $10^{-3}$ |

* derived from de Kok et. al. (66)

The approximated hydrogen partial pressure in the liquid was subsequently used, together with the concentration of $CO_2$ (taking into account the hydration reactions of $CO_2$), 1.8 ± 0.1 mM, and the concentration of acetate, 35 ± 1 mM, in the liquid with a temperature of 30 ˚C, to determine the Gibbs free-energy change of the acetogenesis reaction shown in equation 2 in the results section. The equation used is shown in equation S3.2. The standard Gibbs free-energy change of the reaction used was $\Delta G'^{\circ}$ = -99 ± 24 kJ $mol^{-1}$ (80, 81).

$$(S3.2)\ \Delta_r G' = \Delta_r G'^{o} + RT \ln\left(\frac{products}{substrates}\right)$$

The resulting Gibbs free energy change was -17.3 ± 24 kJ $mol^{-1}$.

# References


1. Kleerebezem R, van Loosdrecht MCM. 2007. Mixed culture biotechnology for bioenergy production. Curr Opin Biotechnol 18:207–212.
2. Angenent LT, Karim K, Al-dahhan MH, Wrenn BA, Domiguez-Espinosa R. 2004. Production of bioenergy and biochemicals from industrial and agricultural wastewater. Trends Biotechnol 22:477–485.
3. Claassen PAM, van Lier JB, Lopez Contreras AM, van Niel EWJ, Sijtsma L, Stams AJM, de Vries SS, Weusthuis RA. 1999. Utilisation of biomass for the supply of energy carriers. Appl Microbiol Biotechnol 52:741–755.
4. Bijerinck MW. 1901. Enrichment culture studies with urea bacteria. Centalblatt f Bacteriol Part II:33–61.
5. Rodríguez J, Kleerebezem R, Lema JM, van Loosdrecht MCM. 2006. Modeling product formation in anaerobic mixed culture fermentations. Biotechnol Bioeng 93:592–606.
6. Zoetemeyer RJ, Matthijsen AJCM, Cohen A, Boelhouwer C. 1982. Product inhibition in the acid forming stage of the anaerobic digestion process. Water Res 16:633–639.
7. Lengeler JW, Drews G, Schlegel HG. 1998. Biology of Prokaryotes. Blackwell, Oxford.
8. Temudo MF, Mato T, Kleerebezem R, van Loosdrecht MCM. 2009. Xylose anaerobic conversion by open-mixed cultures. Appl Microbiol Biotechnol 82:231–239.
9. Lin CY, Cheng CH. 2006. Fermentative hydrogen production from xylose using anaerobic mixed microflora. Int J Hydrogen Energy 31:832–840.
10. Zhang F, Yang JH, Dai K, Chen Y, Li QR, Gao FM, Zeng RJ. 2016. Characterization of microbial compositions in a thermophilic chemostat of mixed culture fermentation. Appl Microbiol Biotechnol 100:1511–1521.
11. Grohmann K, Manthey JA, Cameron RG, Buslig BS. 1998. Fermentation of galacturonic acid and pectin-rich materials to ethanol by genetically modified strains of *Erwinia*. Biotechnol Lett 20:195–200.
12. Richard P, Hilditch S. 2009. D-Galacturonic acid catabolism in microorganisms and its biotechnological relevance. Appl Microbiol Biotechnol 82:597–604.
13. Mohnen D. 2008. Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277.
14. Grohmann K, Bothast RJ. 1994. Pectin-rich residues generated by processing of citrus fruits, apples, and sugar beets - enzymatic hydrolysis and biological conversion to value-added products. ACS Symp Ser Enzym Convers Biomass Fuels Prod 566:372–390.
15. Bayer AG. Why sugar beet matters to us. https://www.cropscience.bayer.com/en/crop-compendium/key-crops/sugar-beets
16. United States Department of Agriculture. 2017. Citrus: world markets and trade. https://apps.fas.usda.gov/psdonline/circulars/citrus.pdf
17. Food and Agriculture Organization of the United Nations. 2016. Crops, Production quantities of sugar beet, citrus fruits and apples in the World. FAOSTAT. http://www.fao.org/faostat/en/#data/QC
18. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH,

Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT. 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 90:391–418.

19. Doran JB, Cripe J, Sutton M, Foster B. 2000. Fermentations of pectin-rich biomass with recombinant bacteria to produce fuel ethanol. Appl Biochem Biotechnol 84–86:141–153.
20. Doran-Peterson J, Cook DM, Brandon SK. 2008. Microbial conversion of sugars from plant biomass to lactic acid or ethanol. Plant J 54:582–592.
21. Ashwell G, Wahba AJ, Hickman J. 1960. Uronic acid metabolism in bacteria I purification and properties of uronic acid isomerase in Escherichia coli. J Biol Chem 235:1559–1565.
22. Hickman J, Ashwell G. 1960. Uronic acid metabolism in bacteria II Purification and properties of D-altronic acid and D-mannonic acid dehydrogenases in *Escherichia coli*. J Biol Chem 235:1566–1570.
23. Smiley JD, Ashwell G. 1960. Uronic acid metabolism in bacteria III purification and properties of D-altronic acid and D-mannonic acid dehydrases in *Escherichia coli*. J Biol Chem 235:1571–1575.
24. Cynkin MA, Ashwell G. 1960. Uronic acid metabolism in bacteria IV purification and properties of 2-keto-3-deoxy-D-gluconokinase in *Escherichia coli*. J Biol Chem 235:1576–1579.
25. Kovachevich R, Wood WA. 1955. Carbohydrate metabolism by *Pseudomonas fluorescens* IV purification and properties of 2-keto-3-deoxy-6-phospho-gluconate aldolase. J Biol Chem 213:757–767.
26. González-Cabaleiro R, Lema JM, Rodríguez J. 2015. Metabolic energy-based modelling explains product yielding in anaerobic mixed culture fermentations. PLoS One 10:1–17.
27. Grohmann K, Baldwin EA, Buslig BS. 1994. Fermentation of galacturonic acid and other sugars in orange peel hydrolysates by the ethanologenic strain of *Escherichia coli*. Biotechnol Lett 16:281–286.
28. Noorman HJ, Heijnen JJ, Luyben KCAM. 1991. Linear relations in microbial reaction systems : a general overview of their origin, form, and use. Biotechnol Bioeng 38:603–618.
29. Brooks JP, Edwards DJ, Harwich Jr MD, Rivera MC, Fettweis JM, Serrano MG, Reris RA, Sheth NU, Huang B, Girerd P, Vaginal Microbiome Consortium (additional members), Strauss JF, Jefferson KK, Buck GA. 2015. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. BMC Microbiol 15:66.
30. Kembel SW, Wu M, Eisen JA, Green JL. 2012. Incorporating 16S gene copy number information improves estimates of microbial diversity and abundance. PLoS Comput Biol 8:16–18.
31. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165.
32. Graham ED, Heidelberg JF, Tully BJ. 2017. BinSanity: unsupervised clustering of environmental microbial assemblies using coverage and affinity propagation. PeerJ 5:e3035.
33. Lu YY, Chen T, Fuhrman JA, Sun F, Sahinalp C. 2017. COCACOLA: Binning

metagenomic contigs using sequence COmposition, read CoverAge, CO-alignment and paired-end read LinkAge. Bioinformatics 33:791–798.
34. Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. 2014. Binning metagenomic contigs by coverage and composition. Nat Methods 11:1144–1146.
35. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics 32:605–607.
36. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 1043–1055.
37. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75.
38. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:206–214.
39. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365.
40. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol 12:635–645.
41. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829.
42. Ragsdale SW. 2008. Enzymology of the Wood-Ljungdahl pathway of acetogenesis, p. 129–136. *In* Annals of the New York Academy of Sciences.
43. Henderson G, Naylor GE, Leahy SC, Janssen PH. 2010. Presence of novel, potentially homoacetogenic bacteria in the rumen as determined by analysis of formyltetrahydrofolate synthetase sequences from ruminants. Appl Environ Microbiol 76:2058–2066.
44. Gagen EJ, Denman SE, Padmanabha J, Zadbuke S, Jassim R Al, Morrison M, Mcsweeney CS. 2010. Functional gene analysis suggests different acetogen populations in the bovine rumen and tammar wallaby forestomach. Appl Environ Microbiol 76:7785–7795.
45. Breznak JA. 2006. The Genus *Sporomusa*, p. 991–1001. *In* Dworkin, M, Falkow, S, Rosenberg, E, Schleifer, KH, Stackebrandt, E (eds.), The Prokaryotes Volume 4: Bacteria: Firmicutes, Cyanobacteria, 3rd ed. Springer-Verlag New York, New York.
46. Shin SH, Kim S, Kim JY, Lee S, Um Y, Oh MK, Kim YR, Lee J, Yang KS. 2012. Complete genome sequence of *Klebsiella oxytoca* KCTC 1686, used in production of 2,3-butanediol. J Bacteriol 194:2371–2372.
47. Ng H, Vaughn RH. 1963. *Clostridium rubrum* sp. n. and other pectinolytic *Clostridia*

from soil. J Bacteriol 85:1104–1113.
48. Burkey LA. 1928. The fermentation of cornstalks and their constituents : I , studies on the pectin-fermenting bacteria. Retrospective Theses and Dissertations. 14219. Iowa State College.
49. Lanigan GW. 1959. Studies on the pectinolytic anaerobes *Clostridium flavum* and *Clostridium laniganii.* J Bacteriol 77:1–9.
50. Thauer RK, Jungermann K, Decker K. 1977. Energy conservation in chemotrophic anaerobic bacteria. Bacteriol Rev 41:100–180.
51. Thauer RK, Kirchniawy FH, Jungermann KA. 1972. Properties and function of the pyruvate formate-lyase reaction in Clostridiae. Eur J Biochem 27:282–290.
52. Valentine DL, Sessions AL, Tyler SC, Chidthaisong A. 2004. Hydrogen isotope fractionation during $H_2/CO_2$ acetogenesis: hydrogen utilization efficiency and the origin of lipid-bound hydrogen. Geobiology 2:179–188.
53. Tremblay P, Höglund D, Koza A, Bonde I, Zhang T. 2015. Adaptation of the autotrophic acetogen *Sporomusa ovata* to methanol accelerates the conversion of $CO_2$ to organic products. Sci Rep 161–168.
54. Franks AH, Harmsen HJM, Gerwin C, Jansen GJ, Schut F, Gjalt W. 1998. Variations of bacterial populations in human feces measured by fluorescent *in situ* hybridization with group-specific 16S rRNA-targeted oligonucleotide probes. Appl Environ Microbiol 64:3336–3345.
55. Moter A, Göbel UB. 2000. Fluorescence *in situ* hybridization (FISH) for direct visualization of microorganisms. J Microbiol Methods 41:85–112.
56. Meehan CJ, Beiko RG. 2014. A phylogenomic view of ecological specialization in the *Lachnospiraceae*, a family of digestive tract-associated bacteria. Genome Biol Evol 6:703–713.
57. Andrews JH, Harris RF. 1986. r- and K-Selection and microbial ecology, p. 99–147. *In* Marshall, KC (ed.), Advances in Microbial Ecology, 9thed. Springer, Bosten, MA.
58. Kuenen JG. 2015. Continuous Cultures (Chemostats). Ref Modul Biomed Res 1–19.
59. Lendenmann U, Snozzi M, Egli T. 1996. Kinetics of the simultaneous utilization of sugar mixtures by *Escherichia coli* in continuous culture. Appl Environ Microbiol 62:1493–1499.
60. Egli T, Lendenmann U, Snozzi M. 1993. Kinetics of microbial growth with mixtures of carbon sources. Antonie Van Leeuwenhoek 63:289–298.
61. Schuchmann K, Müller V. 2014. Autotrophy at the thermodynamic limit of life: A model for energy conservation in acetogenic bacteria. Nat Rev Microbiol 12:809–821.
62. Monod J. 1942. Recherches sur la croissance des cultures bacteriennes. Hermann et Cie, Paris, France.
63. Savage MD, Drake HL. 1986. Adaptation of the acetogen *Clostridium thermoautotrophicum* to minimal medium. J Bacteriol 165:315–318.
64. Daniel SL, Hsu T, Dean SI, Drake HL, Ljungdahl L. 1990. Characterization of the $H_2$ *Clostridium thermoaceticum* and *Acetogenium kivuit*. J Bacteriol 172:4464–4471.
65. Lundie jr. LL, Drake HL. 1984. Development of a minimally defined medium for the acetogen *Clostridium thermoaceticum*. J Bacteriol 159:700–703.
66. De Kok S, Meijer J, van Loosdrecht MCM, Kleerebezem R. 2013. Impact of dissolved hydrogen partial pressure on mixed culture fermentations. Appl Microbiol Biotechnol 97:2617–2625.
67. Heijnen JJ. 1999. Bioenergetics of Microbial Growth, p. 267–291. *In* Flickinger, MC,

Drew, SW (eds.), Encyclopedia of Bioprocess Technology: Fermentation. Biocatalysis, and Bioseparation. John Wiley & Sons.
68. Ragsdale SW, Pierce E. 2008. Acetogenesis and the Wood – Ljungdahl pathway of $CO_2$ fixation. Biochim Biophys Acta - Bioenerg 1784:1873–1898.
69. Temudo MF, Muyzer G, Kleerebezem R, van Loosdrecht MCM. 2008. Diversity of microbial communities in open mixed culture fermentations: Impact of the pH and carbon source. Appl Microbiol Biotechnol 80:1121–1130.
70. Roels J. 1983. Energetics and kinetics in biotechnology. Biomedical Press, Amsterdam.
71. van den Berg EM, Elisário MP, Kuenen JG, Kleerebezem R, van Loosdrecht MCM. 2017. Fermentative bacteria influence the competition between denitrifiers and DNRA bacteria. Front Microbiol 8:1–13.
72. Wang Y, Qian PY. 2009. Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. PLoS One 4:e7401.
73. Johnson K, Jiang Y, Kleerebezem R, Muyzer G, van Loosdrecht MCM. 2009. Enrichment of a mixed bacterial culture with a high polyhydroxyalkanoate storage capacity. Biomacromolecules 10:670–676.
74. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. MetaSPAdes: A new versatile metagenomic assembler. Genome Res 27:824–834.
75. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595.
76. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.
77. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, Banfield JF. 2017. Recovery of genomes from metagenomes via a dereplication, aggregation, and scoring strategy. bioRxiv 107789.
78. Sorokin DY, Abbas B, Geleijnse M, Pimenov NV, Sukhacheva MV, van Loosdrecht MCM. 2015. Methanogenesis at extremely haloalkaline conditions in the soda lakes of Kulunda Steppe (Altai, Russia). FEMS Microbiol Ecol 91:1–12.
79. Janssen L, Warmoeskerken M. 2006. Transport phenomena data companion, 3rd ed. VSSD, Delft.
80. Flamholz A, Noor E, Bar-Even A, Milo R. 2012. EQuilibrator - The biochemical thermodynamics calculator. Nucleic Acids Res 40:770–775.
81. Noor E, Haraldsdottir HS, Milo R, Fleming RMT. 2013. Consistent estimation of Gibbs energy using component contributions. PLoS Comput Biol 9:e1003098.
82. Kempf VAJ, Trebesius K, Autenrieth IB. 2000. Fluorescent *in situ* hybridization allows rapid identification of microorganisms in blood cultures. J Clin Microbiol 38:830–838.
83. Kong Y, He M, McAlister T, Seviour R, Forster R. 2010. Quantitative fluorescence *in situ* hybridization of microbial communities in the rumens of cattle fed different diets. Appl Environ Microbiol 76:6933–6938.
84. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 56:1919–1925.

85. Daims H, Brühl A, Amann R, Schleifer K-H, Wagner M. 1999. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol 22:434–444.

# Chapter 3 | "*Candidatus* Galacturonibacter soehngenii" shows an acetogenic catabolism of galacturonic acid but lacks a canonical carbon monoxide dehydrogenase/acetyl-CoA synthase complex

Laura C. Valk, Martijn Diender, Gerben R. Stouten, Jette F. Petersen, Per H. Nielsen, Morten S. Dueholm, Jack T. Pronk and Mark C.M. van Loosdrecht

Keywords: acetogenesis; $^{13}$C-labelling; meta-transcriptomics; chemostat enrichment culture; Wood-Ljungdahl pathway

This chapter was submitted to: *Frontiers in Microbiology, Special: Acetogens - From The Origin Of Life To Biotechnological Applications.*

## Abstract

Acetogens have the ability to fixate carbon during fermentation by employing the Wood-Ljungdahl pathway, which is highly conserved across Bacteria and Archaea. In a previous study, product stoichometries in galacturonate-limited, anaerobic enrichment cultures of "*Candidatus* Galacturonibacter soehngenii", from a novel genus within the *Lachnospiraceae*, suggested the simultaneous operation of a modified Entner-Doudoroff pathway for galacturonate fermentation and a Wood-Ljungdahl pathway (WLP) for acetogenesis. However, a draft metagenome-assembled genome (MAG) based on short reads did not reveal homologs of genes encoding a canonical WLP carbon-monoxide-dehydrogenase/acetyl-Coenzyme A synthase (CODH/ACS) complex. In this study, $NaH^{13}CO_3$ fed to chemostat-grown, galacturonate-limited enrichment cultures of "*Ca*. G. soehngenii" was shown to be incorporated into acetate. Preferential labelling of the carboxyl group of acetate was consistent with acetogenesis via a WLP in which the methyl group of acetate was predominately derived from formate. This interpretation was further supported by high transcript levels of a putative pyruvate-formate lyase gene and very low transcript levels of a candidate gene for formate dehydrogenase. Reassembly of the "*Ca*. G. soehngenii" MAG with support from long-read nanopore sequencing data produced a complete MAG, which confirmed the absence of canonical CODH/ACS-complex genes homologs. However, high CO-dehydrogenase activities were measured in cell extracts of "*Ca*. G. soehngenii" enrichment cultures, contradicting the absence of corresponding homologues in the MAG. Based on the highly conserved amino-acid motif associated with anaerobic Ni-CO dehydrogenase proteins, a novel candidate was identified which could be responsible for the observed activities. These results demonstrate operation of an acetogenic pathway, most probably as a yet unresolved variant of the Wood-Ljungdahl pathway, in anaerobic, galacturonate-limited cultures of "*Ca*. G. soehngenii".

# Introduction



Over the course of multiple decades, seven carbon-fixing pathways capable of supporting autotrophic growth have been identified and intensively studied; the Calvin-Benson-Bassham (CCB) or reductive pentose-phosphate cycle (rPPP), the reductive citric-acid cycle (Arnon-Buchanan (AB) cycle), the hydroxypropionate (Fuchs-Holo) bi-cycle, the 3-hydroxypropionate/4-hydroxybutyrate cycle, dicarboxylate/hydroxybutyrate cycle, the reductive acetyl-CoA (Wood-Ljungdahl) pathway and the reductive glycine pathway (1–3). The first five pathways are primarily used for carbon fixation and the reductive glycine pathway for recycling of electron carriers. Only the Wood-Ljungdahl pathway (WLP) also acts as a primary pathway for energy conservation in anaerobes (2, 4, 5)

The WLP is highly conserved across Archaea and Bacteria, with only two known variations, one found predominantly in methanogenic archaea and one in acetogenic bacteria. The first has formyl-methanofuran rather than formate as first intermediate, and uses ATP-independent formyl-MFR:tetrahydromethanopterin formyltransferase instead of ATP-consuming formyl-tetrahydrofolate ligase (consuming an ATP). Moreover, methanogens use methanofuran (MFR), tetrahydromethanopterin and coenzyme-$F_{420}$ as cofactors while acetogens rely on NAD(P)H, tetrahydrofolate (THF) and ferredoxin (Fd) (2, 6). Reduction of $CO_2$ to acetate via the WLP requires 8 electrons (equation 1, (4, 7)).

(1) $2\ CO_2 + 8\ H^+ + 8\ e^- + n\ ADP + n\ P_i \rightarrow CH_3COOH + n\ ATP + (2+n)\ H_2O$

The WLP consists of two branches. In acetogens, the WLP methyl branch reduces $CO_2$ to a methyl group by first reducing $CO_2$ to formate via formate dehydrogenase (*fdhA*; EC 1.17.1.9), after which formate is bound to tetrahydrofolate (THF) by formate-tetrahydrofolate ligase (*fhs*, EC 6.3.4.3). Formyl-THF is then further reduced to methenyl-THF, methylene-THF and lastly to methyl-THF by formyl-THF cyclohydrolase and methylene-THF dehydrogenase (*folD*; EC 3.5.4.9 and EC 1.5.1.5) and methylene-THF reductase (*metF*, EC 1.5.1.20), respectively (7, 8). A methyl transferase then transfers the methyl group from THF to a corrinoid iron–sulfur protein (*acsE*, EC 2.3.1.258), which is a subunit of the carbon monoxide (CO) dehydrogenase/acetyl-CoA synthase complex. The carbonyl branch of the WLP reduces $CO_2$ to CO in a reaction catalyzed by another subunit of the canonical WLP, the CO dehydrogenase/acetyl-CoA synthase complex (CODH/ACS, EC 2.3.1.169). Alternatively, CO can be formed by a separate CO dehydrogenase (CODH, EC 1.2.7.4) (9–11). The CODH/ACS complex then links the two WLP branches by coupling the CO- and $CH_3$-groups with CoA, yielding acetyl-CoA (8, 9, 12). The high degree of conservation of WLP genes and their genomic co-localization suggests that their evolution involved interspecies gene transfer events (6, 13). However, two recent studies suggested carbon fixation occurred in the absence of a full complement of structural genes for canonical WLP enzymes (3, 14). These observations suggest that variants of the canonical WLP may still await discovery.

In a recent study on D-galacturonate-limited, anaerobic enrichment cultures, we identified the dominant bacterium as a species from a novel genus within the *Lachnospiraceae*, for which we proposed the name "*Candidatus* Galacturonibacter soehngenii". The *Lachnospiraceae* family is part of the phylum Firmicutes, which includes several genera that harbor acetogens (7, 14–16). Fermentation product stoichiometries of the enrichment cultures were consistent with an acetogenic dissimilation of galacturonate. The overall stoichiometry is shown in equation 2 (14).



(2) $1\ C_6H_{10}O_7 \rightarrow 2.5\ C_2H_4O_2 + 1\ CO_2$

Metagenome analysis of the enrichment culture revealed homologs of most structural genes for WLP enzymes, but no homologs were found for genes encoding subunits of the canonical CODH/ACS complex (EC 2.3.1.169) (14).

The goal of the present study was to further investigate the presence of a possible alternative configuration of the WLP in "*Ca.* G. soehngenii". To analyze *in vivo* activity of the WLP, D-galacturonate-limited enrichment cultures were co-fed with $^{13}C$-labeled bicarbonate, followed by analysis of $^{13}C$ in the methyl and carboxyl groups of acetate. To investigate whether canonical WLP genes might have been overlooked in the initial metagenomics analysis, a fully closed metagenome-assembled genome (MAG) sequence of "*Ca.* G. soehngenii" was constructed using long-read nanopore sequencing, and meta-transcriptome analysis was performed to analyze the expression levels of genes of interest. Additionally, CO dehydrogenase activity was analyzed in cell extracts.

## Results

### Physiological characterization of D-galacturonate-limited enrichment cultures dominated by "*Ca.* G. soehngenii".

Anaerobic, galacturonate-limited chemostat enrichment cultures were used to study the physiology of "*Ca.* G. soehngenii" cultures. In a previous study (14), the relative abundance of "*Ca*. G. soehngenii" in such cultures did not exceed 65 %, based on metagenomic analysis, and formate and $H_2$ were detected in the liquid and gas phases, respectively. It was hypothesized that, in these experiments, a low *in situ* hydrogen partial pressure limited *in vivo* WLP activity, as it was expected that hydrogen was used as reductant for the production of acetate from formate or $CO_2$. To investigate this possibility, head space flushing instead of sparging was applied, using $N_2$ gas. This caused an increase in the hydrogen partial pressure in the media broth (32). Additionally, the dilution rate was decreased from 0.125 $h^{-1}$ to 0.1 $h^{-1}$. Analysis of the abundance of "*Ca*. G. soehngenii" in the resulting enrichment cultures by quantitative fluorescence *in situ* hybridization (qFISH) indicated that 86.5 ± 2.6 % of the bio-volume of qFISH-detectable cells consisted of "*Ca.* G. soehngenii". The major side population *Enterobacteriaceae* represented 13.8 ± 2.4 % of the bio-volume. As these two subpopulations together accounted for 100.2 ± 5.0 % of the

bio-volume, it was assumed that any other, minor, subpopulations did not significantly influence the stoichiometry of catabolic fluxes.

**Table 1 |** Yields (in Cmol ($Cmol_{galacturonate}$)$^{-1}$, unless stated otherwise) and biomass- specific conversion rates (q; mmol $g_{biomass}^{-1}$ $h^{-1}$) of anaerobic, galacturonate-limited chemostat enrichment cultures dominated by "*Ca*. Galacturonibacter soehngenii". Chemostat cultures were operated at dilution rate of 0.1 $h^{-1}$, pH 8 and at 30 °C, with D-galacturonate the sole carbon source. Data are presented as average ± mean deviations, derived from nine measurements each on duplicate steady-state enrichment cultures.



| | Yield ($Cmol_i$ $Cmol_s^{-1}$) | q (mmol $(g_x)^{-1}$ $h^{-1}$) |
|---|---|---|
| **D-Galacturonate** | - | -4.0 ± 0.2 |
| **Biomass ($g_x$ $g_s^{-1}$)** | 0.17 ± 0.01 | - |
| **Acetate** | 0.57 ± 0.03 | 6.9 ± 0.4 |
| **Formate** | 0.02 ± 0.01 | 0.4 ± 0.2 |
| **$CO_2$** | 0.18 ± 0.02 | 4.3 ± 0.3 |
| **$H_2$ (mol $Cmol^{-1}$)** | 0.02 ± 0.01 | 0.2 ± 0.1 |
| **$H_2$ + Formate (mol $Cmol_s^{-1}$)** | 0.04 ± 0.02 | |
| **Acetyl-CoA derivatives (mol $Cmol_s^{-1}$)** | 0.29 ± 0.02 | |

Product yields and biomass-specific conversion rates of the D-galacturonate-limited anaerobic enrichment cultures dominated by "*Ca*. G. soehngenii" (Table 1) showed acetate as dominant catabolic product (0.57 ± 0.03 Cmol ($Cmol_{galacturonate}^{-1}$). Carbon and electron recoveries were 94 % and 92 %, respectively, indicating that all major fermentation products were identified. As observed previously (14), this acetate yield on galacturonic acid was significantly higher than the combined yields of formate and hydrogen. This difference was interpreted as indicative for acetogenesis by one of the dominant organisms, of which only the "*Ca*. G. soehngenii" MAG was shown to harbour homologs for most Wood-Ljungdahl pathway (WLP) structural genes (7, 14). Yields of hydrogen and formate on galacturonate (0.01 ± 0.01 mol $Cmol_{galacturonate}^{-1}$) and 0.02 ± 0.01 Cmol ($Cmol_{galacturonate}^{-1}$), respectively) were significantly lower than found in a previous study on "*Ca*. G. soehngenii" (14). This observation is consistent with a higher in vivo contribution of the WLP as a result of a higher hydrogen partial pressure and/or lower specific growth rate in the present study. Incorporation of $^{13}C$-labeled bicarbonate into acetate corroborates acetogenic fermentation.

## Incorporation of $^{13}C$-labeled bicarbonate into acetate corroborates acetogenic fermentation.

A simple model was constructed to predict formation of labeled acetate, using biomass-specific conversion rates measured in pseudo-steady state enrichment cultures as inputs (Supplementary Material Calculations 2 and Figure S1). Model simulations predicted that, after 8 h, approximately 15 % of the acetate produced by the enrichment culture should be labeled. To investigate if $CO_2$ was indeed incorporated into acetate via acetogenic

fermentation, $^{13}C$-labeled bicarbonate was fed to a "*Ca.* G. soehngenii" enrichment chemostat culture. However, after 8 h, the fraction of $^{13}C$ in the methyl group of acetate increased to 2.0 %. This increase represented only a small increase relative to the 1 % natural abundance of $^{13}C$ (Table 2; (33)). In contrast, after 8 h of $^{13}C$-bicarbonate feeding, the enrichment culture showed a 21.5 % abundance of $^{13}C$ in the carbonyl-group of acetate (Table 2).



**Table 2 |** Percentages of $^{13}C$-labeled methyl and carbonyl groups in total-culture acetate, calculated from proton and carbon NMR spectra. Samples were taken from the "*Ca.* G. soehngenii" chemostat enrichment cultures in bioreactor 2 after switching the alkali supply line from 1 M NaOH to 1 M $NaH^{13}CO_3$ (Time =0 h).

| | Time (h) | % $^{13}C$ |
|---|---|---|
| **Methyl ($CH_3$)** | 0 | 1.0 |
| | 4 | 1.6 |
| | 8 | 2.0 |
| **Carbonyl (CO)** | 8 | 21.8 |

## Significant activity of CO dehydrogenase in cell extracts of "*Ca.* G. soehngenii" enrichment cultures.

In the WLP, $^{13}C$-labeled $CO_2$ incorporation into the carbonyl-group of acetate involves activity of CO dehydrogenase (COOS, EC 1.2.7.4). To investigate the presence of this key enzyme in "*Ca*. G. soehngenii", an anaerobic enzyme activity assay was performed on cell extracts of enrichment cultures, using CO as electron donor and methyl viologen (MV) as electron acceptor (30). These assays revealed a CO dehydrogenase activity of 2.1 ± 0.6 μmol $min^{-1}$ (mg protein)$^{-1}$. Reduction of MV in the absence of either CO or cell extract was below detection limit (< 0.05 μmol $min^{-1}$ (mg protein)$^{-1}$).

## Identification of two putative novel CO dehydrogenase genes in a newly obtained single-scaffold MAG of "*Ca*. G. soehngenii".

Previous analysis of the "*Ca.* G. soehngenii" MAG (14) was based on an assembly made with short-read DNA sequencing data. To identify if putative CODH/ACS complex genes had been missed in this analysis due to incomplete assembly, long-read Oxford Nanopore sequencing (34, 35) was used to improve the previously assembled "*Ca.* G. soehngenii" MAG. The resulting genome assembly consisted of 8 contigs and was estimated to have a 98 % completeness and contained no genetic contamination with sequences from other organisms according to checkM (Table 3). As in the previous study, homologs were detected for most structural genes associated with the WLP (Table 4), but none of the annotated genes in the predicted proteome showed homology with known CODH/ACS genes (8, 14, 24). A search in the newly assembled "*Ca*. G. soehngenii" MAG sequence for homologs of signature genes of the six other known pathways for inorganic carbon

fixation did not point towards their involvement in carbon metabolism (Supplementary Table S2).

**Table 3 |** Statistics of the metagenome-assembled genome (MAG) of "*Ca.* Galacturonibacter soehngenii". Completeness and contamination were estimated with CheckM (68).

| | ***"Candidatus* Galacturonibacter soehngenii"** |
|---|---|
| **Genome size (Mbp)** | 4.1 |
| **Scaffolds** | 1 |
| **Contigs** | 8 |
| **Contigs N50** | 1033779 |
| **Max contig size** | 1514059 |
| **Completeness (%)** | 98 |
| **Contamination (%)** | 0 |
| **GC content (%)** | 34.4 |
| **Protein coding density (%)** | 89 |
| **CDS** | 3924 |
| **rRNA copies** | 5 |



CO dehydrogenases contain highly conserved amino-acid motifs (Pfam or protein-family domains) associated with their nickel-iron-sulfur clusters (10, 13, 29, 36–38). The newly assembled "*Ca*. G. soehngenii" MAG sequence did not reveal hits for the Pfam domain of the CO dehydrogenase α-subunit of the CODH/ACS complex (PF18537) (39). However, two open reading frames F7O84_RS02405 and F7O84_RS11645, harbored the PF03063 Pfam domain, which is associated with the hybrid cluster protein (HCP) and the catalytic center of the Ni-CODH family (40, 41). Although HCP has been associated with hydroxylamine reductase activity, its catalytic activity has not been experimentally confirmed and, moreover, sequence motifs in HCP showed high similarity with the functional domain of Ni-CODHs making it an interesting candidate genes for the CODH function of the WLP in "*Ca*. G. soehngenii" (40, 42–44). A closer inspection of the genetic context of both genes showed many flanking genes encoding hypothetical proteins in their close vicinity, but no genes previously associated with acetogenesis.

**Table 4 |** Genes of the Wood-Ljungdahl pathway from the predictive proteome of the MAG "*Ca*. G. soehngenii" with gene names, EC number, gene or homolog and E-value based on SwissProt alignment (BLASTP version 2.2.28+, MicroScope platform v3.13.2).

| Enzyme | EC | Gene name | E-value | Gene ID |
|---|---|---|---|---|
| **Formate dehydrogenase** | 1.17.1.9 | *fdhA* | 1 $e^{-60}$ | F7O84_ RS07405 |
| **Formate–tetrahydrofolate ligase** | 6.3.4.3 | *fhs* | 0.0 | F7O84_RS05385 |
| **Methenyl-tetrahydrofolate cyclohydrolase/methylene–tetrahydrofolate dehydrogenase** | 3.5.4.9 and 1.5.1.5 | *folD* | 5 $e^{-152}$ | F7O84_RS05380 |
| **Methyl–tetrahydrofolate reductase** | 1.5.1.20 | *metF* | 1 $e^{-87}$ | F7O84_RS08335 |
| **5-Methyl-tetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase** | 2.1.1.258 | *acsE* | 5 $e^{-37}$ | F7O84_RS02745 |
| **CO-Methylating acetyl-CoA synthase** | 2.3.1.169 | *acsBCD* | >10 | |
| **Carbon-monoxide dehydrogenase** | 1.2.7.4 | *cooS* | >10 | |

## Homologs of acetogenesis genes are transcribed in D-galacturonate-limited "*Ca*. G. soehngenii" enrichment cultures.

A meta-transcriptome analysis of the enrichment cultures showed significant transcript levels of most homologs of known WLP genes, which were approximately 10-fold lower than those of homologs of structural genes encoding Entner-Doudoroff-pathway enzymes involved in galacturonate catabolism (Table 5). A notable exception was the extremely low transcript level of a putative formate dehydrogenase gene (F7O84_RS07405; EC 1.17.1.9). A candidate gene for pyruvate-formate lyase (PFL, EC 6.2.1.3) was highly transcribed (F7O84_03160, Table 5). These observations suggested that formate generated by PFL, rather than $CO_2$, was the major substrate for the methyl branch of the WLP in "*Ca*. G. soehngenii".

**Table 5 |** Transcript levels of putative key genes of the adapted Entner-Doudoroff pathway for galacturonate metabolism and the Wood-Ljungdahl pathway for acetogenesis in meta-transcriptome samples of the "*Ca*. G. soehngenii" chemostat enrichment cultures expressed as reads per kilobase million (RPKM, average ± average deviation) based on technical triplicates of duplicate enrichment cultures. N.d.: not detected.

| Protein function | EC number | Gene ID | RPKM |
|---|---|---|---|
| ***Adapted Entner-Doudoroff pathway*** | | | |
| **Uronate isomerase** | 5.3.1.12 | F7O84_RS17360 | 5852 ± 2398 |
| **Tagaturonate reductase** | 1.1.1.58 | F7O84_RS17370 | 3067 ± 1236 |
| **Altronate dehydratase** | 4.2.1.7 | F7O84_RS17375 | 8426 ± 3296 |
| **2-Dehydro-3-deoxygluconokinase** | 2.7.1.45 | F7O84_RS17390 | 3863 ± 1343 |
| **2-Dehydro-3-deoxyphosphogluconate aldolase** | 4.1.2.14 | F7O84_RS17395 | 1752 ± 245 |
| ***Acetate production*** | | | |
| **Pyruvate:ferredoxin oxidoreductase** | 1.2.7.1 | F7O84_RS03200 | 4145 ± 278 |
| **Pyruvate formate lyase** | 6.2.1.3 | F7O84_RS03160 | 1893 ± 651 |
| **Phosphate acetyltransferase** | 2.3.1.8 | F7O84_RS05985 | 1500 ± 176 |
| **Acetate kinase** | 2.7.2.1 | F7O84_RS05980 | 1625 ± 200 |
| ***Wood-Ljungdahl pathway*** | | | |
| **Formate dehydrogenase** | 1.17.1.9 | F7O84_RS07405 | 14 ± 3 |
| **Formate–tetrahydrofolate ligase** | 6.3.4.3 | F7O84_RS05385 | 256 ± 58 |
| **Methenyl-tetrahydrofolate cyclohydrolase/methylene–tetrahydrofolate dehydrogenase** | 3.5.4.9 and 1.5.1.5 | F7O84_RS05385 | 236 ± 9 |
| **Methyl–tetrahydrofolate reductase** | 1.5.1.20 | F7O84_RS08335 | 126 ± 13 |
| **5-methyl-tetrahydrofolate:corrinoid/iron-sulfur protein methyltransferase** | 2.1.1.258 | F7O84_RS08335 | 144 ± 19 |
| **CO-methylating acetyl-CoA synthase** | 2.3.1.169 | | n.d. |
| **CO dehydrogenase** | 1.2.7.4 | | n.d. |
| **Prismane/CO dehydrogenase family** | 1.7.99.1 | F7O84_RS02405 | 40 ± 8 |
| **Prismane/CO dehydrogenase family** | 1.7.99.1 | F7O84_RS11645 | 315 ± 51 |

**Table 5 |** *continued*

| Protein function | EC number | Gene ID | RPKM |
|---|---|---|---|
| ***Energy-metabolism associated genes*** | | | |
| **Electron transport complex protein A** | 7.2.1.2 | F7O84_RS03295 | 58 ± 5 |
| **Electron transport complex protein B** | 7.2.1.2 | F7O84_RS03300 | 261 ± 40 |
| **Electron transport complex protein C** | 7.2.1.2 | F7O84_RS03275 | 329 ± 22 |
| **Electron transport complex protein DG** | 7.2.1.2 | F7O84_RS03290 | 101 ± 13 |
| **Electron transport complex protein E** | 7.2.1.2 | F7O84_RS03285 | 143 ± 9 |
| **Ferredoxin hydrogenase subunit A** | 1.12.7.2 | F7O84_RS09545 | 196 ± 100 |
| **Ferredoxin hydrogenase subunit B** | 1.12.7.2 | F7O84_RS09550 | 356 ± 32 |
| **Ferredoxin hydrogenase subunit C** | 1.12.7.2 | F7O84_RS04820 | 124 ± 86 |

Homologs of Rnf cluster (F7O84_03275-3295; EC 7.2.1.2) and hydrogenase (F7O84_0945-50, F7O84_04820; EC 1.12.7.2) genes, which were previously implicated in acetogenesis (4, 45, 46), showed high transcript levels (Table 5). Of the two candidate genes for CO dehydrogenase, F7O84_RS11645 showed the highest transcript level (Table 5). As, under the experimental conditions, no hydroxylamine reductase activity was expected, this result reinforced the candidature of F7O84_RS11645 as possible CO dehydrogenase gene. In an attempt to directly investigate if F7O84_RS11645 encoded a functional CO dehydrogenase, its open reading frame was cloned into high-copy-number *E. coli* expression vector. However, enzyme assays with cell extracts of the resulting *E. coli* strain did not yield consistent evidence for either CO dehydrogenase or hydroxylamine dehydrogenase activity (Supplementary material Table S3).

## Identification of proteins with a high homology of the putative CODH within other members of the *Lachnospiraceae* species.

A protein BLAST search (28) of the putative CODH (F7O84_RS11645) was done to investigate if presence of the putative CODH gene also coincided with an apparently incomplete WLP in other members of the *Lachnospiraceae* family. Indeed, 13 sequenced members of the *Lachnospiraceae* family showed predicted proteins with a high homology with the putative CODH (Supplemental materials Table S4). 9 of the 13 *Lachnospiraceae* members were present in the KEGG database (27), Supplemental Material Table S5, and subsequently analysed on the presence or absence of the CODH/ACS complex. All

organisms contained only a partial Wood-Ljungdahl pathway, with the ACS genes not identified. In seven of the members, respectively *Lachnoclostridium saccharolyticum*, *Lachnoclostridium phytofermentans, Pseudobutyrivibrio xylanivorans*, *Butyrivibrio fibrisolvens, Pseudobutyrivibrio xylanivorans* and both *Roseburia* species the full CODH/ACS complex was not identified. Further study is required to elucidate the relevance of the putative CODH for acetogenic metabolism.



# Discussion

Incorporation of carbon from $^{13}C$ labeled bicarbonate into the carbonyl group of acetate supported our previous conclusion, based on product profiles, that acetogenesis occurs in anaerobic, galacturonate-limited enrichment culture of "*Ca.* G. soehngenii" (14). A much lower labelling of the methyl group of acetate indicated that, instead of carbon dioxide, the methyl branch of the Wood-Ljungdahl pathway (WLP) in the "*Ca.* G. soehngenii" enrichment cultures predominantly used formate as a substrate, generated in the anaerobic fermentation of galacturonate (Figure 1). This conclusion is consistent with the low transcript levels of the only putative formate dehydrogenase gene (F7O84_RS07405; EC 1.17.1.9; Table 5) identified in the "*Ca.* G. soehngenii" MAG, the high transcript level of a putative pyruvate-formate lyase gene (F7O84_RS03160, EC 6.2.1.3; Table 4) and the low net production rates of formate in the anaerobic enrichment cultures (Table 1). In contrast, previous labelling studies on acetogens harboring the WLP showed marginal preferential labeling of the carboxyl moiety of acetate (47–49), indicating the use of extracellular $CO_2$ as substrate for both the methyl- and carbonyl-groups of acetate.

While the observed labelling pattern was consistent with acetogenic metabolism of galacturonate via a WLP, this did not rule out involvement of another pathway for carbon fixation in acetate. Involvement of the hydroxypropionate bi-cycle, 3-hydroxypropionate/4-hydroxybutyrate cycle and dicarboxylate/hydroxybutyrate cycle were excluded since no homologs were found in the "*Ca*. G. soehngenii" MAG for the majority of genes associated with these three pathways (Supplementary material Table S2). Key genes were also missing for the reductive pentose phosphate cycle (rPPP) and reductive citric acid cycle (rTCA) (Supplementary material Table S2) and, moreover, neither of these pathways could explain preferential labelling of the carboxyl group of acetate (50, 51). No gene candidates were identified for the glycine cleavage (GCV) system (Supplementary material Table S2, Figure S2) and $^{13}C$-labeled bicarbonate fed into this pathway should result in equal labelling of the methyl and carbonyl groups of acetate (3); Supplementary material Figure S2). Additionally, none of the routes would require the high CO dehydrogenase enzyme activity measured in cell extracts of the "*Ca*. G. soehngenii" enrichment culture. This analysis leaves the WLP as the only known carbon fixation pathway consistent with the observed stoichiometry of fermentation products, the labelling pattern of acetate and, with the notable exception of the CODH complex, genome and transcriptome analysis of "*Ca*. G. soehngenii".

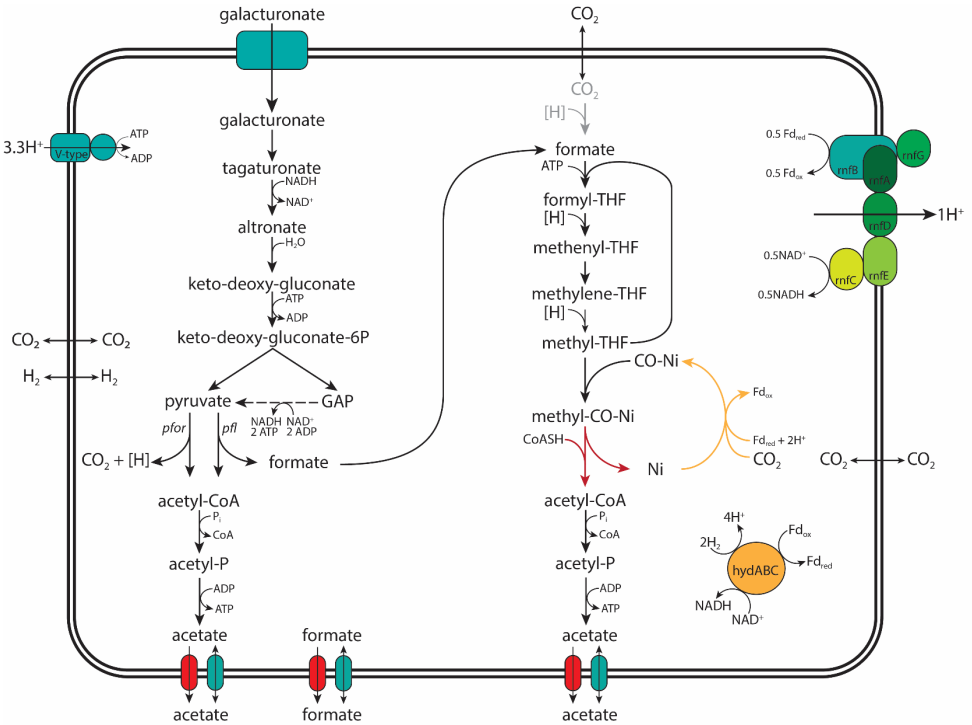



**Figure 1 |** Graphical representation of the proposed pathway for acetogenic galacturonate catabolism in "*Candidatus* Galacturonibacter soehngenii". The conversions of known and annotated genes identified in the MAG and transcribed in the meta-transcriptomic analysis "*Ca*. G. soehngenii" are coloured black, the proposed CO dehydrogenase candidate coloured yellow and the unidentified acetyl-CoA synthase coloured red. With pyruvate:ferredoxin oxidoreductase (*pfor*, EC 1.2.7.1), pyruvate formate lyase (*pfl*, EC 6.2.1.3), ferredoxin hydrogenase (*hydABC*, EC 1.12.7.1) and the Rnf-cluster (*rnfABCDEG*, EC 7.2.1.2) explicitly shown.

Homologs of structural genes encoding enzymes of an adapted Entner-Doudoroff pathway for galacturonate metabolism were highly expressed in the galacturonate-limited, anaerobic "*Ca*. G. soehngenii" enrichment cultures (Table 5). Since conversion of one mole of galacturonate into two moles of pyruvate via this pathway is redox-cofactor neutral, redox equivalents for acetogenesis needed to be derived from pyruvate dissimilation (52, 53). Pyruvate:ferredoxin oxidoreductase (F7O84_RS03200, EC 1.2.7.1) has been reported to couple fermentation and WLP in other anaerobes (4, 54, 55). Strong, highly transcribed homologs of structural genes for PFOR and for a ferredoxin hydrogenase (EC 1.12.7.2) (Table 5; F7O84_RS03200and F7O84_0945-50, F7O84_04820 respectively) indicated that it also fulfils this role in "*Ca*. G. soehngenii".

The significant CO dehydrogenase (CODH) (56) activities in cell extracts enrichment cultures, combined with the incorporation of 13C from bicarbonate in acetate strongly suggested the presence of a functional CODH enzyme in "*Ca*. G. soehngenii". Two highly conserved classes of CODH enzymes have been described (13, 57). Aerobic CODH

enzymes (*coxSML* complex; EC 1.2.5.3) have a Mo-Cu-Se associated active site and only use CO as substrate (58, 59). Strictly anaerobic Ni-Fe-S associated CODH (*cooS*, EC 1.2.7.4) can use also $CO_2$ as substrate (8, 11, 13). A close functional relationship between Ni-CO dehydrogenases and hydroxylamine reductases was shown when a single amino-acid substitution was shown to change a Ni-CO dehydrogenase into a hydroxylamine reductase (44). Since no strong homologs of canonical aerobic or anaerobic CODH genes were identified, the HCP homolog F7O84_RS11645 is therefore the best candidate for the observed CODH activity. Our inability to demonstrate stable CODH activity in cell extracts upon expression of F7O84_RS11645 in *E. coli* could have many causes, including improper folding, metal or co-factor requirements (60, 61) or requirement of additional subunits or other proteins (62–66). The immediate genetic context of F7O84_RS11645 showed many ORFs encoding predicted conserved proteins with unknown function. Co-expression of fosmid libraries (67, 68) of the "*Ca*. G. soehngenii" MAG together with the plasmid used in this study in an E. coli strain, may be helpful in resolving the genetic requirements for CODH activity in this organism.

It remains unclear how the CODH-dependent carbonyl branch and formate-dependent methyl branch of a WLP pathway in "*Ca*. G. soehngenii" organism are linked. The present study is not the first in which carbon fixation linked to the WLP was observed in the absence of a full complement of canonical WLP structural genes (3, 69). However, no clear physiological nor phylogenetic connections were detected between "*Ca*. G. soehngenii" and the organisms studied previously, a strict dehalogenide-respiring *Dehalococcoides mccartyi* strain from the *Chloroflexi* phylum and the phosphite-oxidizing *Deltaproteobacterium* "*Candidatus* Phosphitivorax anaerolimi" Phox-21, respectively.

This study illustrates how quantitative analysis of metabolite formation by chemostat enrichment cultures, combined with $^{13}$C-labelling, (meta-)genome assembly and annotation, meta-transcriptome analysis and biochemical assays can raise new and surprising questions about intensively studied metabolic pathways. Based on our results, involvement of a novel inorganic carbon assimilation pathway, which produces a similar labelling and product profile as the WLP, cannot be fully excluded. However, despite the wide distribution of the CODH/ACS complex in Bacteria and Archaea (46), the available evidence appears to point in the direction of an as yet unidentified link between the methyl and carbonyl branches of the WLP. Further research to resolve this issue may benefit from additional labelling studies with $^{13}$C-bicarbonate, $^{13}$C-formate or partially labelled D-galacturonate combined with metabolome analysis and *in vitro* enzyme activity studies of formate dehydrogenase. Such studies are complicated by our current inability to grow "*Ca*. G. soehngenii" in pure cultures (14). The organisms shown in the Supplemental material Table S4 might be interesting alternative organisms to study in more detail, as they are available in pure culture. It would therefore be relevant to identify if any of these organisms exhibit a similar acetogenic metabolism, with an incomplete complement of WLP enzymes, to further explore this intriguing metabolic conundrum.

## Material and Methods

**Reactor setup and operation.** Chemostat cultures were grown in 1.2 L laboratory bioreactors (Applikon, Delft, The Netherlands), which were stirred at 300 rpm and kept at 30 °C. Anaerobic conditions were maintained by flushing the headspace with nitrogen gas, at a flow rate of 120 mL $min^{-1}$. Culture pH was controlled at 8 ± 0.1 by automatic titration (ADI 1030 Biocontroller, Applikon, Delft, The Netherlands) of 1 M NaOH. The dilution rate was 0.09 ± 0.01 $h^{-1}$ and the working volume of 0.5 L was kept constant by peristaltic effluent pumps (Masterflex, Cole-Parmer, Vernon Hills, USA) coupled to electrical level sensors. Bioreactors were inoculated (10 % v/v) with 50 mL samples of D-galacturonate-limited, anaerobic chemostat enrichment cultures (14), stored in 30 % v/v glycerol at - 20°C. Cultures were run in continuous mode and after at least 6 days (18 generations) stable product composition and biomass concentration were established. System stability was assessed by online monitoring of $CO_2$ production and offline monitoring of fermentation products and optical density. When measurements varied by less than 10 % over multiple volume changes, without a clear upward or downward trend, samples were taken during subsequent cycles.

**Medium.** The cultivation medium contained (g $L^{-1}$): D-galacturonate 4.3; $NH_4Cl$ 1.34; $KH_2PO_4$ 0.78; $Na_2SO_4.10H_2O$ 0.130; $MgCl_2.6H_2O$ 0.120; $FeSO_4.7H_2O$ 0.0031; $CaCl_2$ 0.0006; $H_3BO_4$ 0.0001; $Na_2MoO_4.2H_2O$ 0.0001; $ZnSO_4.7H2O$ 0.0032; $CoCl_2.H_2O$ 0.0006; $CuCl_2.2H_2O$ 0.0022; $MnCl_2.4H_2O$ 0.0025; $NiCl_2.6H_2O$ 0.0005; EDTA 0.10. 19 L of mineral solution (mineral concentration adjusted to the final volume, 20 L) was autoclaved for 20 min at 121°C after which 1 L (86 g $L^{-1}$) D-galacturonate solution was filter sterilized (0.2 µm Mediakap Plus, Spectrum Laboratories, Rancho Dominguez, USA) into the media. 1.5 mL Pluronic PE 6100 antifoam (BASF, Ludwigshafen, Germany) was added per 20 L of mineral solution to avoid excessive foaming.

**Analysis of substrate and extracellular metabolite concentrations.** To determine substrate and extracellular metabolite concentration, reactor sample supernatant was obtained by centrifugation of culture samples (Heraeus Pico Microfuge, ThermoFisher Scientific, Waltman, USA). Concentrations of D-galacturonate and extracellular metabolites were analyzed with an Agilent 1100 Affinity HPLC (Agilent Technologies, Amstelveen, The Netherlands) equipped with an Aminex HPX-87H ion-exchange column (BioRad, Hercules, USA) operated at 60 °C with a mobile phase of 5 mM $H_2SO_4$ and a flow rate of 0.6 mL $min^{-1}$. $CO_2$ and $H_2$ concentrations in the bioreactor exhaust gas were measured using a Prima BT Bench Top mass spectrometer, (ThermoFisher Scientific, Waltman, USA) after the gas was cooled by a condenser (4 °C).

**Biomass dry weight.** 20 mL of culture broth samples were filtered over pre-dried and pre-weighed membrane filters (0.2 µm Supor-200, Pall Corporation, New York, USA), which were then washed with demineralized water, dried in a microwave oven (Robert Bosch GmbH, Gerlingen, Germany) for 20 min at 360 W and reweighed. Carbon and

electron balances were constructed based on the number of carbon atoms and electrons per mole, while biomass composition was assumed to be $CH_{1.8}O_{0.5}N_{0.2}$ (17).

**Quantitative Fluorescent *in situ* hybridization (qFISH) analysis.** Fluorescent *in situ* hybridization was performed as described previously (18), using a hybridization buffer containing 35 % (v/v) formamide. Probes were synthesized and 5' labeled with either 5(6)-carboxyfluorescein-N-hydroxysuccinimide ester (FLUOS) or with one of the sulfoindocyanine dyes (Cy3 and Cy5; Thermo Hybaid Interactiva, Ulm, Germany) (Table 6). The general probe EUB338mix, labeled at both 3' and 5' ends with Cy5, was used to identify all eubacteria in the sample. Microscopic analysis was performed with a LSM510 Meta laser scanning confocal microscope (Carl Zeiss, Oberkochen, Germany). The qFISH analysis was based on at least 29 fields of view at 6730 x magnification, using DAIME (version 2.1) software (DOME, Vienna, Austria; (19)). The bio-volume fractions of "*Ca.* G. soehngenii" and *Enterobacteriaceae* populations were calculated as the ratio of the area hybridizing with specific probes relative to the total area hybridizing with the universal EUBmix probe set (20, 21).



**Table 6 |** Oligonucleotide probes used for the quantitative fluorescence *in situ* hybridization analysis. With W indicating A or T.

| Probe | Sequence (5'-3') | Specificity | Reference |
|---|---|---|---|
| **EUB338mix** | GCWGCCWCCCGTAGGWGT | All bacteria | (20) |
| **ENT** | CTCTTTGGTCTTGCGACG | *Enterobacteriaceae* | (71) |
| **Lac87** | GTGGCGATGCAAGTCTGA | *"Ca.* G. soehngenii" | This study |

**Labelling experiment $^{13}$C-labeled sodium bicarbonate addition.** A 1 M $NaH^{13}CO_3$ solution was used to replace the regular 1 M NaOH solution as a pH titrant in steady-state D-galacturonate-limited enrichment cultures (pH 7.8 ± 0.1, D = 0.1 $h^{-1}$, T = 30 °C). Broth was collected on ice every 2 h for 8 consecutive hours and centrifuged (12,000 x g, Heraeus Pico Microfuge, ThermoFisher Scientific, Waltman, USA) before the supernatant was collected and stored at - 20°C until analysis by NMR. $CO_2$, $H_2$ and $^{13}CO_2$ concentrations in the exhaust gas were measured by MS (Prima BT Bench Top MS, ThermoFisher Scientific, Waltman, USA) after the gas had been cooled by a condenser (4°C).

**Illumina and nanopore sequencing, metagenome assembly and genome binning DNA**. The metagenomic-assembled genome of "*Candidatus* Galacturonibacter soehngenii" described by Valk et al. (2018) was used as template for preparing the metagenome libraries. The DNA extraction, Illumina sequencing, metagenomic assembly and binning process is described in (14). Long-read genomic DNA sequencing was conducted using 1D nanopore sequencing (Oxford Nanopore Technologies, Oxford, UK), following the manufacturer's protocol (LSK-108), omitting the optional DNA shearing and DNA repair steps. The library was loaded on a flow cell (FLO-MIN106) and the MinION Mk1B DNA sequencer (Oxford Nanopore Technologies, Oxford, UK) was used for sequencing

combined with the MinKNOW v. 1.7.3 (Oxford Nanopore Technologies, Oxford, UK) software with the 48 h sequencing workflow (NC_48h_Sequencing_Run_FLO_MIN106_SQK-LSK108.py). Albacore v. 1.2.1 (Oxford Nanopore Technologies, Oxford, UK) was used to base-call the sequencing reads.

**Genome assembly.** The assembling of the contigs from the "*Candidatus* Galacturonibacter soehngenii" genome bin into a single scaffold based on the long Nanopore reads was done using SSPACE-LongRead scaffolder v. 1.1 (22). GapFiller v. 1.11 (23) or by manual read mapping and extension in CLC Genomics Workbench v. 9.5.2 (Qiagen, Hilden, Germany) were used to close gaps in the draft genome with the previously assembled Illumina data. Finally, manual polishing of the complete genome was done to remove SNPs and ensure a high-quality assembly. The meta-genome has been submitted to the sequence read archive (SRA, https://www.ncbi.nlm.nih.gov/sra/) with accession number SRR10674409, under the Bioproject ID PRJNA566068.

**Genome Annotation and Analysis.** The metagenome-assembled genome was uploaded to the automated Microscope platform (24, 25). Manual assessment of pathway annotations was assisted by the MicroCyc (26), KEGG (Kyoto Encyclopaedia of Genes and Genomes;(27)) and SwissProt alignment (BLASTP version 2.2.28+;(28)) databases. The predicted proteome of "*Ca.* G. soehngenii" was submitted to InterProScan (version 5.25-64.0), to identify predictive Pfam domains (29). The annotated genome sequence of "*Candidatus* Galacturonibacter soehngenii" has been submitted to the European Nucleotide Archive (ENA) under the BioProject ID PRJNA566068.

**Genome-centric meta-transcriptomic analyses; RNA extraction and purification.** During pseudo-steady state, broth samples were taken from the enrichment culture, directly frozen in liquid nitrogen and subsequently stored at -80 $^{o}$C. 500 µL samples were thawed on ice, pelleted by centrifugation (21 000 x g, 2 min, 4 °C) and used for total RNA extraction with the RNeasy PowerMicrobiome Kit (Qiagen, Hilden, Germany), following the manufacturer's instruction with the addition of phenol:chloroform:isoamy alcohol (25:25:1) and β-mercaptoethanol (10 µl mL$^{-1}$ final concentration). Cell lysis was with a FastPrep-24 bead beater (MP Biomedicals, Fisher Scientific, Hampton, USA, 4 successive cycles of 40 s at 6.0 m s$^{-1}$, 2 min incubation on ice between cycles). Total RNA extracts were subjected to DNase treatment to remove DNA contaminants by using the DNase Max Kit (Qiagen, Hilden, Germany) and further cleaned up and concentrated with the Agencourt AMpure XP magnetic beads (Beckman Coulter, Brea, USA) before rRNA depletion. Integrity and quality of purified total RNA were assessed on a Tapestation 2200 (Agilent, Santa Clara, USA) with the Agilent RNA screen-tapes (Agilent, Santa Clara, USA) and the concentration was measured using Qubit RNA HS Assay Kit (Thermo Scientific Fisher, Waltman, USA).

**rRNA depletion, library preparation and sequencing.** 500 ng of total RNA from each sample was obtained after rRNA was depleted using the Ribo-Zero rRNA Removal

(Bacteria) Kit (Illumina, San Diego, USA), with 2 µg total RNA as input. Quality of extracted mRNA was checked with Agilent RNA HS screen-tapes (Agilent, Santa Clara, USA) and RNA concentration was determined with a Qubit RNA HS Assay Kit (Thermo Scientific Fisher, Waltman, USA). The TruSeq Stranded mRNA Sample Preparation Kit (Illumina; San Diego, USA) was used to prepare cDNA sequencing libraries according to the manufacturer's instruction. Libraries were sequenced on an Illumina HiSeq2500 using the TruSeq PE Cluster Kit v3-cBot-HS and TruSeq SBS kit v.3-HS sequencing kit (1 x 50 bp; Illumina, San Diego, USA). The raw meta-transcriptome reads have been submitted to the sequence read archive (SRA, https://www.ncbi.nlm.nih.gov/sra/) with accession number SRR10674118-23, under the Bioproject ID PRJNA566068.



**Trimming and mapping of rRNA reads.** Raw RNA reads in FASTQ format were imported into CLC Genomics Workbench v. 9.5.5 and trimmed for quality, requiring a minimum phred score of 20 and a read length of 45. Reads from each sample were hereafter mapped to CDSs obtained from the MAG of "*Ca.* G. soehngenii" with a minimum similarity of 98 % over 80 % of the read length. Reads per kilobase of transcript per million mapped reads (RPKM) were calculated based on raw read-counts and the length of each CDS. The meta-transcriptome mapped to the genome of "*Ca.* G. soehngenii" are shown in Supplemental Data Unconventional_heteroacetogenesis_transcriptome_analysis.xlsx.

**Plasmid and strain construction.** Gene GSOE_50986 was codon optimized for expression in *Escherichia coli* with the GeneArt online tool and integrated behind the TEV recognition site of the pET151/D-TOPO expression vector by GeneArt (GeneArt GmbH, Regensburg, Germany). The resulting plasmid was transformed into a chemically competent *E. coli* strain BL21 according to manufacturer's instructions (NEBuilder HiFi DNA Assembly Master Mix chemical transformation protocol (E2621), New England Biolabs, Ipswich, USA) and named pUD1074. The plasmid sequence of pUD1074 has been deposited at the NCBI GenBank (www.ncbi.nlm.nihgov/genbank) with the corresponding accession number MN498128.

**Heterologous expression of the putative CO dehydrogenase candidate.** All *E.coli* cultures were performed in 120 mL capped bottles with 50 mL of mineral medium (30). Prior to inoculation, the bottles were autoclaved at 120 °C after which the mineral media was supplemented with autoclaved (120°C, 20 min); glucose 5 g $L^{-1}$, peptone (BD Bacto Difco, ThermoFisher Scientific, Waltman, USA) 1 g $L^{-1}$, yeast extract (BD Bacto Difco, ThermoFisher Scientific, Waltman, USA) 2 g $L^{-1}$ and cysteine 1 g $L^{-1}$. Additionally, 0.05 g $L^{-1}$ ampicillin was added and the gas phase was exchanged with air, with a final pressure of 170 kPa. All *E. coli* cultures used for measurements were inoculated with overnight grown pre-cultures (1:50 v/v) and incubated at 37 °C and shaken (300 rpm) until oxygen was depleted (2-3 h). Subsequently 1 mL (250 g $L^{-1}$) glucose, 1 mL reducing agent (0.4 M cysteine) and 1 mL IPTG (40 mM) were added.

After 3 h (at 30 °C, unshaken) of incubation, the cells were harvested and processed anaerobically according to Diender et al., (2016). Enzymatic activity analysis was conducted using a modified method initially described by (30). The essays were performed in an anaerobic environment using 100-300 µL of cell extract with both CO and hydroxylamine as substrate. To increase metal cofactor availability, 1:200 (v/v) metals solution was added to the assay buffer which contained in (g $L^{-1}$); HCl 1.8, $H_3BO_3$ 0.0618, $MnCl_2$ 0.06125, $FeCl_2$ 0.9435, $CoCl_2$ 0.0645, $NiCl_2$ 0.01286, $ZnCl_2$ 0.0677, $CuCl_2$ 0.01335.



**Homology protein BLAST analysis.** The sequence of the putative CODH (F7O84_RS11645) was blasted with the BLASTp (version 2.2.28+, (28) tool of the JGI-IMG/M database (31), with default parameter settings. Finished genomes from members of the *Lachnospiraceae* family in the public JGI-IMG/M database (31) were selected for analysis, Supplemental Material Table S4. The stains identified in the BLAST search, or closely related strains (Supplemental Material Table S5) were subsequently analysed in KEGG (27) for presences of the CODH/ACS complex with pathway map 1200.

## Acknowledgements

This research was supported by the SIAM Gravitation Grant 024.002.002, the Netherlands Organization for Scientific Research. The authors thank J.L. Rombouts for the design of the FISH probe Lac87 used in this study and P. de Waard for the assistance with NMR analysis. The authors additionally thank dr. A.M. Vos and dr. M.D. Verhoeven for critical reading of the manuscript. The authors declare that the research was conducted in the absence of commercial or financial relationships that could be construed as a potential conflict of interest.

### Author Contributions Statement

MvL, JP and LV designed the experiments, interpreted the results and wrote the manuscript. LV did all cultivations and labelling study. LV and MD performed the enzyme activity assays and heterologous experiment. LV and JF performed the qFISH analysis. GS made the model. MSD performed the experimental work for the meta-transcriptomic and meta-genomic analysis. All authors read and approved of the final manuscript.

# Supplemental material



**Supplementary Calculations 1 | Determination of the fraction of Wood-Ljungdahl pathway in the overall acetate production rate.**

The overall reaction of the fermentation of D-galacturonate via the adapted Entner-Doudoroff pathway is shown in equation S2.a, with formate shown as $H_2$ and $CO_2$ for simplification. Equation S2.b shows the overall reaction of the Wood-Ljungdahl pathway.

(S2.a) – 1 D-Galacturonate – 1 $H_2O$ + 2 $H_2$ + 2 $CO_2$ + 2 Acetate
(S2.b) – 4 $H_2$ – 2 $CO_2$ + 1 Acetate +2 $H_2O$

Acetogenesis was hypothesized to occur in "*Candidatus* Galacturonibacter soehngenii" with the total acetate production rate (mmol $h^{-1}$) a function of the two reactions, S2.a and S2.b. The acetate production rate per pathway could be determined, with the D-galacturonate consumption rate and hydrogen and carbon dioxide production rate known, as shown in Table 1 in the main text.

The rate of S2.a could be determined by calculating the catabolic D-galacturonate consumption rate, subtracting the D-galacturonate going towards biosynthesis from the total D-galacturonate consumption rate (0.81 ± 0.01 mmol $h^{-1}$), as D-galacturonate was solely consumed in reaction S2.a.

The total hydrogen production rate (0.18 ± 0.1 mmol $h^{-1}$) was constituted from a hydrogen production rate (S2.a) and consumption rate (S2.b), enabling the determination of the rate of S2.b) (0.35 ± 0.03 mmol $h^{-1}$). Based on these calculations the fraction acetate produced from the WLP was determined to be 0.35.

**Supplementary Calculations 2 | Chemostat model acetogenesis with labeled $NaH^{13}CO_3$**

A principal model was used to estimate the fraction of labelled acetate ($CH_3{}^{13}_{6}COOH$) versus unlabelled acetate ($CH_3{}^{12}_{6}COOH$) over time ($t$) in a chemostat system where at $t = 0$ a constant titration of $NaH^{13}_{6}CO_3$ led to a shift in the fraction of labelled versus unlabelled inorganic carbon in the reactor broth. The model assumed that the carboxyl group ($COOH$) in the anabolism of acetate was directly derived from the dissolved inorganic carbon. Carbon dioxide transfer from the liquid to the gas phase follows the ratio of labelled versus unlabelled inorganic carbon in the liquid phase. Values for the acetate concentration, dissolved inorganic carbon concentration, carbon transfer rate, and volumetric production rates were derived from online and offline measurements (Supplementary Table S1).

**Supplementary Table S1 |** Kinetic and process values for chemostat model.

| Description | Symbol | Value | unit |
|---|---|---|---|
| Dilution rate | D | 0.1 | $h^{-1}$ |
| $CO_2$ production rate | $R_{CO_2}$ | 1.05 | mmol $h^{-1}$ |
| $HCO_3^-$ titration rate | $R_{HCO_3^-}$ | 6.0 | mmol $h^{-1}$ |
| *Acetate* production rate | $R_{Ac}$ | 1.83 | mmol $h^{-1}$ |
| Fraction WLP | $f_{WLP}$ | 0.35 | mmol $mmol^{-1}$ |
| IC[a] $(t = 0, {}^{12}_{6}CO_2)$ | $IC_0^{12}$ | 5 | mmol |
| IC[a] $(t = 0, {}^{13}_{6}CO_2)$ | $IC_0^{13}$ | 0.05 | mmol |
| Carbon transfer rate[c] | CTR | 2.2 | mmol $h^{-1}$ |
| Ac[b] $(t = 0, {}^{12}_{6}CO_2)$ | $Ac_0^{12}$ | 6.40 | mmol |
| Ac[b] $(t = 0, {}^{13}_{6}CO_2)$ | $Ac_0^{13}$ | 0.06 | mmol |

a. Inorganic carbon, dissolved $\boldsymbol{CO_2}$ species of labeled or unlabeled isotopes.
b. Total acetate of which the carboxyl group is either labeled or unlabeled.
c. The carbon transfer rate is the total inorganic carbon transferring from the liquid to the gas phase.

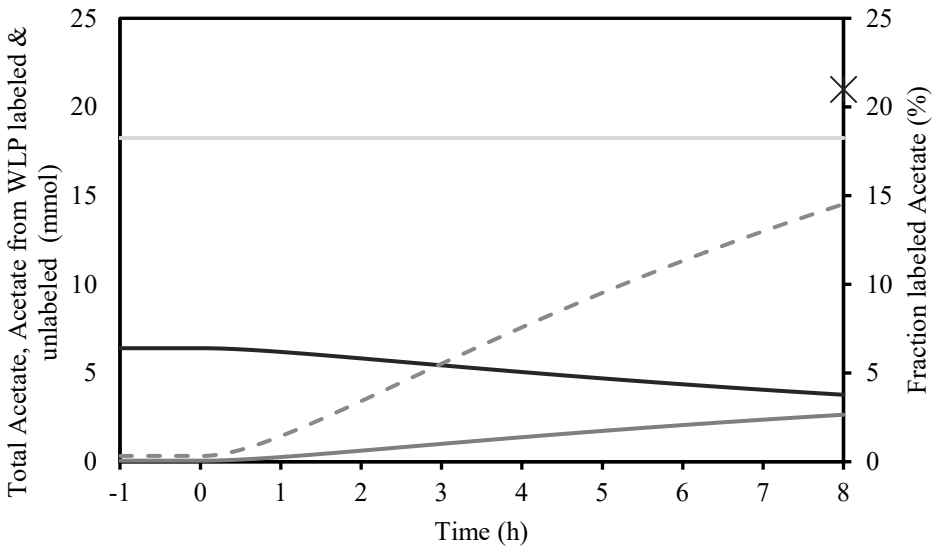
$$f_{IC^{12}} = \frac{IC^{12}}{IC^{12} + IC^{13}} \tag{1}$$

$$\frac{dIC^{12}}{dt} = -D \cdot IC^{12} + R_{CO_2} - f_{IC^{12}} \cdot CTR \tag{2}$$

$$\frac{dIC^{13}}{dt} = -D \cdot IC^{13} + R_{HCO_3^-} - (1 - f_{IC^{12}}) \cdot CTR \tag{3}$$

$$\frac{dAc^{12}}{dt} = -D \cdot Ac^{12} + f_{WLP} \cdot f_{IC^{12}} \cdot R_{Ac} \tag{4}$$

$$\frac{dAc^{13}}{dt} = -D \cdot Ac^{13} + f_{WLP} \cdot (1 - f_{IC^{12}}) \cdot R_{Ac} \tag{5}$$




**Supplementary Figure S1 |** The chemostat model of the "*Ca.* G. soehngenii" enrichment culture of the expected acetate produced over time during the labelling experiment with acetate$_{12}$ (black line (-), mmol) and acetate$_{13}$ (dark grey line (-), mmol), the total amount of acetate produced (light grey line (-), mmol) from the Wood-Ljungdahl pathway, the percentage of labelled acetate versus unlabelled acetate in the reactor broth (dotted grey line (- -), %) and the labelled acetate measured after 8 h (×, %) in bioreactor 2.

**Supplementary Table S2 |** Genes of the six carbon fixation pathways, excluding the Wood-Ljungdahl pathway, from the predictive proteome of the MAG "*Ca*. G. soehngenii" with gene ID and EC number based on SwissProt alignment (BLASTP version 2.2.28+, MicroScope platform v3.13.2).

| Route | EC number | Gene ID |
|---|---|---|
| ***Reductive pentose phosphate cycle (rPPP)*** | | |
| Ribulose-bisphosphate carboxylase | 4.1.1.39 | |
| Phosphoglycerate kinase | 2.7.2.3 | F7O84_RS02895 |
| Glyceraldehyde-3-phosphate dehydrogenase | 1.2.1.12 | F7O84_RS02890 |
| Triose phosphate isomerase | 5.3.1.1 | F7O84_RS02900 |
| Fructose-1,6-bisphosphate aldolase | 4.1.2.13 | F7O84_RS12150 |
| Fructose-1,6-bisphosphatase class 3 | 3.1.3.11 | F7O84_RS00620 |
| Transketolase | 2.2.1.1 | F7O84_RS11220-25 |
| Sedoheptulose-bisphosphatase | 3.1.3.37 | |
| Ribose-5-phosphate isomerase A | 5.3.1.6 | F7O84_RS11205 |
| Ribulose-phosphate 3-epimerase | 5.1.3.1 | F7O84_RS11250 |
| ***Reductive citric acid cycle*** | | |
| Pyruvate ferredoxin:oxidoreductase | 1.2.7.1 | F7O84_RS03200 |
| Aconitate hydratase | 4.2.1.3 | F7O84_RS03605 |
| 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase | 1.2.7.3 and 1.2.7.11 | F7O84_RS03615-20 |
| Malate dehydrogenase | 1.1.1.37 | F7O84_RS05485 |
| Isocitrate dehydrogenase (NADP+) | 1.1.1.42 | F7O84_RS06610 |
| Propionyl-CoA carboxylase | 6.4.1.3 | F7O84_RS07040 |
| Pyruvate phosphate dikinase | 2.7.9.1 | F7O84_RS07755 |
| Pyruvate carboxylase subunit B | 6.4.1.1 | F7O84_RS12740 |
| Aconitate hydratase | 4.2.1.3 | F7O84_RS14165 |
| Fumarate hydratase | 4.2.1.2 | F7O84_RS14815 |
| ATP citrate lyase | 2.3.3.8 | |
| Phosphoenolpyruvate carboxylase | 4.1.1.31 | |
| Fumarate reductase | 1.3.5.4 | |
| Succinate-CoA ligase | 6.2.1.5 | |
| ***Hydroxypropionate bi-cycle*** | | |
| Pyruvate ferredoxin:oxidoreductase | 1.2.7.1 | F7O84_RS03200 |
| pyruvate phosphate dikinase | 2.7.9.1 | F7O84_RS07755 |
| Phosphoenolpyruvate carboxylase | 4.1.1.31 | |
| Malate dehydrogenase | 1.1.1.37 | F7O84_RS05485 |
| Fumarate hydratase | 4.2.1.2 | F7O84_RS14815 |
| Succinate dehydrogenase | 1.3.5.1 | |
| Fumarate reductase | 1.3.4.1 | |
| Fumarate reductase (NADH) | 1.3.1.6 | |
| Fumarate reductase (quinol) | 1.3.5.4 | |
| Succinate-CoA ligase | 6.2.1.5 | |

**Table S2** | *Continued*

| | | |
|---|---|---|
| ***3-Hydroxypropionate/4-Hydroxybutyrate cycle*** | | |
| Acetyl-CoA carboxylase | 6.4.1.2 | |
| Manonyl-CoA reductase (NADP dependent) | 1.2.1.75 | |
| 3-hydroxypropionate dehydrogenase (NADP+) | 1.1.1.298 | |
| 3-hydroxypropionyl-CoA synthase | 6.2.1.36 | |
| 3-hydroxypropionyl-CoA dehydratase | 4.2.1.116 | |
| Acrylyl-CoA reductase (NADPH) | 1.3.1.84 | |
| Propionyl-CoA carboxylase | 6.4.1.3 | F7O84_RS03215 |
| Methylmalonyl-CoA epimerase | 5.1.99.1 | |
| Methylmalonyl-CoA mutase | 5.4.99.2 | |
| Succinyl-coA reductase | 1.2.1.76 | |
| Succinate semialdehyde reductase (NADPH) | 1.1.1.- | |
| 4-hydroxybutyrate-CoA ligase | 6.2.1.40 | |
| 4-hydroxybutanoyl-CoA dehydratase | 4.2.1.120 | |
| Enoyl-CoA hydratase | 4.2.1.17 | |
| 3-hydroxyacyl-CoA dehydrogenase | 1.1.1.35 | |
| Acetyl-CoA acetyltransferase | 2.3.1.9 | |
| Malyl-CoA lyase | 4.1.3.24 | |
| 2-methylfumaryl-CoA hydratase | 4.2.1.148 | |
| 2-methylfumaryl-CoA isomerase | 5.4.1.3 | |
| 3-methylfumaryl-CoA hydratase | 4.2.1.153 | |
| (S)-citramalyl-CoA lyase | 4.1.3.25 | |
| ***Dicarboxylate/hydroxybutyrate cycle*** | | |
| Malyl-CoA lyase | 4.1.3.24 | |
| Succinyl-Coa L-malate-CoA transferase | 2.8.3.22 | |
| Malate dehydrogenase | 1.1.1.37 | F7O84_RS05485 |
| Phosphoenolpyruvate carboxylase | 4.1.1.31 | |
| Pyruvate phosphate dikinase | 2.7.9.1 | F7O84_RS07755 |
| Pyruvate:ferredoxin oxidoreductase | 1.2.7.1 | F7O84_RS03200 |
| ***Reductive glycine pathway*** | | |
| Formate dehydrogenase | 1.17.1.9 | F7O84_RS07405 |
| Ferredoxin hydrogenase | 1.12.7.2 | F7O84_RS09545-50 |
| | | F7O84_RS04820 |
| Formate-tetrahydrofolate ligase | 6.3.4.3 | F7O84_RS05385 |
| Methylenetetrahydrofolate dehydrogenase | 3.5.4.9 | F7O84_RS05380 |
| Glycine cleavage system H protein | | |
| Dihydrolipoamide dehydrogenase | 1.8.1.4 | |
| Glycine cleavage system T protein | 2.1.2.10 | |
| Glycine dehydrogenase subunit A | 1.4.4.2 | |
| Glycine dehydrogenase subunit B | 1.4.4.2 | |
| Glycine hydroxymethyl transferase | 2.1.2.1 | F7O84_RS01545 |
| Serine dehydratase | 4.3.1.17 | F7O84_RS04460-5 |

**Supplementary Table S3 |** Enzyme assay of the putative CO dehydrogenase expressed in an *E. coli* BL21 strain. The mean ± average deviation have been determined from independent duplicate experiments. B.d.l. below detection limit.

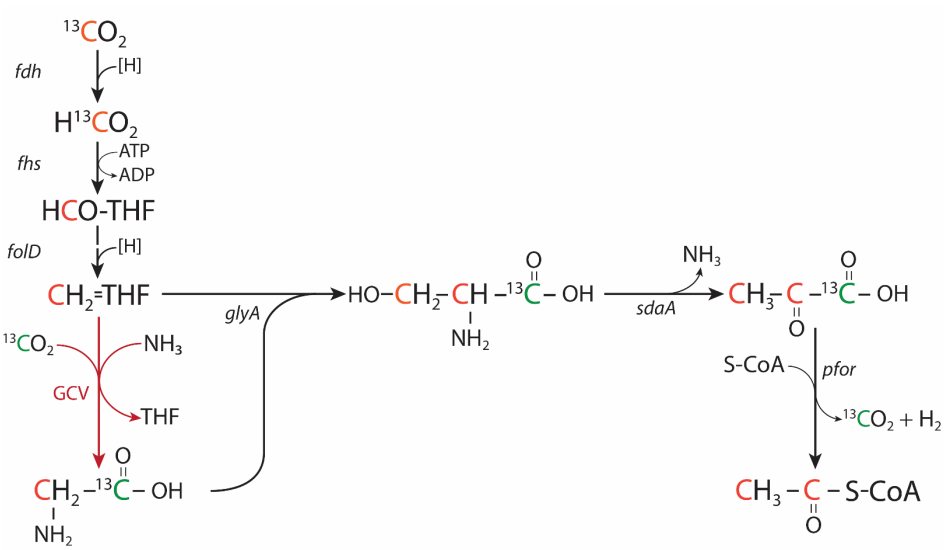
| CO | Hydroxylamine | Metals | pUD1074 | *E. coli* BL21 | Activity (slope/min) |
|---|---|---|---|---|---|
| + | - | + | + | - | 0.013 ± 0.013 |
| - | + | + | + | - | 0.02 |
| + | - | - | + | - | b.d.l. |
| - | - | + | + | - | b.d.l. |
| + | - | + | - | + | b.d.l. |
| - | + | + | - | + | b.d.l. |
| + | - | + | - | - | b.d.l. |

**Supplementary Table S4 |** BLASTp (version 2.2.28+, (28)) search with the organism, genome id, gene (locus tag) and E-value, of the putative CODH (F7O84_RS11645) in *Lachnospiraceae* species obtained from the public JGI-IMG/M database (31).

| Strain | Genome id | Gene | E-value |
|---|---|---|---|
| *Anaerobutyricum hallii* L2-7 | 2826986812 | Ga0265454_2814 | 0 |
| *Pseudobutyrivibrio xylanivorans* MA3014 v2 | 2836274772 | Ga0400149_2313 | 0 |
| *Butyrivibrio fibrisolvens* INBov1 | 2841549333 | Ga0364991_2769 | 0 |
| *Lachnoclostridium phytofermentans* ISDg | 641228486 | Cphy_2846 | 0 |
| *Lachnoclostridium saccharolyticum* WM1, DSM 2544 | 648028018 | Closa_2590 | 0 |
| *Lachnoclostridium* cf. *saccharolyticum* K10 | 650377923 | CLS_14500 | 0 |
| *Roseburia hominis* A2-183, DSM 16839 | 2511231128 | RHOM_10215 | 7.0 $e^{-178}$ |
| *Roseburia intestinalis* L1-82 | 2836830276 | Ga0399605_1213 | 1.0 $e^{-175}$ |
| *Roseburia intestinalis* M50/1 | 650377964 | ROI_37590 | 2.0 $e^{-175}$ |
| *Roseburia intestinalis* XB6B4 | 650377965 | RO1_26150 | 2.0 $e^{-175}$ |
| *Anaerobutyricum hallii* L2-7 | 2826986812 | Ga0265454_1517 | 3.0 $e^{-173}$ |
| *Roseburia hominis* A2-183, DSM 16839 | 2511231128 | RHOM_06150 | 1.0 $e^{-166}$ |
| *Lachnoclostridium saccharolyticum* WM1, DSM 2544 | 648028018 | Closa_2549 | 6.0 $e^{-165}$ |
| *Roseburia intestinalis* L1-82 | 2836830276 | Ga0399605_896 | 8.0 $e^{-160}$ |
| *Roseburia intestinalis* XB6B4 | 650377965 | RO1_03190 | 1.0 $e^{-159}$ |
| *Roseburia intestinalis* M50/1 | 650377964 | ROI_29330 | 2.0 $e^{-159}$ |
| *Lachnoclostridium* sp. YL32 | 2721755796 | Ga0175994_113462 | 4.0 $e^{-152}$ |
| *Lachnoclostridium bolteae* ATCC BAA-613 | 2825694722 | Ga0225995_2847 | 3.0 $e^{-150}$ |
| *Lachnoclostridium phytofermentans* ISDg | 641228486 | Cphy_2847 | 2.0 $e^{-134}$ |

**Table S5 |** Strain, KEGG T number and pathway map used for the KEGG analysis (27) for the presence of the CODH/ACS complex.

| Strain | KEGG T number | Pathway map |
|---|---|---|
| *Anaerobutyricum hallii* EH1 | T05310 | ehl01200 |
| *Pseudobutyrivibrio xylanivorans* MA3014 | T06217 | pxv01200 |
| *Butyrivibrio fibrisolvens* 16/4 | T02581 | bfi01200 |
| *Lachnoclostridium phytofermentans* ISDg | T00619 | cpy01200 |
| *Clostridium saccharolyticum* WM1, DSM 2544 | T01288 | csh01200 |
| *Clostridium* cf. *saccharolyticum* K10 | T02614 | cso01200 |
| *Roseburia hominis* A2-183, DSM 16839 | T01621 | rho01200 |
| *Roseburia intestinalis* M50/1 | T02598 | rim01200 |
| *Roseburia intestinalis* XB6B4 | T02597 | rix01200 |
| *Lachnoclostridium* sp. YL32 | T04430 | lacy01200 |
| *Clostridium bolteae* ATCC BAA-613 | T05466 | cbol01200 |




**Supplementary Figure S2 |** The predicted $^{13}$C-labelling pattern produced by the reductive glycine pathway, with formate dehydrogenase *fdh*, formate-tetrahydrofolate ligase *fhs*, methylenetetrahydrofolate dehydrogenase *folD*, GCV (glycine cleavage system (GcvH, lipoate-binding protein; GcvP, glycine dehydrogenase; GcvT, aminomethyltransferase; Lpd, dihydrolipoyl dehydrogenase); *glyA*, serine hydroxymethyltransferase; *sdaA*, serine deaminase and *pfor*; pyruvate ferredoxin oxidoreductase. The GCV depicted in red was not identified in the genome of "*Ca.* G. soehngenii" (Supplemental material Table S2), the $^{13}$C-labelled carbon integrated via the methyl-branch is depicted in orange and the $^{13}$C-labelled carbon integrated via the GCV is depicted in green (*adapted from* (3))

## References


1. Berg IA. 2011. Ecological aspects of the distribution of different autotrophic $CO_2$ fixation pathways. Appl Environ Microbiol 77:1925–1936.
2. Fuchs G. 2011. Alternative pathways of carbon dioxide fxation: insights into the early evolution of life ? Annu Rev Microbiol 65:631–58.
3. Figueroa IA, Barnum TP, Somasekhar PY, Carlström CI, Engelbrektson AL, Coates JD. 2018. Metagenomics-guided analysis of microbial chemolithoautotrophic phosphite oxidation yields evidence of a seventh natural $CO_2$ fixation pathway. Proc Natl Acad Sci 115:E92–E101.
4. Schuchmann K, Müller V. 2014. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. Nat Rev Microbiol 12:809–821.
5. Bar-Even A, Noor E, Milo R. 2012. A survey of carbon fixation pathways through a quantitative lens. J Exp Bot 63:2325–2342.
6. Adam PS, Borrel G, Gribaldo S. 2018. Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. Proc Natl Acad Sci 115:E1166–E1173.
7. Ragsdale SW, Pierce E. 2008. Acetogenesis and the Wood-Ljungdahl pathway of $CO_2$ fixation. Biochim Biophys Acta - Proteins Proteomics 1784:1873–1898.
8. Ragsdale SW. 2008. Enzymology of the Woods-Ljundahl pathway of acetogenesis. Ann N Y Acad Sci 1125:129–136.
9. Ragsdale SW, Kumar M. 1996. Nickel-containing carbon monoxide dehydrogenase/Acetyl-CoA synthase. Chem Rev 96:2515–2539.
10. Jeoung J-H, Dobbek H. 2011. Ni, Fe-containing carbon monoxide dehydrogenases. Encycl Inorg Bioinorg Chem 1–11.
11. Doukov TI, Iverson TM, Seravalli J, Ragsdale SW, Drennan CL. 2002. A Ni-Fe-Cu center in a bifunctional carbon monoxide dehydrogenase/acetyl-CoA synthase. Science (80- ) 298:567–573.
12. Menon S, Ragsdale SW. 1996. Evidence that carbon monoxide is an obligatory intermediate in anaerobic acetyl-CoA synthesis. Biochemistry 35:12119–12125.
13. Techtmann SM, Lebedinsky A V., Colman AS, Sokolova TG, Woyke T, Goodwin L, Robb FT. 2012. Evidence for horizontal gene transfer of anaerobic carbon monoxide dehydrogenases. Front Microbiol 3:1–16.
14. Valk LC, Frank J, de la Torre-Cortes P, van 't Hof M, van Maris AJA, Pronk JT, van Loosdrecht MCM. 2018. Galacturonate metabolism in anaerobic chemostat enrichment cultures: combined fermentation and acetogenesis by the dominant sp. nov. "*Candidatus* Galacturonibacter soehngenii." Appl Environ Microbiol AEM.01370-18.
15. Drake HL, Gößner AS, Daniel SL. 2008. Old acetogens, new light. Ann N Y Acad Sci 1125:100–128.
16. Schuchmann K, Muller V. 2013. Direct and reversible hydrogenation of $CO_2$ to formate by a bacterial carbon dioxide reductase. Science 342:1382–1386.
17. Roels J. 1983. Energetics and kinetics in biotechnology. Biomedical Press, Amsterdam.
18. Daims H, Stoecker K, Wagner M. 2005. Fluorescence *in situ* hybridization for the detection of prokaryotes, p. 213–239. *In* Osborn, AM, Smith, CJ (eds.), Molecular Microbial Ecology. Taylor & Francis, New York.

19. Daims H, Lücker S, Wagner M. 2006. daime, a novel image analysis program for microbial ecology and biofilm research. Environ Microbiol 8:200–213.
20. Daims H, Brühl A, Amann R, Schleifer K-H, Wagner M. 1999. The domain-specific probe EUB338 is insufficient for the detection of all bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol 22:434–444.
21. Amann RI, Binder BJ, Olson RJ, Chisholm SW, Devereux R, Stahl DA. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Appl Environ Microbiol 56:1919–1925.
22. Boetzer M, Pirovano W. 2014. SSPACE-LongRead: Scaffolding bacterial draft genomes using long read sequence information. BMC Bioinformatics 15:1–9.
23. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13.
24. Vallenet D, Labarre L, Rouy Z, Barbe V, Bocs S, Cruveiller S, Lajus A, Pascal G, Scarpelli C, Médigue C. 2006. MaGe: A microbial genome annotation system supported by synteny results. Nucleic Acids Res 34:53–65.
25. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, Mercier J, Renaux A, Rollin J, Rouy Z, Roche D, Scarpelli C, Medigue C. 2017. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res 45:D517–D528.
26. Caspi R, Foerster H, Fulcher CA, Kaipa P, Krummenacker M, Latendresse M, Paley S, Rhee SY, Shearer AG, Tissier C, Walk TC, Zhang P, Karp PD. 2008. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 36:623–631.
27. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: Back to metabolism in KEGG. Nucleic Acids Res 42:D199–D205.
28. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.
29. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2018. The Pfam protein families database in 2019. Nucleic Acids Res 47:427–432.
30. Diender M, Pereira R, Wessels HJCT, Stams AJM, Sousa DZ. 2016. Proteomic analysis of the hydrogen and carbon monoxide metabolism of *Methanothermobacter marburgensis*. Front Microbiol 7:1–10.
31. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG: The integrated microbial genomes database and comparative analysis system. Nucleic Acids Res 40:115–122.
32. De Kok S, Meijer J, van Loosdrecht MCM, Kleerebezem R. 2013. Impact of dissolved hydrogen partial pressure on mixed culture fermentations. Appl Microbiol Biotechnol 97:2617–2625.
33. Rumble JR, Linde DR, Bruno TJ. 2017. CRC Handbook of Chemistry and Physics, 98th ed. CRC Press LLC.
34. Jain M, Olsen HE, Paten B, Akeson M. 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol 17.

35. Deamer D, Akeson M, Branton D. 2016. Three decades of nanopore sequencing. Nat Biotechnol 34:518–524.
36. Eggen RIL, Geerling ACM, Jetten MSM, De Vos WM. 1991. Cloning, expression, and sequence analysis of the genes for carbon monoxide dehydrogenase of *Methanothrix soehngenii*. J Biol Chem 266:6883–6887.
37. Eggen RIL, van Kranenburg R, Vriesema AJM, Geerling ACM, Verhagen MFJM, Hagen WR, de Vos WM. 1996. Carbon monoxide dehydrogenase from *Methanosarcina frisia* Gö1. J Biol Chem 271:14256–14263.
38. Maupin-Furlow JA, Ferry JG. 1996. Analysis of the CO dehydrogenase/acetyl-coenzyme A synthase operon of *Methanosarcina thermophila*. J Bacteriol 178:6849–6856.
39. Darnault C, Volbeda A, Kim EJ, Legrand P, Vernède X, Lindahl PA, Fontecilla-Camps JC. 2003. Ni-Zn-[Fe4-S4] and Ni-Ni-[Fe4-S4] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. Nat Struct Biol 10:271–279.
40. Wolfe MT, Heo J, Garavelli JS, Ludden PW. 2002. Hydroxylamine reductase activity of the hybrid cluster protein from *Escherichia coli*. J Bacteriol 184:5898–5902.
41. van den Berg WAM, Hagen WR, van Dongen WMAM. 2000. The hybrid-cluster protein ('prismane protein') from *Escherichia coli*. Eur J Biochem 267:666–676.
42. Almeida CC, Romão C V., Lindley PF, Teixeira M, Saraiva LM. 2006. The role of the hybrid cluster protein in oxidative stress defense. J Biol Chem 281:32445–32450.
43. Aragão D, Macedo S, Mitchell EP, Romão C V., Liu MY, Frazão C, Saraiva LM, Xavier A V., LeGall J, Van Dongen WMAM, Hagen WR, Teixeira M, Carrondo MA, Lindley P. 2003. Reduced hybrid cluster proteins (HCP) from *Desulfovibrio desulfuricans* ATCC 27774 and *Desulfovibrio vulgaris* (Hildenborough): X-ray structures at high resolution using synchrotron radiation. J Biol Inorg Chem 8:540–548.
44. Heo J, Wolfe MT, Staples CR, Ludden PW. 2002. Converting the NiFeS carbon monoxide dehydrogenase to a hydrogenase and a hydroxylamine reductase. J Bacteriol 184:5894–5897.
45. Biegel E, Müller V. 2010. Bacterial $Na^+$-translocating ferredoxin: $NAD^+$ oxidoreductase. Proc Natl Acad Sci 107:18138–18142.
46. Schuchmann K, Müller V. 2016. Energetics and application of heterotrophy in acetogenic bacteria. Appl Environ Microbiol 82:4056–4069.
47. Wood HG, Harris DL. 1952. A study of carbon dioxide fixation by mass determination of the types of $C^{13}$-acetate. J Biol Chem 194:905–931.
48. Schulman M, Donald Parker, Ljungdahl LG, Wood HG. 1972. Total synthesis of acetate from $CO_2$ V. Determination by mass analysis of the different types of acetate formed from $^{13}CO_2$ by heterotrophic bacteria. J Bacteriol 109:633–644.
49. O'Brien WE, Ljungdahl LG. 1972. Fermentation of fructose and synthesis of acetate from carbon dioxide by *Clostridium formicoaceticum*. J Bacteriol 109:626–632.
50. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2002. Molecular Biology of the Cell, 4th ed. Garland Science, New York.
51. Shimizu R, Dempo Y, Nakayama Y, Nakamura S, Bamba T, Fukusaki E, Fukui T. 2015. New insight into the role of the Calvin cycle: reutilization of $CO_2$ emitted through sugar degradation. Sci Rep 5:11617.
52. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT. 2006. Alcoholic

fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 90:391–418.
53. Kuivanen J, Biz A, Richard P. 2019. Microbial hexuronate catabolism in biotechnology. AMB Express 9:1–11.
54. Menon S, Ragsdale SW. 1996. Unleashing hydrogenase activity in carbon monoxide dehydrogenase/acetyl-CoA synthase and pyruvate:ferredoxin oxidoreductase. Biochemistry 35:15814–15821.
55. Drake HL, Hu SI, Wood HG. 1981. Purification of five components from *Clostridium thermoaceticum* which catalyze synthesis of acetate from pyruvate and methyltetrahydrofolate. Properties of phosphotransacetylase. J Biol Chem 256:11137–11144.
56. Weghoff MC, Müller V. 2016. CO metabolism in the thermophilic acetogen *Thermoanaerobacter kivui*. Appl Environ Microbiol 82:2312–2319.
57. King GM, Weber CF. 2007. Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. Nat Rev Microbiol 5:107–118.
58. Schübel U, Kraut M, Mörsdorf G, Meyer O. 1995. Molecular characterization of the gene cluster *coxMSL* encoding the molybdenum-containing carbon monoxide dehydrogenase of *Oligotropha carboxidovorans*. J Bacteriol 177:2197–2203.
59. Dobbek H, Gremer L, Meyer O, Huber R. 1999. Crystal structure and mechanism of CO dehydrogenase, a molybdo iron-sulfur flavoprotein containing S-selanylcysteine. Proc Natl Acad Sci 96:8884–8889.
60. Kerby RL, Ludden PW, Roberts GP. 1997. *In vivo* nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: molecular and physiological characterization of *cooCTJ*. J Bacteriol 179:2259–2266.
61. Ensign SA, Campbell MJ, Ludden PW. 1990. Activation of the nickel-deficient carbon monoxide dehydrogenase from *Rhodospirillum rubrum*: kinetic characterization and reductant requirement. Biochemistry 29:2162–2168.
62. Aragão D, Mitchell EP, Frazão CF, Carrondo MA, Lindley PF. 2008. Structural and functional relationships in the hybrid cluster protein family: structure of the anaerobically purified hybrid cluster protein from *Desulfovibrio vulgaris* at 1.35 Å resolution. Acta Crystallogr Sect D Biol Crystallogr 64:665–674.
63. Bar-Even A, Flamholz A, Noor E, Milo R. 2012. Thermodynamic constraints shape the structure of carbon fixation pathways. Biochim Biophys Acta - Bioenerg 1817:1646–1659.
64. Bonams D, Luddent PW. 1987. Purification and characterization of carbon monoxide. J Biol Chem 262:2980–2987.
65. Bonam D, Lehman L, Roberts GP, Ludden PW. 1989. Regulation of carbon monoxide dehydrogenase and hydrogenase in *Rhodospirillum rubrum*: Effects of CO and oxygen on synthesis and activity. J Bacteriol 171:3102–3107.
66. Ensign SA, Ludden PW. 1991. Characterization of the CO oxidation/$H_2$ evolution system of *Rhodospirillum rubrum*: Role of a 22-kDa iron-sulfur protein in mediating electron transfer between carbon monoxide dehydrogenase and hydrogenase. J Biol Chem 266:18395–18403.
67. Ho JCH, Pawar S V., Hallam SJ, Yadav VG. 2018. An improved whole-cell biosensor for the discovery of lignin-transforming enzymes in functional metagenomic screens. ACS Synth Biol 7:392–398.
68. Shizuya H, Birren B, Kim U-J, Mancino V, Slepak T, Tachiiri Y, Simon M. 1992.

Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc Natl Acad Sci 89:8794–8797.

69. Zhuang W, Yi S, Bill M, Brisson VL, Feng X, Men Y, Conrad ME, Tang YJ, Alvarez-Cohen L. 2014. Incomplete Wood–Ljungdahl pathway facilitates one-carbon metabolism in organohalide-respiring *Dehalococcoides mccartyi*. Proc Natl Acad Sci 111:6419–6424.
70. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 1043–1055.
71. Kempf VAJ, Trebesius K, Autenrieth IB. 2000. Fluorescent *in situ* hybridization allows rapid identification of microorganisms in blood cultures. J Clin Microbiol 38:830–838.

# Chapter 4 | A novel D-galacturonate fermentation pathway in *Lactobacillus suebicus* links initial reactions of the galacturonate-isomerase route with the phosphoketolase pathway

L.C. Valk, M.A.H. Luttik, C. de Ram, M. Pabst, M. van den Broek, M.C.M. van Loosdrecht and J.T. Pronk

Keywords: galacturonic acid, pectin degradation, Entner-Doudoroff pathway, enrichment cultivation, heterolactic fermentation, anaerobic metabolism

Manuscript accepted for publication in: *Frontiers in Microbiology, Section Microbial Physiology and Metabolism*.

## Abstract

D-galacturonate, a key constituent of pectin, is a ubiquitous monomer in plant biomass. Anaerobic, fermentative conversion of D-galacturonate is therefore relevant in natural environments as well as in microbial processes for microbial conversion of pectin-containing agricultural residues. In currently known microorganisms that anaerobically ferment D-galacturonate, its catabolism occurs via the galacturonate-isomerase pathway. Redox-cofactor balancing in this pathway strongly constrains the possible range of products generated from anaerobic D-galacturonate fermentation, resulting in acetate as the predominant organic fermentation product. To explore metabolic diversity of microbial D-galacturonate fermentation, anaerobic enrichment cultures were performed at pH 4. Anaerobic batch and chemostat cultures of a dominant *Lactobacillus suebicus* strain isolated from these enrichment cultures produced near-equimolar amounts of lactate and acetate from D-galacturonate. A combination of whole-genome sequence analysis, quantitative proteomics, enzyme activity assays in cell extracts and *in vitro* product identification demonstrated that D-galacturonate metabolism in *L. suebicus* occurs via a novel pathway. In this pathway, mannonate generated by the initial reactions of the canonical isomerase pathway is converted to 6-phosphogluconate by two novel biochemical reactions, catalysed by a mannonate kinase and a 6-phosphomannonate 2-epimerase. Further catabolism of 6-phosphogluconate then proceeds via known reactions of the phosphoketolase pathway. In contrast to the classical isomerase pathway for D-galacturonate catabolism, the novel pathway enables redox-cofactor-neutral conversion of D-galacturonate to ribulose-5-phosphate. While further research is required to identify the structural genes encoding the key enzymes for the novel pathway, its redox-cofactor coupling is highly interesting for metabolic engineering of microbial cell factories for conversion of pectin-containing feedstocks into added-value fermentation products such as ethanol or lactate. This study illustrates the potential of microbial enrichment cultivation to identify novel pathways for the conversion of environmentally and industrially relevant compounds.

# Introduction

Pectin, a commonly occurring polymer in plants, has a linear backbone of D-galacturonate residues or, in the case of rhamnogalacturonan, of alternating D-galacturonate and rhamnose residues. These backbones are decorated with sugars and carbonyl groups of D-galacturonate can be extensively esterified with methyl and acetyl-groups (1, 2). D-galacturonate accounts for approximately 70 % of the mass of pectin (3). During microbial pectin degradation, it is released by the concerted action of pectinases and accessory enzymes (3–5). Knowledge on the pathways involved in microbial fermentation of D-galacturonate is relevant for understanding degradation of plant biomass in anaerobic natural environments and in waste-water treatment processes. It is also highly relevant for the development of anaerobic microbial processes that can convert of large-volume, pectin-rich agricultural residues such as apple pomace, citrus peel and sugar-beet pulp into added-value products (6–8).



Of three experimentally demonstrated pathways for D-galacturonate metabolism, two are found in prokaryotes and one in fungi. The prokaryotic oxidative pathway, first demonstrated in a *Pseudomonas* species (9), converts D-galacturonate to α-ketoglutarate and $CO_2$ via reactions that together reduce 2 moles of $NAD(P)^+$ to NAD(P)H per mole of D-galacturonate (9, 10). In contrast, the reaction sequence that converts D-galacturonate to pyruvate and glycerol in the fungal pathway requires the investment of 2 NAD(P)H per mole of D-galacturonate (11–13). Neither of these routes two routes enable redox-cofactor-neutral, fermentative pathways that generate ATP via substrate-level phosphorylation and they have hitherto only been encountered in microorganisms that are able to respire.

Fermentative, anaerobic metabolism of D-galacturonate is firmly associated with a third pathway. First described in *Escherichia coli* (14–18), this adapted Entner-Doudoroff or isomerase pathway converts D-galacturonate into pyruvate and glyceraldehyde-3-phosphate via 2-keto-3-deoxy-phosphogluconate (KDPG), the characteristic intermediate of the Entner-Doudoroff pathway for sugar dissimilation (19). The canonical isomerase pathway (Figure 1) involves the activity via uronate isomerase (UxaC, EC 5.3.1.12), tagaturonate reductase (UxaB, EC 1.1.1.58), altronate dehydratase (UxaA, EC 4.2.1.7), and 2-keto-3-deoxy-gluconate kinase (KdgK, EC 2.7.1.45) and 2-keto-3-deoxy-phosphogluconate aldolase (KdgA, EC 4.1.2.14). Alternatively, conversion of tagaturonate into 2-keto-3-deoxy-gluconate can be catalysed by tagaturonate 3-epimerase (UxuE, EC 5.1.2.7), fructuronate reductase (UxuB, EC 1.1.1.57) and mannonate dehydratase (UxuA, EC 4.2.1.8) (14–18).

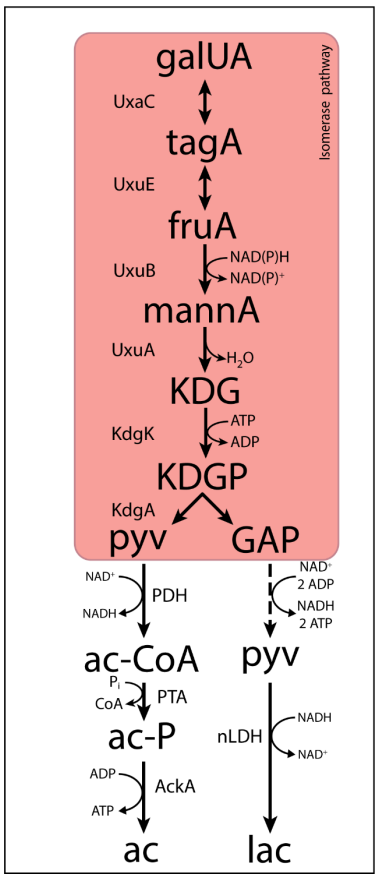



**Figure 1 |** The canonical isomerase pathway for D-galacturonate fermentation. Dashed lines represent multiple conversions. Abbreviations indicate the following metabolites and enzyme activities: galUA, galacturonate; tagA, tagaturonate; fruA, fructuronate; mannA, mannonate; KDG, keto-deoxygluconate; KDGP, keto-deoxy-phosphogluconate; GAP, glyceraldehyde-3-phosphate; ac-P, acetyl-phosphate; pyv, pyruvate; lac, lactate; ac, acetate; and UxaC, uronate isomerase; UxuE, tagaturonate 3-epimerase; UxuB, fructuronate reductase; UxuA, mannonate hydratase; KdgK, keto-deoxy-gluconate kinase; KdgA, keto-deoxy-phosphogluconate aldolase; PDH, pyruvate dehydrogenase; PTA, phosphotransacetylase; AckA, acetate kinase; nLDH, D-/L-lactate dehydrogenase.

In both variants of the isomerase pathway, conversion of D-galacturonate into pyruvate and glyceraldehyde-3-phosphate requires the input of 1 ATP and 1 NAD(P)H. Further conversion of glyceraldehyde-3-phosphate via the lower part of the Embden-Meyerhof glycolysis yields one NADH and two ATP. Use of the isomerase pathway therefore enables redox-cofactor-neutral conversion of D-galacturonate into two moles of pyruvate, with a net ATP yield of 1 mol (mol galacturonate)$^{-1}$ (20–22). This redox-cofactor neutrality constrains the range of fermentation products that can be generated from D-galacturonate. Acetate, which can be formed from pyruvate via redox-cofactor-neutral, ATP-yielding reactions, is typically found as the main product of microbial D-galacturonate fermentation (20, 22–24). For example, in a recent enrichment study on galacturonate performed at pH 8.0, the dominant organism "*Candidatus* Galacturonibacter soehngenii" predominantly produced acetate by a combination of galacturonate fermentation and acetogenesis. In bacteria engineered for ethanol production from D-galacturonate via the isomerase pathway, large amounts of more oxidized by-products are formed (20–22).

As yet undiscovered pathways for D-galacturonate fermentation, that allow for different fermentation product profiles, may exist in nature. Chemical decarboxylation of D-galacturonate to L-arabinose has been reported to occur under relatively mild conditions (25–27) and the possibility that a similar enzyme-catalysed reaction might occur has been proposed (28, 29). Decarboxylation of L-arabinose could enable high-yield microbial production of compounds such as ethanol, lactate or isobutanol. However, no experimental evidence for existence of the required D-galacturonate decarboxylase has yet been found in nature (30, 31).

Based on comparative genome analysis, Rodionova et al., (2012) suggested an alternative pathway for D-galacturonate metabolism involving epimerization of mannonate to gluconate. Subsequent conversion of gluconate by gluconokinase and 6-phosphogluconate dehydrogenase, could also connect of D-galacturonate catabolism to pentose-phosphate metabolism. However, no experimental proof is available for *in vivo* activity of this pathway or for *in vitro* activity of the proposed mannonate epimerase.



Enrichment cultivation remains a powerful tool to select micro-organisms that are optimally adapted to individual and combined sets of environmental parameters from natural or industrial environments. Even when grown on a single substrate, population composition and product profiles in such cultures can be strongly influenced by process parameters such as culture pH, biomass retention time, temperature and dynamic substrate-feeding strategies (33–35). It is well established that, in many fermentative microorganisms, cultivation at low pH induces a transition from the production of acids such as acetic or butyric acid to the formation of non-acidic products such as ethanol or butanol (36, 37). Similarly, enrichment cultivation at pH values below the $pK_a$ of acetic acid (4.76) has been found to select against acetic-acid producing microorganisms, presumably due to a uncoupling of the pH gradient across the cytoplasmic membrane by the undissociated acid (38, 39).

This study explores whether anaerobic enrichment cultivation at pH 4.0 can lead to the discovery of novel pathways for D-galacturonate fermentation. To this end, anaerobic D-galacturonate-limited enrichment cultivation was performed in shake flasks, after which the dominant organism was isolated. Its product profile during anaerobic growth on D-galacturonate, which differed from that in previously described D-galacturonate fermenting bacteria, was characterized in anaerobic bioreactor batch and chemostat cultures and a combination of genomics, proteomics and *in vitro* enzyme-activity analyses was used to investigate the responsible metabolic pathway.

# Results

## *Lactobacillus suebicus* dominates low-pH anaerobic enrichment cultures on D-galacturonate.

Micro-organisms capable of fermenting D-galacturonate at low pH were enriched by serial transfer in anaerobic shake-flask cultures, grown at pH 4.0 on 4 g $L^{-1}$ D-galacturonate. The initial cultures were inoculated with a mixture of rotting orange peels, orange-peel-enriched compost (40). After two transfers to fresh medium, 16S-rRNA gene amplicon sequencing was performed. Based on these short 250-nucleotide amplicon sequences, the dominant taxon was the genus *Lactobacillus* (SINA database 1.2.11, (41), with no other taxa represented by over 1 % of the reads (Supplementary data Figure S1). A pure culture was obtained by repeated anaerobic plating on galacturonate medium. The isolate grew anaerobically on D-galacturonate, L-arabinose, gluconate and D-glucuronate in synthetic media supplemented with 0.4 g $L^{-1}$ yeast extract (Supplementary Material Table S1).



Whole-genome long-read sequencing of the isolate, augmented with short-read sequencing, enabled assembly into one large (2.7 Mbp) and two smaller (89 Kbp and 28 Kbp) contigs (Table 1). The full 16S-RNA gene sequence, derived from the assembled genome sequence showed 100 % identity with that of *Lactobacillus suebicus* (SINA database 1.2.11, (41), a hetero-fermentative lactic acid bacterium previously isolated from apple and pear mashes (42). Analysis with CheckM (43) yielded an estimated completeness of the genome sequence of 99.5 %, with a GC content of 39.1 % and 2811 predicted coding sequences (Table 1). These results are similar to previously published data by Nam et al., (2011) on the genome of *L. suebicus* DSM 5007 shown in Supplemental Material Table S2.

**Table 1 |** Statistical data for the assembled and annotated genome sequence of *Lactobacillus suebicus* LCV1. Completeness and contamination were estimated with CheckM (96).

| | ***Lactobacillus suebicus* LCV1** |
|---|---|
| **Genome size (Mbp)** | 2.8 |
| **Scaffolds** | 3 |
| **Contigs** | 3 |
| **Contigs N50** | 2673450 |
| **Max contig size** | 2673450 |
| **Completeness (%)** | 99.5 |
| **Contamination (%)** | 0 |
| **GC content (%)** | 39.1 |
| **Protein coding density (%)** | 85.5 |
| **Coding density sequences (CDS)** | 2811 |
| **rRNA copies** | 6 |

**D-galacturonate fermentation by *L. suebicus* yields near-equimolar amounts of acetate and lactate.**

The fermentation product profile of D-galacturonate-grown *L. suebicus* LCV1 was investigated in anaerobic bioreactor batch and chemostat cultures. In anaerobic bioreactor batch cultures, *L. suebicus* LCV1 exhibited a specific growth rate on D-galacturonate of 0.20 $h^{-1}$ while acetate and lactate, the sole organic fermentation products, were formed at near-equimolar amounts (Figure 2). A similar product stoichiometry was observed in anaerobic, D-galacturonate-limited chemostat cultures grown at a dilution rate of 0.13 $h^{-1}$, in which biomass-specific production rates of acetate and lactate were 6.0 ± 0.1 mmol $g_{biomass}^{-1}$ $h^{-1}$ and 5.2 ± 0.1 mmol $g_{biomass}^{-1}$ $h^{-1}$, respectively (Table 2). The biomass yield on D-galacturonate in these anaerobic chemostat cultures was 0.09 ± 0.0 g biomass (g galacturonate)$^{-1}$.

**Table 2 |** Physiological parameters of anaerobic, galacturonate-grown chemostat cultures of *Lactobacillus suebicus* LCV1. Cultures were grown at pH 4, at a dilution rate of 0.13 $h^{-1}$ and at 30 °C. Biomass-specific rates of substrate consumption and product formation and product yields are presented as mean ± average deviations of measurements on two independent, steady-state chemostat cultures.

| | **Biomass-specific conversion rates mmol ($g_{biomass}$)$^{-1}$ $h^{-1}$** | **Yield ($mol_i$ mol (galacturonate)$^{-1}$)** |
|---|---|---|
| Galacturonate | -6.9 ± 0.0 | |
| Acetate | 6.0 ± 0.1 | 0.87 ± 0.01 |
| Lactate | 5.2 ± 0.1 | 0.75 ± 0.01 |
| $CO_2$ | 7.2 ± 0.1 | 1.04 ± 0.03 |
| Biomass ($g_{biomass}$ $g_{galacturonate}^{-1}$) | | 0.09 ± 0.0 |

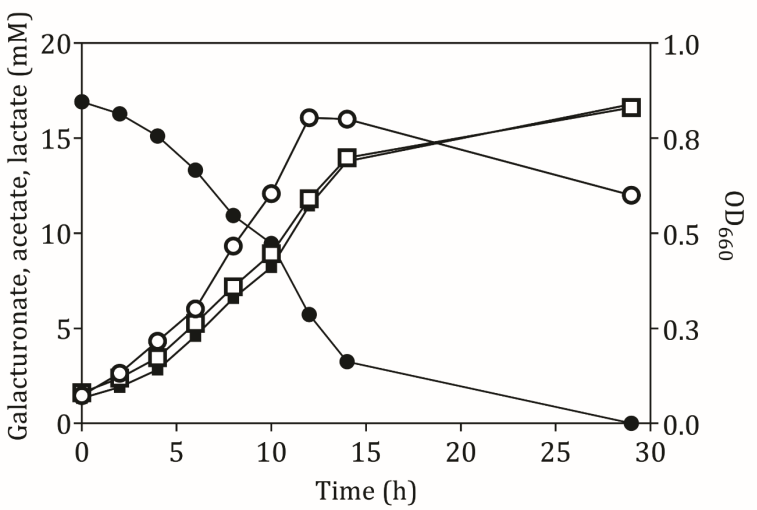



**Figure 2 |** Anaerobic growth and product formation of in an anaerobic bioreactor batch culture of *L. suebicus* LCV1 on D-galacturonate (3.3 g $L^{-1}$) at pH 4 and at 30 °C. Symbols: ● D-galacturonate, ■ acetate, □ lactate and ○ optical density. The data shows one of two independent replicates, data from the second experiment are shown in Supplementary Materials, Figure S2.

## Enzymes involved in the lower half of the isomerase pathway for D-galacturonate metabolism are absent from the proteome of D-galacturonate-grown *L. suebicus* LCV1.

In the canonical isomerase pathway for D-galacturonate fermentation, this substrate is generally first converted into 2 molecules of pyruvate via the isomerase pathway (Figure 1). Since this conversion is redox-cofactor neutral, it could only account for formation of equimolar amounts of acetate and lactate if pyruvate reduction to lactate by a lactate dehydrogenase were stoichiometrically coupled to oxidative decarboxylation of pyruvate to acetate. The latter could, theoretically, be accomplished by the combined action of a pyruvate-dehydrogenase complex, phosphotransacetylase and acetate kinase (Figure 1).

The genome sequence of L. suebicus LCV1 showed a full complement of structural genes for the key enzymes of the isomerase pathway (Figure 3, Table 3). Proteome analysis of D-galacturonate-grown cultures showed high levels of the enzymes of upper part of the adapted ED pathway (uronate isomerase (UxaC, EC 5.3.1.12), tagaturonate 3-epimerase (UxaE, EC 5.1.2.7) and fructuronate reductase (UxuB, EC 1.1.1.57)) (Table 3, Figure 3). However, under the experimental conditions, mannonate dehydratase (UxuA, EC 4.2.1.8), 2-dehydro-3-deoxygluconokinase (KdgK, EC 2.7.1.45) and 2-dehydro-3-deoxyphosphogluconate aldolase (KdgA, EC 4.1.2.14), three key enzymes in the lower part of the isomerase pathway, were not detected in the proteome (Table 3).

The three subunits of the pyruvate-dehydrogenase complex (PDH, EC 1.2.4.1; EC 2.3.1.12; EC 1.8.1.4), as well as phosphate transacetylase (PTA, EC 2.3.1.8), acetate kinase (AckA, EC 2.7.2.1) and both L- and D-lactate dehydrogenase (L(+)-nLDH, EC 1.1.1.27; D-(-)-nLDH EC 1.1.1.28) were all detected in the proteome (Figure 3, Table 3). These results indicated that, under the experimental conditions, enzymes for conversion of pyruvate to lactate and acetate were present in *L. suebicus* LCV1. However, absence of UxuA, KdgK, KdgA in the proteome indicated that, in the anaerobic galacturonate-grown batch cultures, pyruvate formation from galacturonate did not occur via a complete, canonical isomerase pathway.



**Table 3|** Annotated genes identified in the genome of *L. suebicus* LCV1, and associated expressed gene products identified by shot-gun proteomics for the two putative routes the isomerase pathway and the phosphoketolase pathway. Gene name, EC number, genetic nomenclature, E-value based on SwissProt alignment with *Lactobacillus suebicus* DSM 5007 (BLASTP version 2.2.28+, MicroScope platform v3.13.2) and their identified expressed protein products, are listed below. ND = not detected. *Multiple annotated and expressed gene products detected.

| Gene name | EC number | Gene | Identified protein sequence |
|---|---|---|---|
| *Adapted Entner-Doudoroff pathway* | | | |
| Uronate isomerase* | 5.3.1.12 | *uxaC* | A0A0R1WA62; A0A0R1WAR4 |
| Tagaturonate 3-epimerase | 5.1.2.7 | *uxaE* | A0A0R1W3Z7 |
| Fructonate reductase | 1.1.1.57 | *uxuB* | A0A0R1W2X9 |
| Mannonate dehydratase | 4.2.1.8 | *uxuA* | ND |
| 2-Dehydro-3-deoxygluconokinase | 2.7.1.45 | *kdgK* | ND |
| 2-Dehydro-3-deoxyphosphogluconate aldolase | 4.1.2.14 | *kdgA* | ND |
| *Phosphoketolase pathway* | | | |
| 6-phosphogluconate dehydrogenase | 1.1.1.44 | *6pgd* | A0A0R1WAU7 |
| Gluconate kinase* | 2.7.1.12 | *gntK* | A0A0R1W2R7; A0A0R1W6K8 |
| Ribulose-5-phosphate epimerase | 5.1.3.4 | *araD* | A0A0R1W2Q5 |
| Xylulose-5-phosphate phosphoketolase | 4.1.2.9 | *xpkA* | A0A0R1W2X8 |
| *Acid fermentation* | | | |
| Glyceraldehyde-3-phosphate dehydrogenase | 1.2.1.12 | *gapDH* | A0A0R1WD78; A0A0R1W347 |
| Phosphoglycerate kinase | 2.7.2.3 | *pgk* | A0A0R1W7B1 |

**Table 3 |** *Continued*

| Gene name | EC number | Gene | Identified protein sequence |
|---|---|---|---|
| Phosphoglycerate mutase* | 5.4.2.11 | *pgam* | A0A0R1WBW2; A0A0R1W3N5; A0A0R1W8P2; A0A0R1W2R9; A0A0R1W2L2; A0A0R1VYX3; A0A0R1W2D7 |
| Enolase | 4.2.1.11 | *eno* | A0A0R1W6P1 |
| Pyruvate kinase | 2.7.2.3 | *pyk* | A0A0R1W3H0 |
| L - lactate dehydrogenase | 1.1.1.27 | *L -ldh* | A0A0R1W2X2 |
| D - lactate dehydrogenase | 1.1.1.28 | *D-ldh* | A0A0R1W2M1 |
| Pyruvate dehydrogenase subunit E1α | 1.2.4.1 | *pdh E1α* | A0A0R1W4W0 |
| Pyruvate dehydrogenase subunit E1β | 1.2.4.1 | *pdh E1β* | A0A0R1W442 |
| Pyruvate dehydrogenase subunit E2 | 2.3.1.12 | *pdh E2* | A0A0R1W4J5 |
| Pyruvate dehydrogenase subunit E3 | 1.8.1.4 | *pdh E3* | A0A0R1WEK4 |
| Phosphate transferase | 2.3.1.8 | *pta* | A0A0R1W6L8 |
| Acetate kinase | 2.7.2.1 | *ackA* | A0A0R1VTZ1 |
| Transhydrogenase | 1.6.1.2 | *pntAB* | A0A0R1W642; A0A0R1WE53; A0A0R1W6T3 |
| NADH oxidase | 1.6.3.4 | *nox* | A0A0R1WBZ1 |

## Proteome analysis and activities of 'conventional' enzymes in cell extracts indicate involvement of a hybrid isomerase/phosphoketolase pathway.

The phosphoketolase (PK) pathway in lactic acid bacteria is a well-known route for hetero-fermentative dissimilation of sugars to equimolar amounts of lactate, ethanol and CO2 (45–48). It has previously been proposed that D-galacturonate metabolism via the initial reactions of the isomerase pathway can be linked to the PK pathway by epimerization of mannonate to gluconate (32), although this mechanism has not been experimentally demonstrated. Consistent with the involvement of such hybrid pathway, proteome analysis of galacturonate-grown cultures of *L. suebicus* LCV1 showed high levels of a gluconate kinase (GntK, EC 2.7.1.12), 6-phosphogluconate dehydrogenase (GndA, EC 1.1.1.44) and phosphoketolase (XpkA, EC 4.1.2.9), three key enzymes of the PK pathway (Table 3, Figure 3).

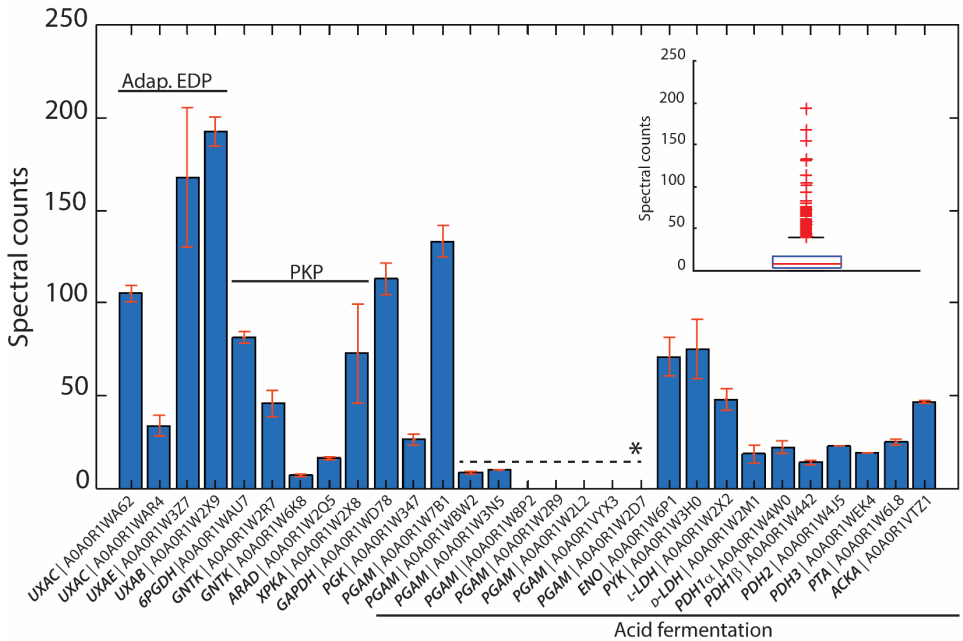



**Figure 3 |** Identified proteins from the isomerase pathway (adap. EDP), phosphoketolase pathway (PKP) and acid fermentation plotted as normalised spectral counts in LC-MS based proteomic analysis. The insert shows the box plot from all identified proteins within a range of 0-250 normalised counts. Only 3 protein hits were above this range, which are not shown. All proteins detected from the isomerase pathway, phosphoketolase pathway and acid fermentation were above the 75th percentile of the complete proteome dataset. Bars represent the average value of two biological reactor duplicates, and error bars show the standard deviation between runs. * Paralogous genes.

*In vitro* activity assays were used to investigate the presence of known enzyme reactions involved in the proposed hybrid isomerase-PK pathway in galacturonate-grown *L. suebicus* LCV1. Under aerobic conditions, cell extracts rapidly oxidized NADH (Table 4). This activity was attributed to an NADH oxidase (NoxA, EC 1.6.3.4), for which a candidate gene (A0A0R1WBZ1, Table 3) and high expression levels of the encoded protein (Figure 3) were detected. Enzyme activity assays that involved $NAD^+$ or NADH were therefore performed under anaerobic conditions. Fructuronate reductase (UxuB, EC 1.1.1.57), a key enzyme of the upper part of the isomerase pathway, was measured with D-tagaturonate and NADH or NADPH as substrates and showed an activity of 0.99 ± 0.01 µmol (mg protein)-1 min-1 and 0.2 ± 0.01 µmol (mg protein)$^{-1}$ min$^{-1}$, respectively. Cell extracts also showed high activities of $NADP^+$-dependent 6-phosphogluconate dehydrogenase (GndA, EC 1.1.1.44; Table 4), the first enzyme of the PK pathway. Activities of NAD+-dependent lactate dehydrogenase (L(+)-nLDH, EC 1.1.1.27; D-(-)-nLDH EC 1.1.1.28) in cell extracts of D-galacturonate-grown *L. suebicus* LCV1 (Table 4) were similar to those observed in cell extracts of sugar-grown cultures of other Lactobacilli (49, 50). Pyruvate-dehydrogenase activity in cell extracts was below detection limit (Table 4).

**Table 4 |** Specific enzyme activities of key enzymes of the isomerase and phosphoketolase pathways of cell extract of galacturonate-grown *L. suebicus* LCV1 (D=0.13 $h^{-1}$, T= 30 °C and pH 4). Both gluconate kinase (GntK) and fructuronate reductase (UxuB) were measured by a coupled reaction with a naturally expressed epimerase, (6-phosphomannonate 2-epimerase and tagaturonate 3-epimerase (EC 5.1.2.7, UxuE), respectively). Mean ± standard deviations were derived from duplicate experiments.

| Enzyme name | Enzyme | Co-factor | Substrate | Specific activity (µmol $mg_{protein}^{-1}$ $min^{-1}$) |
|---|---|---|---|---|
| Fructuronate reductase | UxuB | NADPH | Tagaturonate | 0.24 ± 0.1 |
| Fructuronate reductase | UxuB | NADH | Tagaturonate | 0.99 ± 0.1 |
| Gluconate kinase | GntK | ATP and $NADP^+$ | Mannonate | 0.08 ± 0.0 |
| 6-phospholgluconate dehydrogenase | 6PGD | $NADP^+$ | 6-phospho-gluconate | 1.18 ± 0.1 |
| Lactate dehydrogenase | nLDH | NADH | Pyruvate | 5.80 ± 0.7 |
| NADH oxidase | NOX | NADH | $O_2$ | 0.19 ± 0.0 |
| Pyruvate dehydrogenase | PDH | $NAD^+$ | Pyruvate | < 0.05 |

## Demonstration of mannonate kinase and 6-phosphomannonate-2-epimerase activities in cell extracts.

Conversion of mannonate to 6-phosphogluconate, a key conversion in the proposed hybrid isomerase-PK pathway, has not previously been demonstrated. To investigate whether this conversion occurs in *L. suebicus* LCV1, cell extracts were incubated with different substrates and products were analysed by GC-MS. Incubation (0.5 h) of cell extracts with mannonate did not lead to formation of gluconate (Figure 4A, black line). When ATP was also added, mannonate was predominantly converted to 6-phosphomannonate, thus indicating the activity of a mannonate kinase in the cell extracts. In addition, small amounts of 6-phosphogluconate were formed (Figure 4A, blue line). Upon prolonged incubation of cells extract with mannonate and ATP, formation of 6-phosphomannonate was clearly observed (Figure 4B, red line). In addition, incubation of cell extracts with 6-phosphogluconate led to the formation of 6-phosphomannonate (Figure 4C, black line). These results demonstrated the presence, in galacturonate-grown *L. suebicus*, of 6-phosphomannonate 2-epimerase activity.

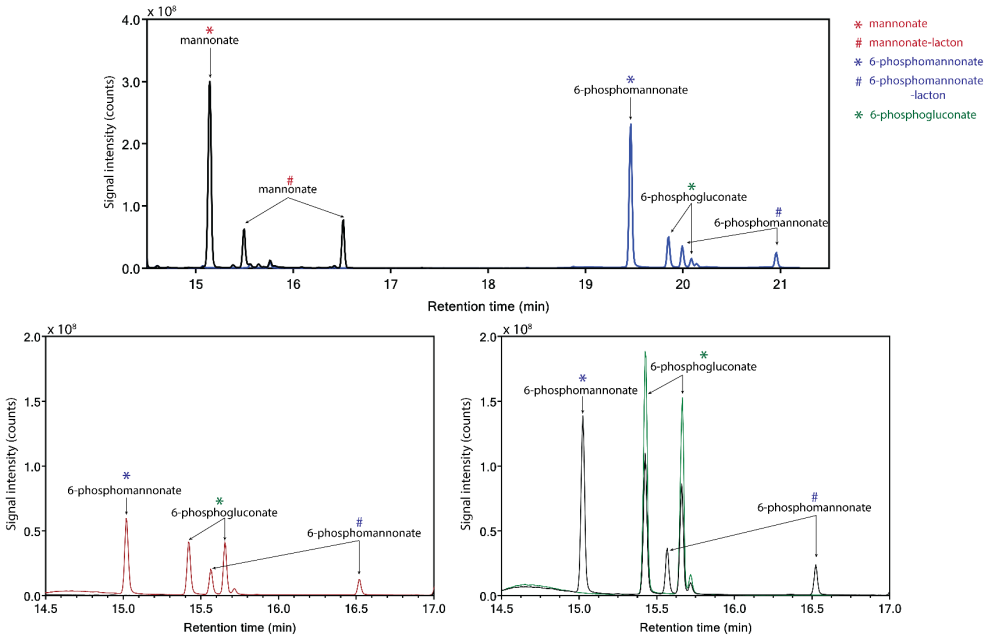



**Figure 4 |** Identification of products from the conversion of mannonate or 6-phosphogluconate, in the presence and absence of ATP, with cell extract from *L. suebicus* LCV1, grown on galacturonate in anaerobic chemostat cultures (D = 0.13 $h^{-1}$, 30 °C, pH 4). A) 0.5 h incubation of mannonate with cell extract in the absence (black line) and presence of ATP (blue line), B) incubation of mannonate with cell extract and ATP, start of incubation (black line) and sample taken after 3 h of incubation (red line), C) incubation of 6-phosphogluconate, start of incubation (green line) and sample taken after 3 h of incubation (black line). For each experiment, one of two of independent biological duplicate experiments is shown. Data for the duplicate experiments are shown in Supplementary Material Figure S3. Different retention times of compounds in the panels are caused by a reduced column length for the experiments shown in panel C. Retention times of the standards (mannonate and 6-phosphogluconate) and the annotated GC-MS profile of mannonate are shown in Supplemental material Figure S4 and S5, respectively.

Continuous enzyme-activity assays with cell extracts, in which activity of gluconate kinase was coupled to 6-phosphogluconate dehydrogenase, revealed a $K_M$ for gluconate of 10 mM and a $V_{max}$ of 0.24 µmol (mg protein)$^{-1}$ min$^{-1}$ (Figure 5A). This $K_M$ was two to three orders of magnitude higher than reported for bacterial gluconate kinases (49–51). The same continuous assay, using the native, unidentified 6-phosphomannonate epimerase activity as a coupling enzyme, was used to investigate the kinetics of mannonate kinase in *L. suebicus* LCV1. Although this assay did not contain the demonstrated excess of coupling-enzyme activity required for reliable kinetic analyses of mannonate kinase activity, it indicated the presence of a high-affinity mannonate kinase activity (Figure 5B, estimated $K_M$ of 0.4 mM). The observed maximum reaction rate and apparent substrate inhibition at higher substrate concentrations could reflect properties of either the mannonate kinase or of the unidentified epimerase enzyme(s) under the conditions of the *in vitro* assay.

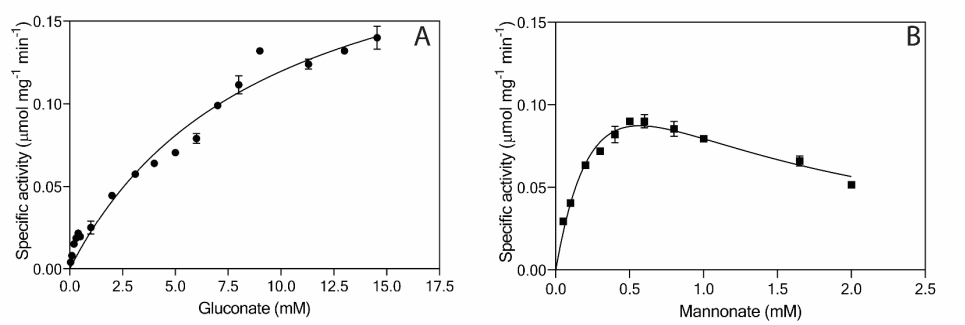





**Figure 5 |** Effect of A) gluconate (●, mM) and B) mannonate concentrations (■, mM) on the specific enzymatic activity (µmol $mg_{protein}^{-1}$ $min^{-1}$) of gluconate kinase and mannonate kinase, respectively, in cell extracts of *L. suebicus* LCV1, pre-grown on galacturonate in anaerobic chemostat cultures (D = 0.13 $h^{-1}$, 30 °C, pH 4). The average ± mean deviations were derived from duplicate measurements.

## Discussion

Enrichment cultivation at low pH enabled the isolation of a lactic acid bacterium with a product profile that had not previously been observed during anaerobic fermentation of D-galacturonate. In contrast to previously described D-galacturonate-fermenting organisms, the isolate, identified as a strain of Lactobacillus suebicus, did not produce acetate as single major fermentation product but, instead, produced lactate and acetate at near-equimolar ratios. Due to its lower $pK_a$ and lower lipid solubility, the uncoupling effect of lactate on the proton gradient across biological membranes is much less pronounced than that of acetic acid (54, 55). Our results are therefore in line with previous studies in which enrichment cultivation at low pH values generated a negative selective pressure against the formation of acetate as predominant fermentation product (56–59). At low extracellular concentrations of lactate, some anaerobic bacteria can couple lactate export via a proton symporter to energy conservation (60, 61). If active in *L. suebicus*, such an 'end-product-efflux' mechanism might provide an additional, condition-dependent advantage over formation of only acetate.

*Lactobacilli* are known for their ability to ferment a wide range of sugars under mildly acidic conditions (62, 63). However, metabolic pathways for D-galacturonate metabolism and the resulting product profiles in *Lactobacilli* have not previously been studied in detail. In the context of the present study, it is relevant to note that apple and pear mashes, from which *L. suebicus* has previously been isolated, are rich in pectin and that, during growth on sugars, the metabolism of *L. suebicus* was characterized as heterofermentative (42).

We initially hypothesized that the observed product stoichiometry reflected a simultaneous, redox-cofactor- balanced reduction and oxidative decarboxylation of pyruvate, generated via a canonical isomerase pathway for D-galacturonate metabolism

(Figure 1; (24, 64)). Although pyruvate dehydrogenase (PDH) complexes are typically inhibited at high NADH/$NAD^+$ ratios (65, 66), such a concerted action of the PDH complex and an $NAD^+$-linked lactate dehydrogenase has been implicated in fermentation of pyruvate by *Oenococcus oeni* and *Leucococcus mesenteroides* (64, 67). However, its involvement in D-galacturonate fermentation by *L. suebicus* was rejected based on the absence of key enzymes of the lower half of the canonical isomerase pathway in the proteome of galacturonate-grown cultures (Table 3, Figure 3).

Based on investigations on *Thermotoga maritima*, a deep-branching, anaerobic, hyperthermophilic bacterium able to degrade a wide range of carbohydrates including pectin (68), Rodionova *et al*., (2012) proposed that the initial reactions of the isomerase pathway for galacturonate might be coupled to the phosphoketolase pathway (Figure 6A) by gluconate 2-epimerase (*gntE*), and a gluconate kinase, (*gntK*, EC 2.1.7.12). This proposal was based on the organization of hexuronate and pectin utilization loci, in which gntE, which was assumed to encode the epimerase activity, was found between the fructuronate reductase (uxuA, EC 1.1.1.58) and gluconate kinase genes. A homology search indicated that gntE homologs also occurred in other bacterial genera, including *Lactobacillus, Acidobacterium* and *Terriglobus* (32). However, no experimental evidence was provided to show that gntE indeed encoded the proposed epimerase.



Discontinuous and continuous enzyme activity assays and product identification by GC-MS showed that, instead of a direct epimerization of mannonate to gluconate followed by phosphorylation of gluconate (32), *L. suebicus* LCV1 first phosphorylated mannonate to 6-phosphomannonate, followed by an epimerization to 6-phosphogluconate (Figure 6B). This newly discovered conversion, which couples the isomerase pathway to the PK pathway, involves two enzyme activities, a mannonate kinase and a 6-phosphomannonate/6-phoshogluconate 2-epimerase, that have not previously been described.

The BRENDA database (69, 70) does not list specific mannonate kinases, nor does it list mannonate as a known substrate for gluconate kinases. Indeed, *E. coli* gluconate kinase (GntK, EC 2.7.1.12) was reported to be unable to use D-mannonate as substrate (71). The high $K_M$ of gluconate kinase activity and the low apparent $K_M$ of mannonate kinase in cell extracts (Figure 4) strongly suggest that at least one of the two *L. suebicus* genes with strong homology to gluconate kinase genes (A0A0R1W2R7, A0A0R1WAU7), whose products were both detected in the proteome of galacturonate-grown cells (Table 3, Figure 3), encodes a kinase with a much higher affinity for mannonate than for gluconate.

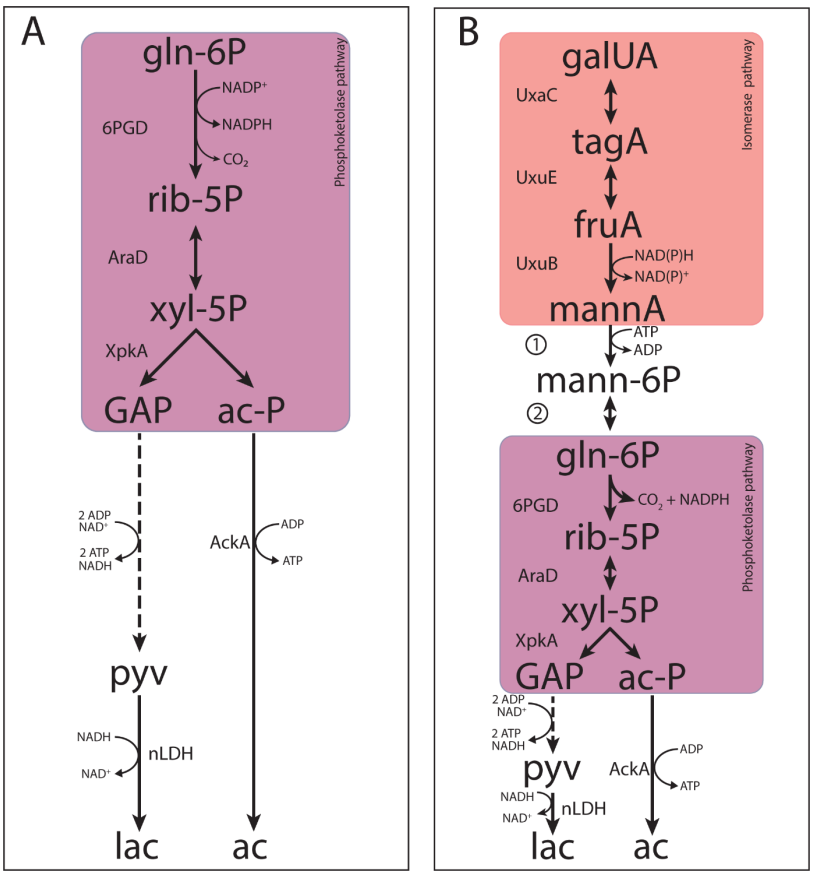



**Figure 6** | Metabolic pathways contributing to fermentative galacturonate fermentation in *L. suebicus* LCV1 as discussed in this paper. A) phosphoketolase pathway and B) novel pathway proposed in this study that combines known reactions of the upper part of the canonical isomerase pathway and the phosphoketolase pathway, coupled by the concerted action of a mannonate kinase and a 6-phosphomannonate 2-epimerase. Dashed lines represent multiple conversions. Abbreviations indicate the following metabolites and enzyme activities: galUA, galacturonate; tagA, tagaturonate; fruA, fructuronate; mannA, mannonate; mann-6P, 6-phosphomannonate; gln-6P, 6-phosphogluconate; rib-5P, ribulose-5-phosphate; xyl-5P, xylulose-5-phosphate; GAP, glyceraldehyde-3-phosphate; ac-P, acetyl-phosphate; pyv, pyruvate; lac, lactate; ac, acetate; and UxaC, uronate isomerase; UxuE, tagaturonate 3-epimerase; UxuB, fructuronate reductase, 1, mannonate kinase; 2, 6-phosphomannonate 2-epimerase; 6pgd, 6-phosphogluconate dehydrogenase (decarboxylating); AraD, ribulose-5-phosphate 4-epimerase; XpkA, phosphoketolase; AckA, acetate kinase; nLDH, D-/L-lactate dehydrogenase.

Four Structural Classification of Proteins (SCOP) families harbour 2-epimerases; NAD+-dependent epimerases or dehydratases (CEP1), N-acylglucosamine epimerases (CEP4), NanE-like (CEP5) and UDP-N-acetylglucosamine 2-epimerases (CEP9). All these epimerases act on compounds with a CDP or UDP side-group or on dimers (72, 73). No homology to structural genes for either of the enzyme groups was identified in the *L. suebicus* LCV1 genome. Based on co-localization with hexuronate and D-gluconate

fermentation pathway genes, Rodionova et al., (2012) proposed that the TM0042 gene in *Thermotoga maritima* encoded a gluconate-mannonate 2-epimerase. The predicted protein sequence showed similarity to an aldose 1-epimerase (EC 5.1.3.3) or a 6-phosphoglucose 1-epimerase (EC 5.1.3.15), but no biochemical evidence for its activity was provided. A homologous gene was identified in *L. suebicus* LCV1 (A0A0R1W1H9) and although the gene was not co-localized with genes of the D-galacturonate fermentation pathway, the encoded protein product was successfully identified with high sequence coverage in the D-galacturonate-grown cultures. Further research is required to investigate whether this protein encodes the 6-phosphomannonate 2-epimerase active in these cultures.



Consistent with the literature on other organisms (71, 74, 75), activity of 6-phosphogluconate dehydrogenases (GndA, EC 1.1.1.44) in cell extracts of *L. suebicus* showed high activities NADP+ as electron acceptor. While fructuronate reductases typically use NADH as co-factor (15, 76–78), cell extracts of *L. suebicus* LCV1 also showed activity of this enzyme with NADPH as electron donor (Table 3). A matching cofactor use off these two enzymes, possibly further facilitated by the expression of a transhydrogenase (PntAB; EC 1.6.1.2, A0A0R1W642; A0A0R1WE53; A0A0R1W6T3), enables redox-cofactor balancing in the integrated isomerase-PK pathway (Figure 6B).

Enzyme purification and characterization, possibly combined with heterologous expression studies and/or generation and analysis of *L. suebicus* mutants, should resolve the genetic basis for the mannonate kinase and 6-phosphomannonate/6-phosphogluconate-2-epimerases in *L. suebicus* LCV1. In addition to generating fundamental knowledge on anaerobic metabolism in natural environments of galacturonate, a ubiquitous monomer in plant biomass, such studies are likely to have a considerable industrial relevance. In contrast to the canonical isomerase pathway (79), the integrated isomerase-PK pathway demonstrated in this study enables redox-cofactor neutral conversion of D-galacturonate to a ribulose-5-phosphate and $CO_2$ (Figure 6B). Linking these reactions to the non-oxidative pentose-phosphate pathway would pave the way for new metabolic engineering strategies for high-yield, anaerobic conversion of galacturonate-containing feedstocks to compounds such as ethanol, isobutanol or lactate.

## Material and Methods

**Growth media.** Unless stated otherwise, liquid medium contained (g $L^{-1}$): D-galacturonate 4.3; $NH_4Cl$ 1.34; $KH_2PO_4$ 0.78; $Na_2SO_4.10H_2O$ 0.130; $MgCl_2.6H_2O$ 0.120; FeSO4. 7H2O 0.0031; $CaCl_2$ 0.0006; $H_3BO_4$ 0.0001; $Na_2MoO_4.2H_2O$ 0.0001; $ZnSO_4.7H2O$ 0.0032; $CoCl_2.H_2O$ 0.0006; $CuCl_2.2H_2O$ 0.0022; $MnCl_2.4H_2O$ 0.0025; $NiCl_2.6H_2O$ 0.0005; EDTA 0.10, tryptone (BD Bacto Difco, ThermoFisher Scientific, USA) 0.6, yeast extract (BD Bacto Difco, ThermoFisher Scientific, USA) 0.4 and vitamin solution (80) 1 mL $L^{-1}$. 19 L of mineral solution with yeast extract (concentration adjusted to the final volume, 20 L) was

autoclaved for 20 min at 121°C. Subsequently, 86 g D-galacturonate, 8 g tryptone and 20 mL vitamin solution in 1 L demineralized water was filter sterilized (0.2 µm Mediakap Plus, Spectrum Laboratories, Rancho Dominguez, USA) and added to the autoclaved medium. For bioreactor cultivation, 0.075 mL $L^{-1}$ Pluronic PE 6100 antifoam (BASF, Ludwigshafen, Germany) was added to the mineral medium before autoclaving.

**Isolation and maintenance.** Anaerobic 50-mL shake flasks containing 30 mL liquid medium were inoculated with 2 % (v/v) of a mixed inoculum consisting of rumen content of a grass-fed cow, provided by an artisanal butcher (Slager Jonkers, Est, The Netherlands), rotting orange peels and orange-peel-enriched compost. Triplicate shake flask cultures were grown at 30 °C in an anaerobic chamber (Bactron III, Shell Lab, Cornelius, USA, gas composition 89 % $N_2$, 6% $CO_2$ and 5% $H_2$) and transferred once, upon depletion of the carbon source. After the second enrichment phase, the enrichment cultures were plated on 1 % agar plates (4 g $L^{-1}$ D-galacturonate, 0.4 g $L^{-1}$ yeast extract, pH 5) under anaerobic conditions. Single colonies were re-streaked thrice and grown in shake flasks on liquid medium. Stocks were stored in 30 % (v/v) glycerol at -80 °C.

**Bioreactor cultivation.** Batch and chemostat cultures were grown in 1.2 L laboratory bioreactors (Applikon, Delft, The Netherlands) at a temperature of 30 °C. Reactors were stirred at 300 rpm and, to maintain anaerobiosis, sparged with nitrogen gas at a flow rate of 120 mL $min^{-1}$. Culture pH was controlled at pH 4.0 ± 0.1 by automatic titration (ADI 1030 Biocontroller, Applikon, Delft, The Netherlands) with 1 M NaOH. For chemostat cultivation, the dilution rate was set at 0.13 ± 0.01 $h^{-1}$ and a working volume of 0.50 L was maintained by a peristaltic effluent pump (Masterflex, Cole-Parmer, Vernon Hills, USA) coupled to a level sensor.

**Metabolite analysis in culture supernatants.** Cell-free supernatants were obtained by centrifugation (Heraeus Pico Microfuge, ThermoFisher Scientific, Waltman, USA) of culture samples. Concentrations of galacturonate and extracellular metabolites were analysed on an Agilent 1100 Affinity HPLC (Agilent Technologies, Santa Clara, USA) equipped with an Aminex HPX-87H ion-exchange column (BioRad, Hercules, USA), operated at 60 °C with a mobile phase of 5 mM H2SO4 and at a flow rate of 0.6 mL $min^{-1}$. Concentrations of $CO_2$ and $O_2$ in bioreactor exhaust gas were measured with a Prima BT Bench Top mass spectrometer MS (Thermo Scientific Fisher, Waltman, USA) after cooling the gas with a condenser (4 °C).

**Biomass dry weight.** 20 mL of culture broth samples were filtered over pre-dried and pre-weighed membrane filters (0.2 µm Supor-200, Pall Corporation, New York, USA), washed with demineralized water, dried in a microwave oven (Robert Bosch GmbH, Gerlingen, Germany) for 20 min at 360 W and reweighed. Carbon and electron balances were constructed based on the number of carbon atoms and electrons per mole, with an assumed biomass composition of $CH_{1.8}O_{0.5}N_{0.2}$ (81).

Microbial community analysis. 2-mL samples from triplicate shake-flask enrichment cultures were centrifuged (13000 x g, Microfuge, ThermoFisher Scientific, Waltman, USA) and cell pellets were stored -80 °C until analysis. Genomic DNA was extracted using the UltraClean DNA isolation kit (Qiagen Inc., CA, USA), following the manufacturer's instructions. 16S-rRNA gene sequences in the enrichment cultures were analysed by amplicon sequencing on an Illumina HiSeq sequencer (Novogene Bioinformatics Technology Co., Ltd, Beijing, China). Primers 341F (5'- CCTAYGGGRBGCASCAG - 3') and 805R (5' – GGACTACNNGGGTATCTAAT - 3') were used to generate 250 bp paired end reads and sequences were analysed as described previously (82). Representative sequences for the dominant operational taxonomic unites (OTUs) (>1 %) were submitted for taxonomic analysis in the SILVA database (SINA, version 1.2.11, (41) using default settings. The amplicon sequences of the 16S rRNA gene analysis are shown in the Supplemental Material Table S3.



**Whole-genome sequencing.** A 250-mL sample from a steady-state chemostat culture of *L. suebicus* LCV1 was centrifuged for 10 min at maximum speed (4700 x g at 4 °C, Sorvall Legend XTR ThermoFisher Scientific, Waltman, USA). The cell pellet was washed with TE buffer (pH 8) and stored at -20 °C. DNA was extracted with the Genomic-tip 100/G Kit (Qiagen Inc, CA, USA) according to the manufacturer's protocol except for the addition of 2.6 mg $mL^{-1}$ zymolyase (20T, Amsbio, UK) and 4 mg $mL^{-1}$ lysozyme (Qiagen, Hilden, Germany) to facilitate cell lysis. The amount of extracted DNA was quantified with a Qubit dsDNA BR assay kit (Thermo Fisher Scientific, Waltman, USA) and its quality was assessed with NanoDrop 2000 (Thermo Fisher Scientific, Waltman, USA) and Tapestation 2200 technology (Agilent Technologies, Santa Clara, USA). Sequencing was performed in-house on an Illumina MiSeq Sequencer (Illumina, San Diego, USA), with MiSeq Reagent Kit v3 with 2 x 300 bp read length to obtain a 300 cycle paired-end library with an insert-size of 550 bp using TruSeq PCR-free library preparation yielding 5.92 million reads with a total quantity of 1.78 gigabase sequence. Long-read sequencing was performed with the MinION platform (Oxford Nanopore Technologies, Oxford, UK). The MinION genomic library was prepared using Nanopore 1-D ligation sequencing kit (SQK-LSK108), using 2-3 ug of input genomic DNA fragmented in a Covaris g-Tube (Covaris, Brighton, UK) with the 8-10 kb fragments settings according to manufacturer's instructions. All R9 flow cells (FLO MIN106) were primed with priming buffer and libraries were loaded following manufacturer's instructions. MinKNOW software (Oxford Nanopore) was used for quality control of active pores and for sequencing. Raw files generated by MinKNOW were base called using Albacore (version 1.2.5; Oxford Nanopore). Reads, in fastq format, with minimum length of 1000 bp were extracted, yielding 1.63 Gb sequence with an average read length of 6.28 Kbp.

**De novo assembly.** Base-calling was performed with Albacore v1.2.6 (Oxford Nanopore Technologies, Oxford, United Kingdom), FASTQ files were obtained and filtered on size (> 1000 bp). The long-read genome sequences generated with the MinION platform were de novo assembled using Canu v1.4 (settings: genomesize = 3 m) (83). 2.78 Mbp genome into

three contigs of which the largest with a length of 2.67 Mb while the two smaller with a length of 89 and 28 Kbp Paired-end Illumina library was aligned, using BWA (84), to the assembly and the resulting BAM file (Binary alignment map file) was processed by Pilon (85) for polishing the assembly (for correcting assembly errors), using correction of only SNPs and short indels (–fix bases parameter).

**Genome annotation and analysis.** The assembled genome was uploaded to the automated Microscope platform (86, 87). Manual assessment of pathway annotations was assisted by the MicroCyc (88), KEGG (Kyoto Encyclopaedia of Genes and Genomes; (89)) and SwissProt alignment (BLASTP version 2.2.28+; (90)) databases. The genome sequence of Lactobacillus suebicus LCV1 has been deposited at the NCBI archive with the corresponding BioProject ID PRJNA578870. Common and unique coding sequences in the two strains, with *L. suebicus* DSM5007 annotated with Prokka (91), were identified with OrthoVenn2 (92).



**Proteome analysis.** A published protocol (93) was modified to prepare whole protein extracts. Approximately 20 mg biomass (wet weight) were collected and solubilised in a solution consisting of 175 µL B-PER reagent (Illumina, San Diego, USA) and 175 µL triethylammonium bicarbonate (TEAB) buffer (50 mM TEAB, 1 % (w/w) sodium deoxycholate (NaDOC), adjusted to pH 8.0). After addition of 0.1 g glass beads (acid washed, 0.1 mm diameter), cells were disrupted by bead beating with a Fast Prep FP120 (MP Biomedicals, Fisher Scientific, Hampton, USA; 4 bursts of 20 s at a setting of 6 m s-1). Cell debris was removed by centrifugation (15 min at 14,000 x g) and at 4 °C in a 5424R centrifuge (Eppendorf AG, Hamburg, Germany). The supernatant was transferred to a new Eppendorf tube and kept at 4°C until further processing. Proteins were precipitated by adding 4 volumes of ice-cold acetone. The solution was incubated at -20 °C for 30 min and centrifuged 15 min at 14,000 x g) and at 4 °C in a 5424R centrifuge (Eppendorf AG, Hamburg, Germany). The protein pellet was washed twice with 200 µL ice-cold acetone and re-dissolved in 200 mM ammonium bicarbonate containing 8 M urea, to a final concentration of approximately

100 µg protein $\mu L^{-1}$. 30 µL of a 10 mM dithiothreitol solution was added to 100 µL of the resulting protein solution. After 1 h incubation at 37 °C, 30 µL of a freshly prepared 20 mM 3-idoleacetic acid (IAA) solution was added, followed by 30 min incubation in the dark at room temperature. The solution was then diluted to below 1 M urea with a 200 mM bicarbonate buffer. An aliquot of approximately 25 µg protein was digested overnight at 37 °C using sequencing grade trypsin (Promega, Madison, USA), at a trypsin-to-protein ratio of approximately 1 : 50. Peptides were desalted using an Oasis HLB 96 well plate (Waters, Mildford, USA) according to the manufacturer's protocol. The purified peptide eluate was dried with a speed vac concentrator (Thermo Fisher Scientific, Waltman, USA). The dried peptide fraction was resuspended in 3 % acetonitrile and 0.1 % formic acid solution by careful vortexing. An aliquot corresponding to approximately 50-100 ng protein digest was analysed using a one-dimensional shot-gun proteomics (EASY nano LC

connected with a QE plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltman, USA)) (94). Raw mass spectrometry data were analysed using PEAKS Studio X (Bioinformatics Solutions Inc., http://www.bioinfor.com) by searching against a global Lactobacillus suebicus database downloaded from UniProtKB (July 2018). The search included a GPM crap contaminant database (https://www.thegpm.org/crap/) and used a decoy fusion for determining false discovery rates. Search parameters included 20 ppm parent ion and 0.02 Da fragment ion mass error tolerance, up to 3 missed cleavage sites, carbamidomethylation as fixed and methionine oxidation and N/Q deamidation as variable modifications. Peptide spectrum matches were filtered against 1 % false discovery rate (FDR) and protein identifications were accepted as being significant when at least 2 unique peptides were identified. The mass spectrometry proteomics data have been deposited to the Proteome Xchange Consortium via the PRIDE (95) partner repository with the dataset identifier PXD015964**.**



**Discontinuous enzyme assay**. Cell extract of galacturonate-grown *L. suebicus* LCV1 (5.1 ± 0.4 g $L^{-1}$) was incubated for 30 or 180 min at 30 °C in a reaction mixture containing triethanolamine (TEA) buffer (100 mM, pH 7.6), 2.5 mM $MgCl_2.6H_2O$, 5 mM ATP and 5 mM mannonate or no ATP and with 5 mM 6-phosphogluconate, unless stated otherwise. Samples were centrifuged (13,000 x g, Microfuge, ThermoFisher Scientific, Waltman, USA) after incubation and the supernatant was collected. 100 µL of the supernatant was frozen at -80 °C (U101 Innova freezer, Eppendorf, Hamburg, Germany) and freeze-dried over-night (Mini Lyotrap freeze-dryer, LTE Scientific Ltd, Greenfield, Oldham, UK). The supernatant was derivatized according to (96) without addition of AAL-mix and analysed using a 7890A gas chromatography system (Agilent, Santa Clara, CA, USA) coupled to a 5975C MSD single quadrupole mass spectrometer (Agilent, Santa Clara, CA, USA) according to (97); split ratio of 1 : 50 for standards, split ratio of 1 : 10 for samples). Identification of the peaks was done via MassHunter Workstation Qualitative Analysis software (Agilent, version B06.00) and comparison to the NIST Standard Reference Database (version 2.0).

**Enzyme-activity assays in cell extracts.** Cells were harvested from steady-state chemostat cultures by centrifugation (5 min at 4696 x g and at 4 °C, Sorvall Legend X1R, ThermoFisher Scientific, USA) and the pellet was stored at -20 °C. Cells were washed and suspended in 100 mM potassium-phosphate buffer (pH 7.5) with 2 mM $MgCl_2$ and protease inhibitors (cOmpleteTM Protease inhibitor cocktail, Merck group, Darmstad, Germany). For lactate dehydrogenase activity assays, 1 mM dithiothreitol was added instead of protease inhibitors. After cell disruption with a Fast Prep FP120 (MP Biomedicals, Fisher Scientific, Hampton, USA; 4 bursts of 20 s at a setting of 6 m $s^{-1}$), cell debris was removed by centrifugation (20 min at 4 °C and at 20,000 rpm (rotor SS-34) Sorval RC 5C plus) and the resulting cell extract was directly used for enzyme assays. Protein concentrations were determined with the Lowry assay (98).

Activity of lactate dehydrogenase (LDH) was measured according to Van Maris et al., (2004), fructuronate reductase (UxuB) according to Linster and Schaftingen, (2004) and pyruvate dehydrogenase (PDH) according to Flikweert et al., (1997) in an anaerobic environment (Vinyl Anaerobic Chamber; gas phase 2.3 % H2 and 97.7% N2; Coy Lab, USA)) with an AvaSpec-3648 high-resolution spectrometer with AvaLight-DH-S light source (Avantes, Apeldorn, The Netherlands). Aerobic enzyme-activity assays were measured at 340 nm and 30 $^{o}$C with a Hitachi model U-3010 spectrophotometer (Sysmex Europe GmbH, Norderstedt, Germany). Fructuronate reductase (UxuB) was measured in a reaction mixture containing triethanolamine (TEA) buffer (100 mM, pH 7.6), 2.5 mM $MgCl_2.6H_2O$, 0.6 mM NADPH and cell extract of galacturonate-grown *L. suebicus* LCV1 (5.1 ± 0.4 g protein $L^{-1}$). The reaction was stated with the addition of tagaturonate (0.5 mM) (101). 6-phosphogluconate dehydrogenase (GndA) was measured as described previously (102), but with a 50 mM potassium-phosphate buffer (pH 7.5). The reaction was started by addition of 6-phosphogluconate (5 mM). Gluconate kinase was measured via a coupled reaction with 6-phoshogluconate dehydrogenase (0.8 U, Saccharomyces cerevisiae, Sigma Aldrich, St. Louis, USA). The reaction mixture contained TEA buffer (100 mM, pH 7.6), 2.5 mM $MgCl_2.6H_2O$, 1 mM ATP, 0.15 mM NADP+ and the reaction was started by addition of gluconate (0.5 mM, unless stated otherwise). The same assay was used for estimation of mannonate-kinase activities, by starting the reaction with mannonate (0.2 mM unless stated otherwise) instead of gluconate and using native *L. suebicus* 6-mannonate 2-epimerase activity in cell extracts as coupling enzyme. All assays were performed in duplicate and reaction rates were proportional to the amount of cell extract added.



## Acknowledgements

This research was supported by the SIAM Gravitation Grant 024.002.002, the Netherlands Organization for Scientific Research. The authors thank B. Abbas for his help with the 16S rRNA-gene amplicon sequencing, P. de la Torre-Cortés for whole-genome sequencing, and S.L. de Jong and M. Wissink for their contributions to the experimental work as part of their MSc and BSc thesis projects, respectively. The authors declare that the research was conducted in the absence of commercial or financial relationships that could be construed as a potential conflict of interest.

### Author contribution statement

ML, JP and LV designed the experiments, interpreted the results and wrote the manuscript. LV performed the enrichment, isolation and bioreactor experiments. LV and MB analysed the *L. suebicus* genome. ML and LV performed the continuous and discontinuous enzymatic assays. MP and CR performed and interpreted MS studies on proteome and metabolites, respectively. All authors read and approved of the final manuscript.

# Supplemental material

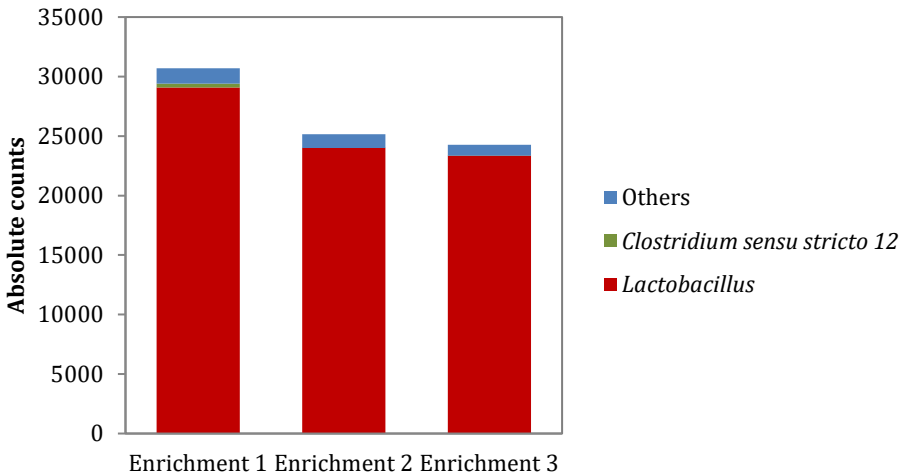





**Figure S10 |** Absolute counts of 16S-rRNA gene amplicon analysis of three independent anaerobic enrichments on D-galacturonate (pH 4), supplemented with 0.1 g $L^{-1}$ yeast extract. Enrichment cultures were inoculated with a mixture of rotting orange peels, cow-rumen content and orange-peel-enriched compost. Red bars represent *Lactobacillus* genus, green bar represents the *Clostridium* sensu stricto 12 genus and blue bars represent the combined operational taxonomic units (OTUs) with an abundance below 1 % (others).

**Table S7 |** Anaerobic growth and catabolic products of *Lactobacillus suebicus* LCV1 on different carbon sources in synthetic media with 4 g $L^{-1}$ carbon source, supplemented with 0.4 g $L^{-1}$ yeast extract. All experiments were performed in an anaerobic chamber (gas phase 5 % $CO_2$, 5 % $H_2$, and 90 % $N_2$) at 25 °C in shake flasks at pH 4. Growth was measured as optical density at 660 nm upon carbon depletion, along with concentrations of catabolic products. Shake flasks were incubated for 12 d. Standard errors were derived from duplicate experiments. N.d.: 'not detected'.

| | **Optical Density (660 nm)** | **Lactate (mM)** | **Acetate (mM)** | **Ethanol (mM)** |
|---|---|---|---|---|
| Galacturonate | 0.68 ± 0.01 | 18.0 ± 0.1 | 19.5 ± 0.0 | 3.8 ± 0.7 |
| Glucose | n.d. | n.d. | n.d. | n.d. |
| Arabinose | 0.94 ± 0.00 | 23.2 ± 0.1 | 23.1 ± 0.0 | 6.2 ± 0.1 |
| Gluconate | 0.80 ± 0.07 | 17.3 ± 0.3 | 7.8 ± 0.2 | 22.5 ± 0.2 |
| Glucuronate | 0.80 ± 0.02 | 18.6 ± 0.0 | 20.2 ± 0.1 | 7.0 ± 0.2 |
| Pyruvate | n.d. | n.d. | n.d. | n.d. |

**Table S8 |** Statistical data for the assembled and annotated genome sequence of *Lactobacillus suebicus* LCV1 and *Lactobacillus suebicus* DSM 5007(44).

| | ***Lactobacillus suebicus* LCV1** | ***Lactobacillus suebicus* DSM 5007** |
|---|---|---|
| **Genome size (Mbp)** | 2.8 | 2.6 |
| **Scaffolds** | 3 | 143 |
| **Contigs** | 3 | 143 |
| **Contigs N50 (bp)** | 2673450 | 54666 |
| **Max contig size (bp)** | 2673450 | 151146 |
| **GC content (%)** | 39.1 | 39.0 |
| **Protein coding density (%)** | 85.5 | n.d. |
| **Coding sequences (CDS)** | 2811 | 2543 |
| **Unique CDS** | 149 | 8 |
| **rRNA copies** | 6 | 3 |



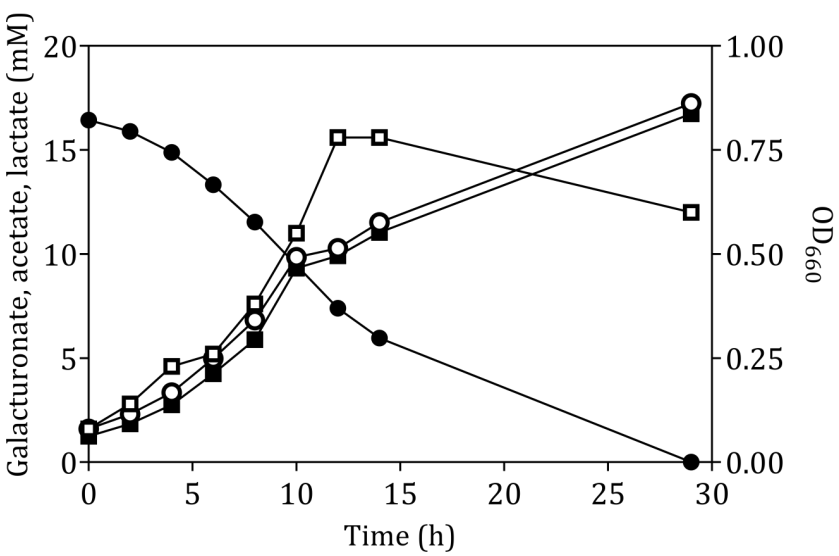



**Figure S2 |** Anaerobic growth and product formation of an anaerobic bioreactor batch culture of *L. suebicus* LCV1 on D-galacturonate (3.3 g $L^{-1}$) at pH 4 and at 30 °C. Symbols: ● D-galacturonate, ■ acetate, □ lactate and ○ optical density. The data shows one of two independent replicates, bioreactor 2.

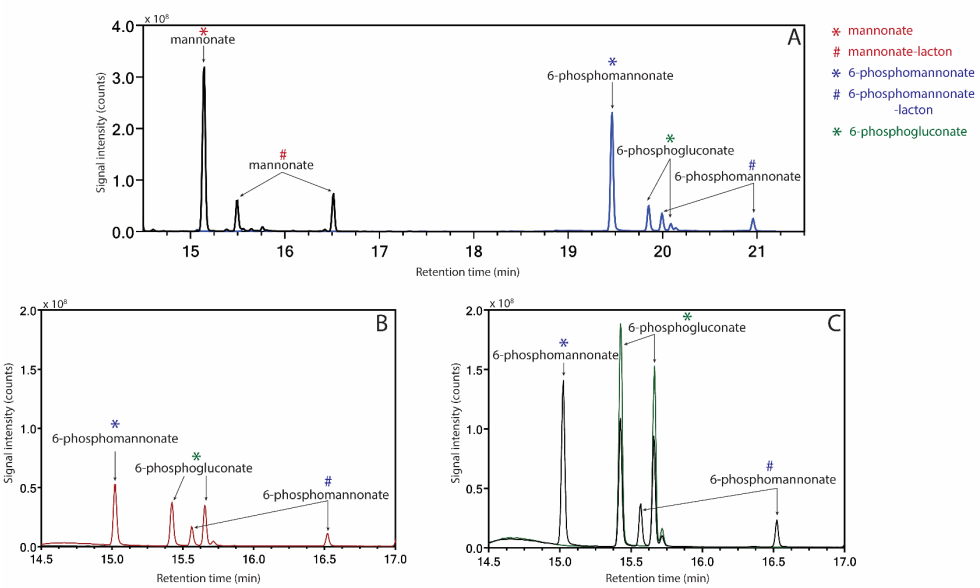



**Figure S3 |** Identification of products from the conversion of mannonate or 6-phosphogluconate in the presence and absence of ATP with cell extract from galacturonate-grown *L. suebicus* LCV1 (D=0.13 $h^{-1}$, T= 30 ºC and pH 4). Panel A) 0.5 h incubation, mannonate with cell extract and no ATP (black line) and ATP (blue line), B) mannonate with cell free extract and ATP, start of the incubation (black line) and 3h incubation (red line), C) 6-phosphogluconate with cell extract start of the incubation (green line) and 3h incubation (black line). One of two of independent biological duplicate experiments are shown. The retention times of the compounds differ between panel A and B, C due to shortening of the column for maintenance purposes.

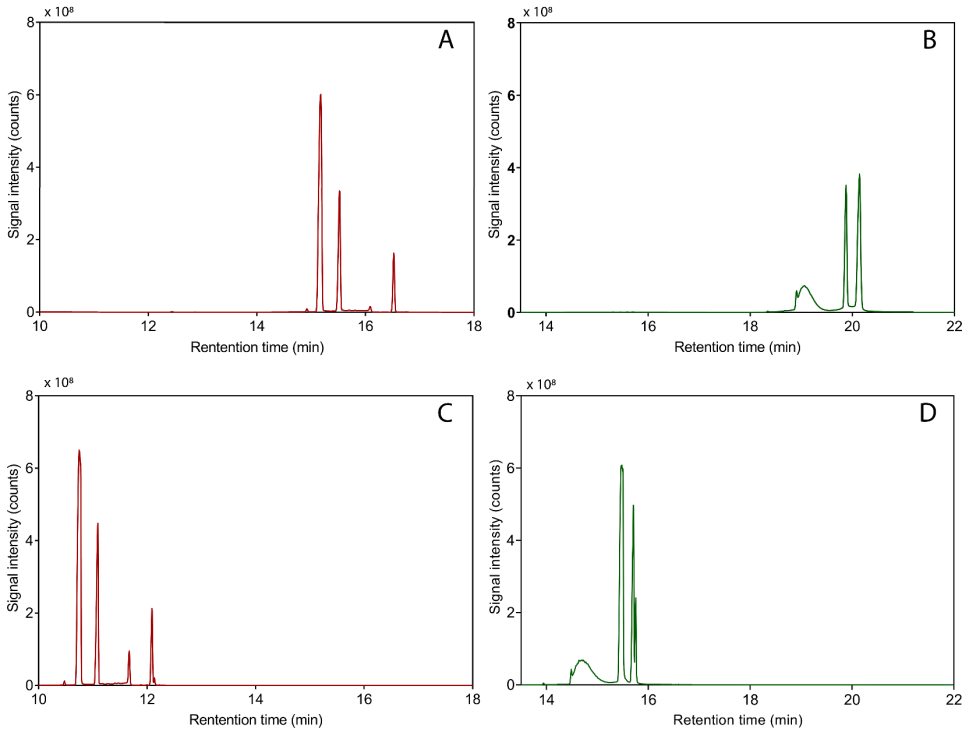



**Figure S4 |** Signal intensities in counts and retention times in minutes of the standards used for the GC-MS analysis before and after reduction of the column. A) mannonate standard before reduction. B) 6-phosphogluconate standard before reduction. C) mannonate standard after reduction. D) 6-phosphogluconate standard after reduction.

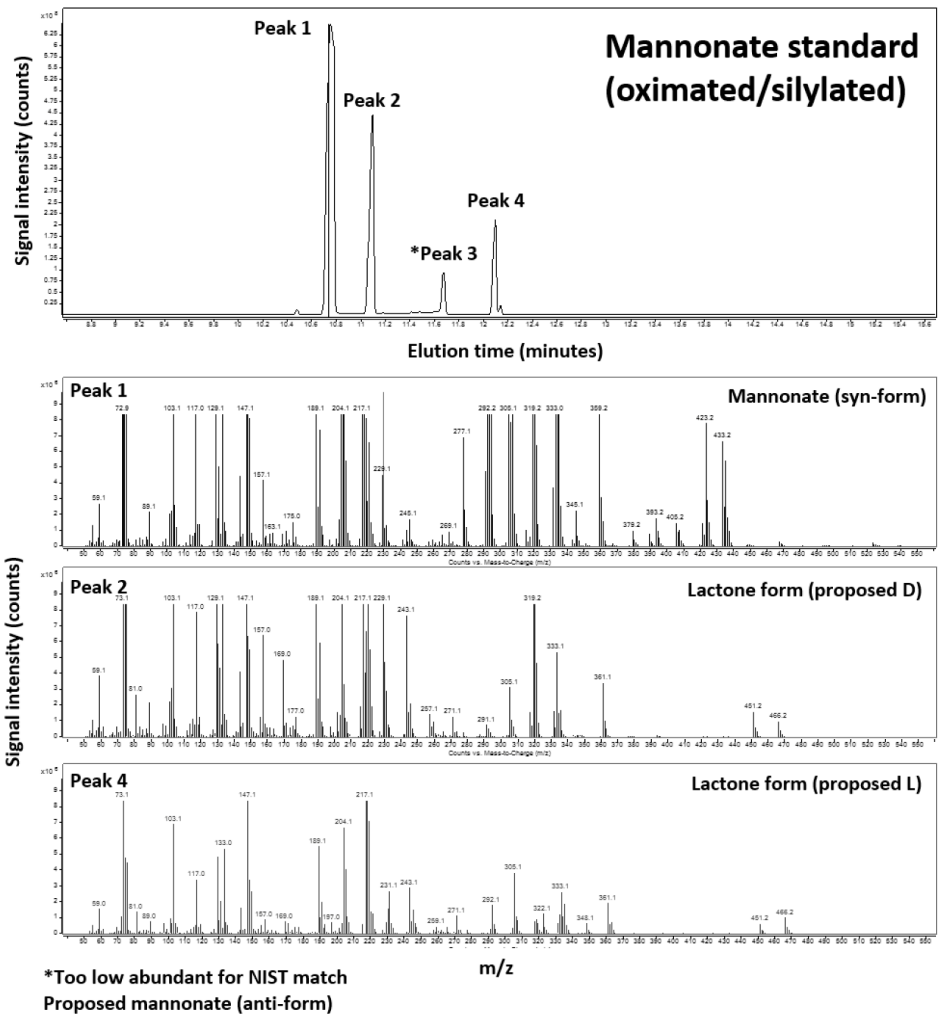



**Figure S5 |** The annotated GC-MS profile of mannonate with the signal intensity in counts and the mass spectrometry profiles for each peak in m/z.

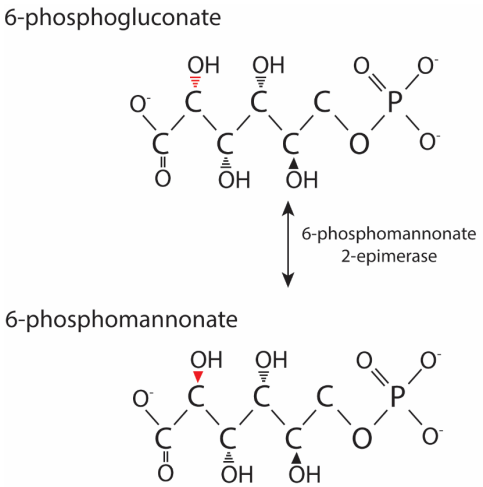



**Figure S6 |** The proposed reaction catalysed by 6-phosphomannonate 2-epimerase, with the affected hydroxyl moiety depicted in red.

**Table S3** | 16S rRNA gene amplicon sequences of the genera shown in Figure S1.

| Genus | Amplicon sequence |
|---|---|
| *Lactobacillus* | TAGGGAATCTTCCACAATGGGCGCAAGCCTGATGGAGCAACACCGCGTGAGTGAAG AAGGGTTTCGGCTCGTAAAACTCTGTTGTTAAAGAAGAACGTATCTAAGAGTAACT GCTTAGGTAGTGACGGTATTTAACCAGAAAGCCACGGCTAACTACGTGCCAGCAGC CGCGGTAATACGTAGGTGGCAAGCGTTATCCGGATTTATTGGGCGTAAAGCGAGCG CAGGCGGTTttttAAGTCTGATGTGAAAGCCTTCGGCTTAACCGAAGAAGTGCATCGG AAACTGGGAGACTTGAGTGCAGAAGAGGACAGTGGAACTCCATGTGTAGCGGTGAA ATGCGTAGATatatGGAAGAACACCAGTGGCGAAGGCGGCTGTCTGGTCTGTAACTG ACGCTGAGGCTCGAAAGCATGGGTAGCGAACAGG |
| *Clostridium* sensu stricto 12 | TGGGGAATATTGCACAATGGGCGAAAGCCTGATGCAGCAACGCCGCGTGAGTGATG ACGGTCTTCGGATTGTAAAGCTCTGTCTTTTGGGACGATAATGACGGTACCAAAGG AGGAAGCCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTGGCAAGCG TTGTCCGGATTTACTGGGCGTAAAGGATGTGTAGGCGGATACTTAAGTGAGATGTG AAAGCCCCGAGCTTAACTTGGGGACTGCATTTCAAACTGGGTGTCTAGAGTGCAGG AGAGGAAAGCGGAATTCCTAGTGTAGCGGTGAAATGCGTAGAGATTAGGAAGAAC ATCAGTGGCGAAGGCGGCTTTCTGGACTGTAACTGACGCTGAGGCATGAAAGCGTG GGGAGCAAACAGG |

# References


1. Khanh NQ, Ruttkowski E, Leidinger K, Albrecht H, Gottschalk M. 1991. Characterization and expression of a genomic pectin methyl esterase-encoding gene in *Aspergillus niger*. Gene 106:71–77.
2. Searle-van Leeuwen MJF, Vincken JP, Beldman G. 1996. Acetyl esterases of Aspergillus niger: purification and mode of action on pectinsProc. Int. Symp. on Progress in Biotechnology 14: Pectins and pectinases. Elsevier, Amsterdam.
3. Mohnen D. 2008. Pectin structure and biosynthesis. Curr Opin Plant Biol 11:266–277.
4. Biz A, Farias FC, Motter FA, Paula DH De, Richard P, Krieger N, Mitchell DA. 2014. Pectinase activity determination: An early deceleration in the release of reducing sugars throws a spanner in the works! PLoS One 9:e109529.
5. de Vries RP, Visser J. 2001. *Aspergillus* enzymes involved in degradation of plant cell wall polysaccharides. Microbiol Mol Biol Rev 65:497–522.
6. Marzo C, Diaz AB, Caro I, Blandino A. 2019. Status and perspectives in bioethanol production from sugar beet, p. 158–161. *In* Bioethanol Production from Food Crops. Elsevier Inc., Cadiz.
7. Food and Agriculture Organization of the United Nations. 2016. Crops, Production quantities of sugar beet, citrus fruits and apples in the World. FAOSTAT.
8. Edwards MC, Doran-Peterson J. 2012. Pectin-rich biomass as feedstock for fuel ethanol production. Appl Microbiol Biotechnol 95:565–575.
9. Zajic JE. 1959. Hexuronic dehydrogenase of *Agrobacterium tumefaciens*. J Bacteriol 78:734–735.
10. Chang YF, Feingold DS. 1970. D-Glucaric acid and galactaric acid catabolism by *Agrobacterium tumefaciens*. J Bacteriol 102:85–96.
11. Kuorelahti S, Kalkkinen N, Penttilä M, Londesborough J, Richard P. 2005. Identification in the mold *Hypocrea jecorina* of the first fungal D-galacturonic acid reductase. Biochemistry 44:11234–11240.
12. Martens-Uzunova ES, Schaap PJ. 2008. An evolutionary conserved D-galacturonic acid metabolic pathway operates across filamentous fungi capable of pectin degradation. Fungal Genet Biol 45:1449–1457.
13. Zhang L, Thiewes H, van Kan J a L. 2011. The D-galacturonic acid catabolic pathway in *Botrytis cinerea*. Fungal Genet Biol 48:990–997.
14. Ashwell G, Wahba AJ, Hickman J. 1960. Uronic acid metabolism in bacteria I Purification and properties of uronic acid isomerase in *Escherichia coli*. J Biol Chem 235:1559–1565.
15. Hickman J, Ashwell G. 1960. Uronic acid metabolism in bacteria II Purification and properties of D-altronic acid and D-mannonic acid dehydrogenases in *Escherichia coli*. J Biol Chem 235:1566–1570.
16. Smiley JD, Ashwell G. 1960. Uronic acid metabolism in bacteria III purification and properties of D-altronic acid and D-mannonic acid dehydrases in *Escherichia coli*. J Biol Chem 235:1571–1575.
17. Cynkin MA, Ashwell G. 1960. Uronic acid metabolism in Bacteria IV Purification and properties of 2-keto-3-deoxy-D-gluconokinase in *Eschericia coli*. J Biol Chem 235:1576–1579.
18. Kovachevich R, Wood WA. 1955. Carbohydrate metabolism by *Pseudomonas fluorescens* IV purification and properties of 2-keto-3-deoxy-6-phospho-gluconate aldolase. J Biol Chem 213:757–767.
19. Peekhaus N, Conway T. 1998. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. J Bacteriol 180:3495–3502.

20. Grohmann K, Baldwin EA, Buslig BS. 1994. Fermentation of galacturonic acid and other sugars in orange peel hydrolysates by the ethanologenic strain of *Escherichia coli*. Biotechnol Lett 16:281–286.
21. Grohmann K, Manthey JA, Cameron RG, Buslig BS. 1998. Fermentation of galacturonic acid and pectin-rich materials to ethanol by genetically modified strains of *Erwinia*. Biotechnol Lett 20:195–200.
22. Doran JB, Cripe J, Sutton M, Foster B. 2000. Fermentations of pectin-rich biomass with recombinant bacteria to produce fuel ethanol. Appl Biochem Biotechnol 84–86:141–153.
23. Kuivanen J, Biz A, Richard P. 2019. Microbial hexuronate catabolism in biotechnology. AMB Express 9:1–11.
24. Valk LC, Frank J, de la Torre-Cortes P, van 't Hof M, van Maris AJA, Pronk JT, van Loosdrecht MCM. 2018. Galacturonate metabolism in anaerobic chemostat enrichment cultures: combined fermentation and acetogenesis by the dominant sp. nov. "*Candidatus* Galacturonibacter soehngenii." Appl Environ Microbiol AEM.01370-18.
25. Link KP, Niemann C. 1930. The action of weak mineral acids on uronic acids. J Am Chem Soc 52:2474–2480.
26. Ruff O. 1898. Über die Verwandlung der D-gluconsäure in D-arabinose. Berichte der Dtsch Chem Gesellschaft 31:1573–1577.
27. McKinnis RB. 1928. An investigation of the hypothetical combined pentose and the so-called free pentose with inferences on the composition of pectin. J Am Chem Soc 50:1911–1915.
28. Gulland JM. 1943. Some aspects of the chemistry of nucleotides. J Chem Soc 208–217.
29. Hirst EL. 1942. Recent progress in the chemistry of the pectic materials and plant gums. J Chem Soc 70–78.
30. Fan DF, Feingold DS. 1972. Biosynthesis of uridine diphosphate D-xylose V. UDP D-glucuronate and UDP D-galacturonate carboxy-lyase of *Ampullariella digitata*. Arch Biochem Biophys 148:576–580.
31. Marounek M, Dušková D. 1999. Metabolism of pectin in rumen bacteria *Butyrivibrio fibrisolvens* and *Prevotella ruminicola*. Lett Appl Microbiol 29:429–433.
32. Rodionova IA, Scott DA, Grishin NV., Osterman AL, Rodionov DA. 2012. Tagaturonate-fructuronate epimerase UxaE, a novel enzyme in the hexuronate catabolic network in *Thermotoga maritima*. Environ Microbiol 14:2920–2934.
33. Kleerebezem R, van Loosdrecht MCM. 2007. Mixed culture biotechnology for bioenergy production. Curr Opin Biotechnol 18:207–212.
34. Johnson K, Jiang Y, Kleerebezem R, Muyzer G, van Loosdrecht MCM. 2009. Enrichment of a mixed bacterial culture with a high polyhydroxyalkanoate storage capacity. Biomacromolecules 10:670–676.
35. Zoetemeyer RJ, Arnoldy P, Cohen A, Boelhouwer C. 1982. Influence of temperature on the anaerobic acidification of glucose in a mixed culture forming part of a two-stage digestion process. Water Res 16:313–321.
36. Grimmler C, Janssen H, Krauße D, Fischer RJ, Bahl H, Dürre P, Liebl W, Ehrenreich A. 2011. Genome-wide gene expression analysis of the switch between acidogenesis and solventogenesis in continuous cultures of *Clostridium acetobutylicum*. J Mol Microbiol Biotechnol 20:1–15.
37. Bahl H, Andersch W, Braun K, Gottschalk G. 1982. Effect of pH and butyrate concentration on the production of acetone and butanol by *Clostridium acetobutylicum* grown in continuous culture. Eur J Appl Microbiol Biotechnol 14:17–20.
38. Salmond C V., Kroll RG, Booth IR. 1984. The effect of food preservatives on pH homeostasis in *Escherichia coli*. J Gen Microbiol 130:2845–2850.

39. Eklund T. 1983. The antimicrobial effect of dissociated and undissociated sorbic acid at different pH levels. J Appl Bacteriol 54:383–389.
40. Voragen AGJ, Coenen GJ, Verhoef RP, Schols HA. 2009. Pectin, a versatile polysaccharide present in plant cell walls. Struct Chem 20:263–275.
41. Pruesse E, Peplies J, Glöckner FO. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics 28:1823–1829.
42. Kleynmans U, Heinzl H, Hammes WP. 1989. *Lactobacillus suebicus* sp. nov., an obligately heterofermentative *Lactobacillus species* isolated from fruit mashes. Syst Appl Microbiol 11:267–271.
43. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res 1043–1055.
44. Nam S-H, Choi S-H, Kang A, Kim D-W, Kim RN, Kim D-S, Kim A, Park H-S. 2011. Genome Sequence of *Lactobacillus suebicus* KCTC 3549. J Bacteriol 193:5532–5533.
45. Gunsalus IC, Gibbs M. 1952. The heterolactic fermentation: II. Position of $C^{14}$ in the products of glucose dissimilation by *Leuconostroc mesenteroides*. J Biol Chem 194:871–875.
46. DeMoss RD, Bard RC, Gunsalus IC. 1951. The mechanism of the heterolactic fermentation: a new route of ethanol formation. J Biol Chem 62:499–511.
47. Busse M, Kindel PK, Gibbs M. 1961. The heterolactic fermentation III position of $C^{14}$ in the products of fructose dissimilation by *Leuconostoc mesenteroides*. J Biol Chem 236:2850–2853.
48. Gibbs M, Sokatch JT, Gunsalus IC. 1955. Product of labeling of glucose-1-$C^{14}$ fermentation by homofermentative and heterofermentative lactic acid bacteria. J Biol Chem 70:572–576.
49. de Vries W, Kapteijn WMC, van der Beek EG, Stouthamer AH. 1970. Molar growth yields and fermentation balances of *Lactobacillus casei* L3 in batch cultures and in continuous cultures. J Gen Microbiol 63:333–345.
50. Thomas TD, Ellwood DC, Longyear VMC. 1979. Change from homo- to heterolactic fermentation by *Streptococcus lactis* resulting from glucose limitation in anaerobic chemostat cultures. J Bacteriol 138:109–117.
51. Izu H, Adachi O, Yamada M. 1996. Purification and characterization of the *Escherichia coli* thermoresistant gluconokinase encoded by the *gntK* gene. FEBS Lett 394:14–16.
52. Tong S, Porco A, Isturiz T, Conway T. 1996. Cloning and molecular genetic characterization of the *Escherichia coli gntR*, *gntK* and *gntU*, the main system for gluconate metabolism. J Bacteriol 178:1584–1590.
53. Zachariou M, Scopes R. 1985. Gluconate kinase from *Zymomonas mobilis*: isolation and characteristics. Biochem Int 10:367–371.
54. Abbott DA, Suir E, Van Maris AJA, Pronk JT. 2008. Physiological and transcriptional responses to high concentrations of lactic acid in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. Appl Environ Microbiol 74:5759–5768.
55. Leyer GJ, Wang LL, Johnson EA. 1995. Acid adaptation of *Escherichia coli* O157:H7 increases survival in acidic foods. Appl Environ Microbiol 61:3752–3755.
56. Zoetemeyer RJ, van den Heuvel JC, Cohen A. 1982. pH influence on acidogenic dissimilation of glucose in an anaerobic digestor. Water Res 16:303–311.
57. Fang HHP, Liu H. 2002. Effect of pH on hydrogen production from glucose by a mixed culture. Bioresour Technol 82:87–93.
58. Horiuchi JI, Shimizu T, Tada K, Kanno T, Kobayashi M. 2002. Selective production of organic acids in anaerobic acid reactor by pH control. Bioresour Technol 82:209–213.
59. Temudo MF, Kleerebezem R, van Loosdrecht M. 2007. Influence of the pH on (open) mixed culture fermentation of glucose : A chemostat study. Biotechnol Bioeng 98:69–

79.
60. Ten Brink B, Konings WN. 1986. Generation of a protonmotive force in anaerobic bacteria by end-product efflux. Methods i Enzymol 125:492–510.
61. Trchounian A, Trchounian K. 2019. Fermentation revisited : How do microorganisms survive under energy-limited conditions ? Trends Biochem Sci 44:391–400.
62. Siezen RJ, Bayjanov JR, Felis GE, van der Sijde MR, Starrenburg M, Molenaar D, Wels M, van Hijum SAFT, van Hylckama Vlieg JET. 2011. Genome-scale diversity and niche adaptation analysis of Lactococcus lactis by comparative genome hybridization using multi-strain arrays. Microb Biotechnol 4:383–402.
63. van Hylckama Vlieg JE, Rademaker JLW, Bachmann H, Molenaar D, Kelly WJ, Siezen RJ. 2006. Natural diversity and adaptive responses of *Lactococcus lactis*. Curr Opin Biotechnol 17:183–190.
64. Wagner N, Tran QH, Richter H, Selzer PM, Unden G. 2005. Pyruvate fermentation by *Oenococcus oeni* and *Leuconostoc mesenteroides* and role of pyruvate dehydrogenase in anaerobic fermentation. Appl Environ Microbiol 71:4966–4971.
65. Snoep JL, de Graef MR, Westphal AH, de Kok A, Teixeira de Mattos MJ, Neijssel OM. 1993. Differences in sensitivity to NADH of purified pyruvate dehydrogenase complexes of *Enterococcus faecalis*, *Lactococcus lactis*, *Azotobacter vinelandii* and *Escherichia coli*: Implications for their activity *in vivo*. FEMS Microbiol Lett 114:279–283.
66. Snoep JL, Westphal AH, Benen JAE, Teixeira de Mattos MJ, Neijssel OM, de Kok A. 1992. Isolation and characterisation of the pyruvate dehydrogenase complex of anaerobically grown Enterococcus faecalis NCTC 775. Eur J Biochem 203:245–250.
67. Hansen RG, Henning U. 1966. Regulation of pyruvate dehydrogenase activity in Escherichia coli K12. Biochim Biophys Acta 122:355–358.
68. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Nelson WC, Ketchum KA, McDonald L, Utterback TR, Malek JA, Linher KD, Garrett MM, Stewart AM, Cotton MD, Pratt MS, Phillips CA, Richardson D, Heidelberg J, Sutton GG, Fleischmann RD, Eisen JA, White O, Salzberg SL, Smith HO, Venter JC, Fraser CM. 1999. Evidence for lateral gene transfer between archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323–329.
69. Schomburg I, Jeske L, Ulbrich M, Placzek S, Chang A, Schomburg D. 2017. The BRENDA enzyme information system–From a database to an expert system. J Biotechnol 261:194–206.
70. Jeske L, Placzek S, Schomburg I, Chang A, Schomburg D. 2019. BRENDA in 2019: A European ELIXIR core data resource. Nucleic Acids Res 47:D542–D549.
71. Cohen SS. 1950. Gluconokinase and oxidative path of glucose-6-phosphate utilization. J Bacteriol 617–629.
72. Van Overtveldt S, Verhaeghe T, Joosten HJ, van den Bergh T, Beerens K, Desmet T. 2015. A structural classification of carbohydrate epimerases: From mechanistic insights to practical applications. Biotechnol Adv 33:1814–1828.
73. Senoura T, Taguchi H, Ito S, Hamada S, Matsui H, Fukiya S, Yokota A, Watanabe J, Wasaki J, Ito S. 2009. Identification of the cellobiose 2-epimerase gene in the genome of *Bacteroides fragilis* NCTC 9343. Biosci Biotechnol Biochem 73:400–406.
74. McNair Scott DB, Cohen SS. 1952. The oxidative pathway of carbohydrate metabolism in *Escherichia coli*. I The isolation and properties of glucose 6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase. Biochem J 55:23–32.
75. McNair Scott DB, Cohen SS. 1956. The oxidative pathway of carbohydrate metabolism in *Escherichia coli*. V. Isolation and identification of ribulose phosphate produced from 6-phosphogluconate by the dehydrogenase of *E. coli*. Biochem J 65:686–689.
76. Portalier RC, Stoeber FR. 1972. La D-mannonate: NAD-oxydoréductase d'*Escherichia*

*coli* K12. Purification, propriétés et individualité. Eur J Biochem 26:290–300.
77. Linster CL, van Schaftingen E. 2004. A spectrophotometric assay of D-glucuronate based on *Escherichia coli* uronate isomerase and mannonate dehydrogenase. Protein Expr Purif 37:352–360.
78. Mandrand-Berthelot M-A, Lagarde AE. 1977. "Hit-and-run" mechanism for D-glucuronate reductase catalyzed by D-mannonate:NAD oxidoreductase of *Escherichia coli*. Biochim Biophys Acta 483:6–23.
79. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT. 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 90:391–418.
80. Verduyn C, Dijken JP van, Postma E, Scheffers WA. 1992. Effect of benzoic acid on metabolic fluxes in yeasts : a continuous-culture study on the regulation of respiration and alcoholic fermentation. Yeast 8:501–517.
81. Roels J. 1983. Energetics and kinetics in biotechnology. Biomedical Press, Amsterdam.
82. van den Berg EM, Elisário MP, Kuenen JG, Kleerebezem R, van Loosdrecht MCM. 2017. Fermentative bacteria influence the competition between denitrifiers and DNRA bacteria. Front Microbiol 8:1–13.
83. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736.
84. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595.
85. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9.
86. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, Mercier J, Renaux A, Rollin J, Rouy Z, Roche D, Scarpelli C, Medigue C. 2017. MicroScope in 2017: An expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res 45:D517–D528.
87. Vallenet D, Labarre L, Rouy Z, Barbe V, Bocs S, Cruveiller S, Lajus A, Pascal G, Scarpelli C, Médigue C. 2006. MaGe: A microbial genome annotation system supported by synteny results. Nucleic Acids Res 34:53–65.
88. Caspi R, Foerster H, Fulcher CA, Kaipa P, Krummenacker M, Latendresse M, Paley S, Rhee SY, Shearer AG, Tissier C, Walk TC, Zhang P, Karp PD. 2008. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 36:623–631.
89. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. 2014. Data, information, knowledge and principle: Back to metabolism in KEGG. Nucleic Acids Res 42:D199–D205.
90. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 25:3389–3402.
91. Seemann T. 2014. Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.
92. Xu L, Dong Z, Fang L, Luo Y, Wei Z, Guo H, Zhang G, Gu YQ, Coleman-Derr D, Xia Q, Wang Y. 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res 47:W52–W58.
93. Hansen SH, Stensballe A, Nielsen PH, Herbst F-A. 2014. Metaproteomics: evaluation of protein extraction from activated sludge. Proteomics 14:2535–2539.
94. Köcher T, Pichler P, Swart R, Mechtler K. 2012. Analysis of protein mixtures from

whole-cell extracts by single-run nanoLC-MS/MS using ultralong gradients. Nat Protoc 7:882–890.
95. Vizcaíno JA, Csordas A, Del-Toro N, Dianes JA, Griss J, Lavidas I, Mayer G, Perez-Riverol Y, Reisinger F, Ternent T, Xu QW, Wang R, Hermjakob H. 2016. 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44:D447–D456.
96. Niedenführ S, ten Pierick A, van Dam PTN, Suarez-Mendez CA, Nöh K, Wahl SA. 2016. Natural isotope correction of MS/MS measurements for metabolomics and $^{13}C$ fluxomics. Biotechnol Bioeng 113:1137–1147.
97. de Jonge LP, Buijs NAA, ten Pierick A, Deshmukh A, Zhao Z, Kiel JAKW, Heijnen JJ, van Gulik WM. 2011. Scale-down of penicillin production in *Penicillium chrysogenum*. Biotechnol J 6:944–958.
98. Lowry OH, Rosenbourgh NJ, Farr AL, Randall RJ. 1951. Protein measurements with the folin phenol reagent. J Biol Chem 193:265–276.
99. van Maris AJA, Winkler AA, Porro D, van Dijken JP, Pronk JT. 2004. Homofermentative lactate production cannot sustain anaerobic growth of engineered *Saccharomyces cerevisiae* : possible consequence of energy-dependent lactate export. Appl Environ Microbiol 70:2898–2905.
100. Flikweert MT, Van Dijken JP, Pronk JT. 1997. Metabolic responses of pyruvate decarboxylase-negative *Saccharomyces cerevisiae* to glucose excess. Appl Environ Microbiol 63:3399–3404.
101. Huisjes EH, Luttik MAH, Almering MJH, Bisschops MMM, Dang DHN, Kleerebezem M, Siezen R, Maris AJA Van, Pronk JT. 2012. Toward pectin fermentation by *Saccharomyces cerevisiae* : Expression of the first two steps of a bacterial pathway for D-galacturonate metabolism. J Biotechnol 162:303–310.
102. Scott WA, Abramsky T. 1973. *Neurospora* 6-phosphogluconate dehydrogenase I. Purification and characterization of the wild type enzyme. J Biol Chem 248:3535–3541.

# Outlook

The research described in this thesis led to the identification of novel ways to anaerobically ferment D-galacturonate, an important monomer of plant biomass, uncovered by applying different process conditions to anaerobic enrichment cultures grown on this carbon source.

The research described in **Chapters 2** and **3**, which was based on enrichment in D-galacturonate-limited cultures, led to stable enrichment cultures in which the dominant microorganisms showed acetogenic growth. From an ecophysiological point of view, acetogenesis can be interpreted as a means to improve ATP yields under energy-limited conditions. Further analysis indicated that acetogenesis occurred via an as yet unknown and unidentified variant of the carbon monoxide dehydrogenase/acetyl-CoA synthase complex (CODH/ACS) of the Wood-Ljungdahl pathway, one of the most important autotrophic $CO_2$-fixing pathways in anaerobic natural environments. Additional research is required to elucidate the reactions and enzymes involved in this pathway variant. Previous research on the canonical Wood-Ljungdahl pathway (1, 2) can guide the way in unravelling this metabolic conundrum. In-vitro enzyme activity assays, enzyme purification, all performed under anaerobic conditions to prevent potential oxygen inactivation of key enzymes, are a logical first step in future research (3, 4). Alternatively, anaerobic gel-based isolation of the responsible complex might be explored (4). Mass-spectrometry based identification of the responsible proteins could then rapidly lead to the responsible structural genes. If possible, generation of a defined co-culture of the *Klebsiella oxytoca* and *candidatus* Galacturonibacter soehngenii, the two predominant organisms in the enrichment cultures described in **Chapters 2** and **3**, could help to simplify such biochemical studies. As a possible alternative to follow-up research with mixed cultures, microorganisms that are closely related to *candidatus* Galacturonibacter soehngenii but can be grown in pure culture, such as *Lachnotalea glycerini* (5) (93 % identity) could be tested for presence of the unknown Wood-Ljungdahl pathway variant. In case of a positive result, such pure cultures could provide a simpler and therefore better experimental platform for elucidation of the unidentified enzymes and reactions.

While the research described in **Chapters 2** and **3** increased fundamental knowledge on D-galacturonate metabolism, it did not solve a major challenge identified at the outset of this PhD project: the limited range of fermentation products enabled by the conventional 'isomerase pathway' for D-galacturonate metabolism (**Chapter 1**). In **Chapter 4**, enrichment cultivation at low pH led to a product stoichiometry that was not compatible with operation of the canonical isomerase pathway. Instead, genomic, proteomic and biochemical analysis of the dominant microorganism, a strain of *Lactobacillus suebicus*, enabled elucidation of a novel pathway, which connected the isomerase pathway and the phosphoketolase pathway.

Although *in vitro* experiments demonstrated activity of key enzymes of the novel pathway, some of the responsible genes remain to be identified. The already identified candidate mannonate kinase gene could be functionally analysed by expressing it in a suitable gluconate kinase-negative host strain (e.g. *Escherichia coli* or *Saccharomyces cerevisiae*), followed by in vitro analysis of mannonate phosphorylation. Alternatively, the mannonate kinase gene could be identified via chromatography of cell extracts of *Lactobacillus suebicus* LCV1, enzyme-activity assays and mass spectrometry. Overexpression and purification of mannonate kinase would then enable its use as a coupling enzyme for the subsequent identification, by a similar approach, of the 6-phosphomannonate epimerase. Alternatively, comparison of the transcriptomes of D-galacturonate and D-gluconate-grown *L. suebicus* LCV1 cultures might provide a fast identification of candidate genes.

The novel pathway described in **Chapter 4** can provide a redox-cofactor-neutral link of D-galacturonate metabolism with the non-oxidative pentose-phosphate pathway. The latter route is already used for industrial production of ethanol by the well-known microbial cell factories *Klebsiella oxytoca* and *Saccharomyces cerevisiae* (6, 7). Many 'second generation' feedstocks for ethanol production contain significant amounts of D-galacturonate (8). After identification of the kinase and epimerase genes, it would therefore be highly interesting to engineer a *Saccharomyces cerevisiae* strain able to covert D-galacturonate to ethanol via enzymes of the novel pathway described in **Chapter 4** and the non-oxidative pentose phosphate pathway.

At the outset of this study, use of a previously proposed but never experimentally demonstrated D-galacturonate decarboxylase pathway (**Chapter 1**) was considered as one of the possible outcomes of enrichment experiments. Since this pathway, which involves a direct decarboxylation of D-galacturonate to L-arabinose, is thermodynamically possible and of potential industrial interest, further research to explore its existence in nature remains relevant. One interesting strategy is to perform enrichment studies in sequential batch cultures at low pH, with alternating batch-cultivation cycles on D-galacturonate and L-arabinose. This strategy, for which only preliminary experiments were done during this PhD project, is based on two characteristics of the elusive decarboxylase pathway. First, its use could obviate the need for production of organic acids, which could provide a selective advantage at low pH. Additionally, it could minimize the need for energetically expensive protein synthesis (9) during transition between D-galacturonate and L-arabinose.

The research described in this thesis underlines the continued importance of enrichment cultivation for unlocking the diversity of microbial metabolism, especially when combined with (meta-)genomic analysis of the resulting mixed populations and newly isolated microorganisms. However, even for exploring the diversity of anaerobic microbial metabolism of a single substrate, D-galacturonate, the research described in this thesis covered only a minute fraction of the conceivable enrichment regimes. This limited

coverage of conceivable experiments is linked to the duration and complexity of many enrichment experiments.

Especially when slow-growing micro-organisms are involved, enrichment cultivation can be a time-consuming process. This point is powerfully illustrated by the recent decade-long enrichment experiment that enabled isolation of a slow-growing Lokiarchaeota-related Asgard archaeon from marine sediment (10). In addition to the duration of some enrichment experiments, exploration of dynamic selection regimes (e.g. feast-famine regimes or alternating cultivation cycles on different carbon sources) requires long-term operation of controlled bioreactor cultures. Investments in dedicated cultivation facilities for automated, multi-parallel exploration of different enrichment strategies, analogous to recently described set-ups for automated laboratory evolution (11), could therefore be highly valuable for accelerating research in this field. Knowledge on microbial metabolism generated in such facilities will contribute to our understanding of microbial metabolism in natural environments subject to change, e.g. as a result of climate change. In addition, it will enable the development of mixed-culture based bioprocesses (12, 13) and provide new options for metabolic engineering of established microbial cell factories used in industrial biotechnology.

## References


1. Ragsdale SW. 2008. Enzymology of the Wood-Ljungdahl pathway of acetogenesis, p. 129–136. *In* Annals of the New York Academy of Sciences.
2. Ragsdale SW, Pierce E. 2008. Acetogenesis and the Wood-Ljungdahl pathway of $CO_2$ fixation. Biochim Biophys Acta - Proteins Proteomics 1784:1873–1898.
3. Ragsdale SW, Wood HG. 1985. Acetate biosynthesis by acetogenic bacteria. Evidence that carbon monoxide dehydrogenase is de condensing enzyme that catalyzes the final steps of the synthesis. J Biol Chem 260:3970–3977.
4. Terlesky KC, Nelson MJK, Ferry JG. 1986. Isolation of an enzyme complex with carbon monoxide dehydrogenase activity containing corrinoid and nickel from acetate-grown *Methanosarcina thermophila*. J Bacteriol 168:1053–1058.
5. Jarzembowska M, Sousa DZ, Beyer F, Zwijnenburg A, Plugge CM, Stams AJM. 2016. *Lachnotalea glycerini* gen. nov., sp. nov., an anaerobe isolated from a nanofiltration unit treating anoxic groundwater. Int J Syst Evol Microbiol 66:774–779.
6. Bothast RJ, Saha BC, Vance Flosenzier A, Ingram LO. 1994. Fermentation of L-arabinose, D-xylose and D-glucose by ethanolegenic recombinant *Klebsiella oxytoca* strain P2. Biotechnol Lett 16:401–406.
7. Verhoeven MD, de Valk SC, Daran JMG, van Maris AJA, Pronk JT. 2018. Fermentation of glucose-xylose-arabinose mixtures by a synthetic consortium of single-sugar-fermenting *Saccharomyces cerevisiae* strains. FEMS Yeast Res 18:1–12.
8. van Maris AJA, Abbott DA, Bellissimi E, van den Brink J, Kuyper M, Luttik MAH, Wisselink HW, Scheffers WA, van Dijken JP, Pronk JT. 2006. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: Current status. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol 90:391–418.
9. Stouthamer AH. 1973. A theoretical study on the amount of ATP required for synthesis of microbial celt material. Antonie Van Leeuwenhoek 39:545–565.
10. Imachi H, Nobu MK, Nakahara N, Morono Y, Ogawara M, Takaki Y, Takano Y, Uematsu K, Ikuta T, Ito M, Matsui Y, Miyazaki M, Murata K, Saito Y, Sakai S, Song C, Tasumi E, Yamanaka Y, Yamaguchi T, Kamagata Y, Tamaki H, Takai K. 2019. Isolation of an archaeon at the prokaryote-eukaryote interface. bioRxiv.
11. Sandberg TE, Salazar MJ, Weng LL, Palsson BO, Feist AM. 2019. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. Metab Eng 56:1–16.
12. González-Cabaleiro R, Lema JM, Rodríguez J. 2015. Metabolic energy-based modelling explains product yielding in anaerobic mixed culture fermentations. PLoS One 10:1–17.
13. Regueira A, González-Cabaleiro R, Ofiţeru ID, Rodríguez J, Lema JM. 2018. Electron bifurcation mechanism and homoacetogenesis explain products yields in mixed culture anaerobic fermentations. Water Res 141:349–356.

# Acknowledgements

The PhD thesis in front of you would not have been possible without the help, advice, assistance and contribution of many people I have come to know over the years. I had the privilege to work in two large research groups, be part of a large institute and work with colleagues abroad. I want to convey my gratitude to many of you in this part of the thesis and I can only hope I will do everyone justice.

Jack, you really inspired me throughout my project with your (seemingly) boundless knowledge of biochemistry, metabolism, energetics and fundamental microbiology. I highly enjoyed our discussions about my research, the science behind it and your enthusiasm for the discoveries we made within the project. Additionally, I want to thank you for your help with all the manuscripts in this thesis, it was paramount for its completion, and I can only aspire to one day achieve similar writing abilities. Next to this I want to express my gratitude you for your support, advice and the space you provided during the period of my mothers' passing.

Mark, thank you for always having the trust that I would finish my thesis, even if that trust evaded me sometimes. I highly enjoyed our meetings, made inspirational by your wide variety of knowledge and creative ideas on experiments and measurement methods. I want to additionally thank you for the many presentation opportunities, for the grand application tips and for the freedom to dive deeply into the metabolisms described in this thesis.

Ton, thank you so much for the trust you had in me when you hired me, and for allowing me to start this epic journey. Although you left after one year to start your adventure in Sweden, I still adhere to your philosophy of Occam's razor or parsimony, which helped a lot during my thesis.

I was very fortunate to have a lot of roommates: Harmen, Matthijs, Angel, Hannes, Jasmijn, Bart Eveline, Simon, Danny, Ana-Maria, and Stefan. Thank you all so much for all the jokes in the office, your patience and compassion when I needed to vent and help and tips on how to best submit a paper or make a poster. I additionally wat to thank you for reminding me when there was cake, important presentation or other events I would have missed otherwise.

At the start of my PhD I was fortunate to be part of the ethanolics group (or was it alcoholics?), Maarten, Ioannis and Jasmine, thank you so much for helping me in the first couple of years, with your advice, comments and ideas! I really enjoyed our Friday-morning sessions and our combined journey to produce as much ethanol as possible,

although you probably succeeded better than I did in the end. When I was the last one left, I was adopted by the ElOxY group, with Sanne, Jonna, Wijb, Aurin and Raul. Each one of you was very supportive and your unbridled enthusiasm for my prokaryote work, even though it was a bit outside your eukaryote-comfort zone, helped me through the later years of my PhD. Thank you so much for allowing me to talk about my work and present an endless stream of strange pathways.

Despite much of their work not ending up in this thesis, I would not have been able to finish any of the chapters without my students; Max, Lotte, Martijn R., Sam, Martijn W., Kelly and Alessandro. Thank you so much for your work, desire to learn and to explore, as you did a lot of explorative work and not a lot was known about the methods or analysis we used. However, your unwavering enthusiasm and our joint learning curve not only taught me a lot but also kept it fun (I hope also for all of you).

The experiments which are part of this thesis would not have been possible without the help of the people who keep the lab running. Marcel, Erik, Rob, Ben, Dirk, Zita, Rhody, Pilar and Marijke thank you all so much for the help on numerous occasions. Additionally,your wealth of knowledge on how to best approach a certain experiment or how to best design an experimental set-up was invaluable for my thesis. I learned a great deal from all of you.

Marijke, thank you so much for all your help with the enzyme assays, especially the work shown in chapter 4, which not only ensured I was able to finish my thesis on time, it also made the work which is now submitted much better! I additionally want to thank you for all enzymatic knowledge you passed on to me and for the talks we had when we both lost someone close and dear to us. Pilar, I really enjoyed our escape room adventures with Bart and Maarten, the Noel Gallagher-concert we went to and the time I crashed on your couch! Thank you so much for your kind words and friendship over the years and of course your exceptional sequencing skills which I had the privilege to use on multiple occasions.

Lieve dames van de MSD, Jannie, Apilena, Astrid, dank jullie wel voor al jullie snelle werk, voor de ontelbare keren dan ik toch last-minute een 20L vat bij de autoclaaf mocht zetten of dat jullie de autoclaaf iets later aanzetten zodat mijn bioreactoren er ook bij konden. Zonder jullie kan er niet gecultiveerd worden en door jullie is de media keuken een fijne geoliede machine.

Martijn, Daan, Sigrid (ISME-buddy) and Sara thank you so much for the cool and fun times we had at the SIAM retreats, symposia and writing retreats! Every one of those occasions is a highlight in my PhD and felt like a small holiday. Martijn, thank you so much for inviting me to Wageningen and our cool work on the CODH. I highly enjoyed working with you and I think we made a very efficient anaerobic enzymatic assay-team at the end of my stay.

Needless to say, Sanne and Ingrid, thank you so much for agreeing to be my paranympfs, for helping me organize this memorable moment in my life and for standing by my side during my defense. Additionally, I want to thank the members of my thesis committee for joining my defense and for reading my thesis.

Xavier thank you so much for all the lunches and tea moments we had, and of course organizing the first IMB lab-retreat and employee meetings. Your care and concern for the people around you and your desire for improvement of your surroundings inspires me a lot. So, in answer your question, you definitely returned the favour! Gerben and Aurin, thank you so much for all the advice you gave me on my results, research and writing. Your boundless enthusiasm for research, microbiology and metabolism made discussing my data and results very fun and educational.

Sanne, Eveline, Maaike and Marissa thank you so much for all the tea and coffee moments and your kind compassion for my experiences during my PhD. Your advice helped me through the tough moments but most of all the 'gezelligheid' provided by the coffee breaks enriched my PhD enormously.

Louise dank je wel voor alle gesprekken die wij hebben gehad, ik denk nog vaak terug aan wat je gezegd hebt.

Lieve Kobalt; Lennaert, Kristel, Teatske, Ingeborg en Mareike, het was altijd superleuk op het water en ik mis de zaterdagochtend en zondagmiddag trainingen nu al! Samen commitment tonen voor elkaar in de wedstrijd was echt tof en ik ben blij dat we hiermee een trend konden beginnen bij DDS.

Philine, Nine, Merel en Lot, dank jullie wel voor jullie vriendschap, steun en verhalen gedurende mijn PhD. Vriendschappen maken het leven leuk en ik voel mij bevoorrecht om elk van jullie al lang te mogen kennen!

Lieve papa en mama, dank jullie wel dat jullie mij altijd gesteund hebben en mij altijd hebben aangemoedigd in het maken van mijn eigen keuzes. Daarnaast hebben jullie er ook voor gezorgd dat ik altijd kritisch en nieuwsgierig blijf, wat de wetenschap leuk maakt. Lieve Sofie, dank je wel dat je er altijd bent, het delen van onze ervaringen in de wetenschap en voor de inspiratie voor stelling nr. 4 natuurlijk. Hoewel mama er niet meer is, gaan we toch samen verder, en zo blijft ze nog een beetje bij ons.

Allerliefste Maarten, het beste van deze thesis is natuurlijk dat ik jou heb leren kennen tijdens het maken ervan. Jij bent mijn rots in de branding en maakt elke plek voor mij een thuis. Samen hebben we al veel lief en leed gedeeld, en ik ben kijk uit naar de avonturen die er voor ons in het verschiet liggen.

# Curriculum vitae

Laura Christina Valk was born in Broek in Waterland on the 9th of November 1989. After completing pre-university education (VWO) at the Bernard Nieuwentijt College in 2008, she enrolled in the Architectural Engineering Bsc program at Delft University of Technology. However, after a year she decided it was not for her and enrolled in 2009 in the Life, Science and Technology (LST) Bsc program jointly taught at Leiden University and Delft University of Technology. Next to her studies Laura was also active at her rowing club D.S.R.V. LAGA where she participated in competitive crew rowing at a national level and the organisation of a summer camp for youths centred around rowing. Her Bsc thesis at the Leiden University at the lab of Prof. dr. R. Dame focussed on gel and isotope-based analysis of the DNA binding proteins of *Sulfolobus solfataricus* and *Escherichia coli*. After completion of her Bsc she started the Cell Factory track within the Msc program of LST at the Delft University of Technology. She conducted her Msc thesis on cultivation and selection for lipid accumulating algae at the group of Prof. dr. ir. M.C.M. van Loosdrecht under the supervision of dr. P.R. Mooij. Laura completed her Msc studies with a three months industrial internship at Chr. Hansen A/S in Hørsholm, Denmark. After graduation she started her PhD project within both the Environmental Biotechnology group and Industrial Microbiology group at the Delft University of Technology. This project, supervised by Prof. dr. J.T. Pronk and Prof. dr. ir. M.C.M. van Loosdrecht focussed on the exploration of the microbial and metabolic diversity of D-galacturonate fermentation in an anaerobic environment. The project was also part of the Soehngen Institute of Anaerobic Microbiology (SIAM), allowing for many collaborations with affiliated Universities in the Netherlands. The results of this research are described in this thesis.

## List of publications

1. **Laura C. Valk**, Jeroen Frank, Pilar de la Torre-Cortés, Max van 't Hof, Antonius J. A. van Maris, Jack T. Pronk, Mark C. M. van Loosdrecht. 'Galacturonate metabolism in anaerobic chemostat enrichment cultures: combined fermentation and acetogenesis by the dominant sp. nov. "*Candidatus* Galacturonibacter soehngenii"'. *Applied and Environmental Microbiology* (2018). 84:e01370-18
2. **Laura C. Valk**, Martijn Diender, Gerben R. Stouten, Jette F. Petersen, Per H. Nielsen, Morten S. Dueholm, Jack T. Pronk and Mark C.M. van Loosdrecht. '"*Candidatus* Galacturonibacter soehngenii" shows an acetogenic catabolism of galacturonic acid but lacks a canonical carbon monoxide dehydrogenase/acetyl-CoA synthase complex'. *Manuscript submitted for publication*.
3. **Laura C. Valk**, Marijke A.H. Luttik, Carol de Ram, Martin Pabst, Marcel van den Broek, Mark C.M. van Loosdrecht and Jack T. Pronk. 'A novel D-galacturonate fermentation pathway in *Lactobacillus suebicus* links initial reactions of the galacturonate-isomerase route with the phosphoketolase pathway'. *Manuscript submitted for publication*.

## Conference contributions

1. SIAM symposium (2016) - *Exploring the natural microbial and metabolic diversity of anaerobic galacturonic acid fermentation.*
2. Microbiology Centennial symposium (2017) - *Physiology of anaerobic microbial enrichment cultures fed with galacturonic acid*.
3. ISME 17 (2018) - *Galacturonate fermentation by anaerobic chemostat enrichment cultures.* (*Poster*)
4. SIAM symposium (2018) – *Metabolic diversity of galacturonate fermenters.*
5. SIAM symposium (2019) – 5 years of SIAM research! - *Unconventional hetero-acetogenesis in a D-galacturonate-limited enrichment culture.*

## Other activities

1. Co-chair *Environmental Microbiology* session - KNVM meeting (2019)
2. SIAM talent grant - 2018

Cover design by Laura Valk
This thesis was printed on 100% recycled paper.
Printed by: ProefschiftMaken || www.proefschriftmaken.nl