Technische Universität München, Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt, Lehrstuhl für Technische Mikrobiologie

Structure, Biosynthesis and Possible Ecological Role of a Novel Heteropolysaccharide from *Kozakia baliensis*

Julia Brandt

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Prof. Dr. Karl-Heinz Engel

Prüfer der Dissertation:

1. Prof. Dr. Rudi F. Vogel
2. Prof. Dr. Wolfgang Liebl

Die Dissertation wurde am 12. Februar 2018 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 25. Mai 2018 angenommen.

VORWORT

Diese Arbeit wurde durch die Gelder des German Federal Ministry for Economic Affairs and Energy via the German Federation of Industrial Research Associations (AiF) (FEI); Projektnummer AiF 18071 N, unterstützt.

Zu allererst möchte ich mich besonders bei meinem Doktorvater Prof. Dr. Rudi F. Vogel bedanken, der mir die Möglichkeit gab, meine Doktorarbeit an seinem Institut, an einem stets spannenden Thema durchführen zu dürfen. Zudem möchte ich mich für seine Betreuung, das entgegengebrachte Vertrauen und die Möglichkeit mich frei zu entfalten, bedanken.
Dieser Dank gebührt ebenfalls meinem Mentor Dr. Frank Jakob. Danke für all deine hilfreichen Hinweise, dein kritisches Auge und die moralische Unterstützung und Motivation während der letzten Jahre.
Mein Dank gilt weiterhin allen Kollegen, allen voran meiner Projektpartnerin Tharalinee Ua-Arak, für die freundliche Aufnahme, die Hilfsbereitschaft und vor allem für das hervorragende Arbeitsklima.

Mein besonderer Dank gilt zudem unseren Kooperationspartnern, Prof. Dr. Mirko Bunzel und Dr. Daniel Wefers, für die chemische Charakterisierung der *K. baliensis* Polysaccharide und der gemeinschaftliche Arbeit an der erschienenen Publikation. Des Weiteren möchte ich mich bei Prof. Dr. Wolfgang Liebl und Dr. David Kostner für die überaus nette Hilfe und Unterstützung bei der Konstruktion der *K. baliensis* Deletionsvektoren bedanken.

Ein großer Dank gilt besonders meinen Freunden, die mir stets zur Seite stehen, immer ein offenes Ohr für alles haben und die mit ihrem Esprit und sprudelnder Lebensfreude mein Leben bereichern. Ein genauso großer Dank gilt dir Dennis, vielen Dank, dass die Zeit mit dir zusammen nie langweilig ist und wir die Fähigkeit haben aus allem etwas Schönes zu machen. Danke, dass du mich immer antreibst und unterstützt. Zuletzt sei insbesondere meinen Eltern gedankt, die mir das Studieren ermöglicht haben und auf die ich mich immer verlassen kann. Ich freu mich sehr, dass ich mein Leben mit euch allen Teilen darf!

"It is hard work and great art to make life not so serious"
(John Irving, The Hotel New Hampshire)

ABBREVIATIONS

A.	*Acetobacter*
aa	amino acid
AAB	acetic acid bacteria
AF4	asymmetric flow field-flow fractionation
BLASTP	Basic Local Alignment Search Tool
bp	base pair
C	carbon
cfu	colony forming units
cm	centimeter
Cys	cysteine
Da	Dalton
DGC	diguanylate cyclase
DNA	deoxyribonucleic acid
D_2O	deuterium oxide
DP	degree of polymerization
DSM	German Collection of Microorganisms and Cell Cultures
DTT	dithiothreitol
E.	*Escherichia*
EDTA	ethylene diamine tetraacetic acid
EPS	exopolysaccharide
Fig.	figure
FTF	fructosyltransferase
ftf	fructosyltransferase gene
g (centrifugation)	relative centrifugal force
g (weight)	gram
G.	*Gluconobacter*
Ga.	*Gluconacetobacter*
Gal	galactose
GH-68	glycoside hydrolase 68 family
Glc	glucose
GlcA	glucuronic acid
Gln	glutamine
GT	glycosyltransferase
gum	protein involved in HePS biosynthesis (derived from *X*. *campestris*)
H	hydrogen
h	hour
HePS	heteropolysaccharide
HGAP	hierarchical genome-assembly process
HoPS	homopolysaccharide
HP	hypothetical protein
HPLC	high-performance liquid chromatography
HPMC	hydroxypropylmethylcellulose
Hz	Hertz
IS	insertion sequence
K	Kelvin
K.	*Kozakia*
Ko.	*Komagataeibacter*
L	liter
L.	*Leuconostoc*
LAB	lactic acid bacteria
LB	Luria-Bertani
LC-MS/MS	liquid chromatography coupled to tandem mass spectrometry
M	molar

m	meter
MALS	multi-angle light scattering
Man	mannose
M_i	molar mass
min	minute
mMRS	modified de Man, Rogosa, Sharp medium
M_N	number average molar mass
M_W	weight average molar mass
MWCO	molecular weight cut off
M_{Wi}	molecular weights
N	Newton
N.	*Neoasaia*
NaG	natrium/sodium-gluconate medium
NBRC	National Institute of Technology and Evaluation (NITE) Biological Resource Center,
NCBI	Japan national center for biotechnology information
n_i	number of molecules
NMR	nuclear magnetic resonance
nt	nucleotide
OD	optical density
ORF	open reading frame
p	probability
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PDI	polydispersity index
pol	protein involved in pellicle formation
ppm	parts per million
R_Gi	radiuses of gyration
RI	refractive index
rms	root mean square
rRNA	ribosomal ribonucleic acid
rpm	revolutions per minute
RT	retention time
s	second
SAP	shrimp alkaline phosphatase
SDS	sodium dodecyl sulfate
si	standard inverse
SMRT	single-molecule real-time sequencing
TCA	tricarboxylic acid
TEMED	tetramethylethylendiamine
T_m	melting temperature
TMW	Technische Mikrobiologie Weihenstephan
U	unit
UDP	uridin diphosphate
V	volt
v_G	hydrodynamic coefficient
v/v	volume/volume
w/v	weight/volume
w/w	weight/weight
Z.	*Zymomonas*
X.	*Xanthomonas*

CONTENT

PUBLICATIONS INCLUDED IN THIS THESIS

Brandt et al. (2016)

Brandt, J. U., Jakob, F., Behr, J., Geissler, A.J., Vogel, R. F. (2016). **Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*.** Microbial Cell Factories, 15 (1), 170.

Brandt et al. (2017)

Brandt, J. U., Jakob, F., Geissler, A.J., Behr, J., Vogel, R. F. (2017). **Multiple Genome Sequences of Heteropolysaccharadie-Forming Acetic Acid Bacteria.** Genome Announcements, 5 (16).

Brandt et al. (2017)

Brandt, J. U., Born, F., Jakob, F., Vogel, R.F. (2017). **Environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in *Kozakia baliensis* NBRC 16680.** BMC Microbiology, 17 (172).

Brandt et al. (2017)

Brandt, J. U., Welfers, D., Jakob, F., Bunzel, M., Vogel, R. F. (2017). **Characterization of an acetan-like heteropolysaccharide produced by *Kozakia baliensis* NBRC 16680.** International Journal of Biological Macromolecules, (106).

SUMMARY

Bacteria produce a wide range of different exopolysaccharides (EPS), which can be either homopolysaccharides (HoPS) or heteropolysaccharides (HePS). There is a variety of well-established HoPS on the market, such as cellulose, dextran and levan. In addition to the already well-established HoPS, also HePS get increasingly into the focus of interest. Due to their highly branched structures, they show strong viscosifying effects even in low concentrations. This makes them to ideal additives when it comes to the thickening of food and cosmetic products. Xanthan is probably the most prominent example of an industrially used HePS. It is derived from *Xanthomonas (X.) campestris* and represents the first industrially produced biopolymer that is manufactured in large-scale fermentations. However, in relation to their energy-intensive biosynthesis, HePS are commonly formed in small amounts, making newly discovered HePS more expensive for the industrial production. Nevertheless, HePS can be commercialized by introducing them into specific market niches, such as directly *in situ* produced during food fermentation. To guarantee the safety of the respective products, it is important to use non-pathogenic, food-associated bacteria as a HePS forming starter culture. A HePS producing group of bacteria that has not been intensively studied yet but meets this specification are the acetic acid bacteria (AAB). AAB represent a group of food-grade, non-pathogenic bacteria, in which various members produce a wide range of different exopolysaccharides (EPS). The most prominent AAB produced EPS is cellulose, being the main component of the "mother of vinegar".

Kozakia (*K.*) *baliensis* is a relatively new member of AAB, which produces ultra-high molecular weight levan from sucrose. Throughout cultivation of two *K. baliensis* strains (DSM 14400, NBRC 16680) on sucrose- deficient media, we found that both strains still produce high amounts of mucous, water-soluble substances from mannitol and glycerol as (main) carbon sources. This indicated, that both *Kozakia* strains additionally produce new classes of so far not characterized HePS.

The present work aims to describe this new polysaccharide of *K. baliensis* (DSM 14400, NBRC 16680), starting with its genomic origin and biosynthesis, conjoined with its structural elucidation, followed by the attempt to clarify the physiological significance of the HePS for the bacterium. A first neutral sugar monomer analysis of the isolated *K. baliensis* (DSM 14400, NBRC 16680) EPS, should confirm the heteropolysaccharide character of the EPS. Whereas a

subsequent comparative genomic approach should further identify associated genomic clusters. For this purpose, a preliminary HPLC analysis of both *K. baliensis* (DSM 14400, NBRC 16680) HePS was carried out, which verified that both EPS consist of glucose, galactose and mannose, respectively. Thereupon, both *K. baliensis* strains (DSM 14400, NBRC 16680) were sequenced by PacBio single-molecule real-time (SMRT) sequencing. With the received data, *de novo* assemblies were generated and circularized genomes could be established for each strain. For *K. baliensis* DSM 14400, seven contigs were generated upon sequencing, which were further assembled to one complete genome, consisting of one circularized chromosome (2,888,029 bp) and four circularized plasmids, as well as two partial plasmids. Genome sequencing of *K. baliensis* NBRC 16680 resulted in six contigs, which were assembled to one chromosome (2,807,246 bp) and three plasmids, as well as two partial plasmids.

Via a comparison of the genomes of both *K. baliensis* strains (DSM 14400, NBRC 16680) with related AAB genomes, under consideration of known literature about the xanthan biosynthesis, EPS forming clusters were identified. As expected, complete ORFs coding for levansucrases could be detected in both *Kozakia* strains. In contrast to *K. baliensis* NBRC 16680, solely the DSM 14400 strain harbors a plasmid encoded cellulose synthase operon and fragments of a truncated levansucrase operon. Additionally, in both *K. baliensis* strains, clusters for putative HePS producing genes were identified, including a "*gum*-like cluster" of 25 kb. It could be demonstrated that the *gum*-like cluster is related to the well characterized *gum* cluster coding for the xanthan synthesis in *X. campestris*. The *gum*-like clusters of both *Kozakia* strains involve in total 19 ORFs, whereas the homologous *gum*-like genes are *gumB*, *-C*, *-D*, *-E*, *-H*, *-J*, *-K*, and *–M*. While the *gum*-like clusters of both *K. baliensis* strains (DSM 14400, NBRC 16680) share identical genetic organizations among each other, they are closely related to *gum*-like HePS clusters from other AAB such as *Gluconacetobacter (Ga.) diazotrophicus* and the acetan forming strain, *Komagataeibacter (Ko.) xylinus*. By comparison of the different HePS clusters, including both AAB clusters, the well characterized *gum*-cluster of *X. campestris* and the *K. baliensis* cluster, specific functions could be assigned to several *K. baliensis gum*-like genes. The identified genes, *gumD*, *gumH* and *gumM*, code for enzymes that are necessary for the incorporation of the *K. baliensis* HePS constituents glucose and mannose, which are also components of both AAB HePS and xanthan. In addition to the identified glycosyltransferase genes, the *K. baliensis gum*-like clusters harbor genes, which are necessary for the HePS polymerization and export (e.g. *gumB*, *gumC*, *gumE*) and are also present in the mentioned AAB and xanthan HePS clusters. However, the AAB *gum*-like clusters also vary from each

other and show different gene arrangements. The acetan cluster of *Ko xylinus*, as well as the HePS cluster of *Ga*. *diazotrophicus*, include a putative acetyltransferase gene (*gumF)*, coding for a protein that incorporates acetyl residues into the particular HePS, which is not present in the respective *K*. *bali*ensis clusters. This suggests that both *K*. *baliensis* HePS are free of acetyl residues. Further, no rhamnosyl- transferase gene (*aceR)* is located in the two *K*. *baliensis* clusters, unlike to the acetan cluster of *Ko*. *xylinus*. Acetan consists of D -glucose, D -rhamnose, D -mannose and D-glucuronic acid, accordingly only in this AAB cluster the *aceR* gene is involved. The acetan cluster additionally includes a *gumK* gene, which is also detected in both *K*. *baliensis* clusters and suggests the existence of an additional glucuronic acid (GlcA) residue in the HePS, which could not be detected via HPLC sugar monomer analysis.

The comparative genomics got adjuvant insights into the biosynthesis of this novel *gum*-like HePS and allowed to establish a hypothetical composition of the *K*. *baliensis* HePS, which seems to be similar to the one from acetan and xanthan. In order to get better insight into the *K*. *baliensis* HePS structure and to evaluate if the genomic predictions can be verified a detailed structural analysis of the *K*. *baliensis* NBRC 16680 HePS was carried out by using HPAEC-PAD, methylation analysis and NMR. Because *K*. *baliensis* DSM 14400 harbors a cellulose synthase operon in the genome and separation of the formed HePS from the produced cellulose cannot be guaranteed, the structure assignment was solely made with the NBRC 16680 HePS. Methylation analysis of the *K*. *baliensis* NBRC 16680 HePS allowed the detection of 1,4-/1,3,4-substituted glucose units, 1,2-substituted mannose and 1,4-substituted glucuronic acid units. Furthermore, a putative, unidentified uronic acid and 1,6-substituted galactose units were found to be HePS constituents. With the obtained data, it was not possible to elucidate the exact structure of the *K*. *baliensis* formed HePS. However, a comparison of the HePS cluster of *K*. *baliensis* NBRC 16680 with the *gum*-cluster of *X*. *campestris* and the acetan cluster of *Ko*. *xylinus* E25, together with the NMR data, finally lead to a verisimilar HePS structure. We hypothesize that the *K*. *baliensis* NBRC 16680 HePS has a (1→4)-linked acetan-like backbone, composed of 1,4- and 1,3,4-substituted glucose units, with ramifications at position O3. This hypothesis is supported by roughly equal portions of 1,4-Glcp and 1,3,4-Glcp. The branching side chains are possibly composed of 1,2-substituted mannose, 1,4-substituted glucuronic acid and 1,6-substituted glucose. In addition, 1,6-substituted galactose units as well as a putative uronic acid might also be part of the side chains of the *K*. *baliensis* HePS, but cannot be correlated with the acetan biosynthesis machinery, due to missing corresponding biosynthesis

genes. Rhamnose units and acetyl residues, which are structural elements of acetan, were not detected in the HePS of *K. baliensis* NBRC 16680. This observation is supported by the absence of corresponding homologous enzymes in the *K. baliensis* HePS cluster, which incorporate rhamnose and acetyl residues into acetan. Instead of the terminal rhamnose in the side chains of acetan, a further attachment of 1,6-substituted glucose and 1,6-substituted galactose units can be assumed, marking this HePS as a unique EPS that differs from known AAB EPSs.

To address the structural reproducibility of the *K. baliensis* HePS, its chemical composition was investigated in media with different initial carbon sources or supplemented magnesium. Different, media dependent HePS amounts were detected. Additionally, different rheological properties and elution profiles during AF4 and HPSEC separations could be observed. Irrespective of the altering media, the *K. baliensis* NBRC 16680 HePS showed an analogue monomer composition, with similar amounts of glucose, galactose and mannose. As the HePS biosynthesis is a directed process, we presume that this is the result of a differentiated regulation process of enzymes involved during the HePS polymerization. Especially magnesium in the medium could result in a longer main chain of the *K. baliensis* NBRC 16680 HePS, resulting in larger molecules with a constant monomer composition.

By considering a non-HePS producing *K. baliensis* strain, it could be excluded that these deviations are due to a simultaneously formed polysaccharide. This mutant strain of *K. baliensis,* NBRC 16680, exhibited a transposon insertion in front of the *gumD* gene of its *gum*-like cluster, leading to a morphology switch, from smooth wild-type (S) to rough mutant colonies (R), with no HePS production in shaking medium. On the one hand, this demonstrates the essential role of the *gum* genes for the HePS production of *K. baliensis*. On the other hand, a clear difference between the two *Kozakia* strains becomes clear, since only *K. baliensis* NBRC 16680 comprises this specific mobile element in his genome. To clarify whether this has a physiological effect on the NBRC 16680 strain, and whether the transposon insertion is a random or directed event, maybe triggered by environmental factors, we analyzed the mutation frequency from smooth to rough colonies under different growth conditions. Here, we could demonstrate that the mutation frequency from smooth to rough strains was significantly increased upon growth on ethanol and acetic acid supplemented medium, which suggests a better adaption of the R strain. However, this connection cannot be considered without the inclusion of another HePS forming cluster of *K. baliensis*. *K. baliensis* (NBRC 16680, DSM 14400) harbors two main HePS clusters, a *gum*-cluster coding for the previously described secreted HePS, and a *polABCD*E cluster coding for a HePS, forming a surface pellicle during

static cultivation. Both *K*. *baliensis* phenotypes, R and S, are still able to form a pellicle on the surface of the medium. However, the morphologic switch went along with changes in the sugar contents of the produced pellicle HePS and therefore could also play a role in the different assertiveness against ethanol and acetic acid. To get better insights into the composition and physiological function of the pellicle, a deletion of the *K*. *baliensis* NBRC 16680 R *polE* gene was conducted, which is reported to be involved in pellicle formation in other AAB. As expected, the *K*. *baliensis* NBRC 16680 R:Δ*polE* mutant strain was not able to form a regular pellicle, but surprisingly secreted a HePS into the medium, with a similar sugar monomer composition as the pellicle HePS formed by the R strain. Nevertheless, solely the *K*. *baliensis* NBRC 16680 rough mutant strain showed in static and shaking growth experiments a markedly increased tolerance towards acetic acid and ethanol. Accordingly, the succeeded *polE* gene deletion in *K*. *baliensis* NBRC 16680 R, showed an involvement of the *polE* gene in this process. We suppose that the increased tolerance is caused by a transformation of the *K*. *baliensis* NBRC 16680 EPS, from a secreted HePS in the wild type to a capsular HePS in the rough mutant strain, in which the *polE* gene seems to play an essential role in the formation of a cell-associated, capsular polysaccharide. Thus, the results show that the alteration from a smooth to rough phenotype is presumably a genetically triggered event, which is related to an increased acetic acid and ethanol tolerance of *Kozakia* NBRC 16680. If *K*. *baliensis* DSM 14400 is comprised in this context, which shows a comparatively high intrinsic acetic acid tolerance from the beginning, it is likely that both strains have developed different adaptive mechanisms to cope with changing natural habitats, like increasing acetic acid or ethanol concentrations. Hence, the genetic and morphological switch of *K*. *baliensis* NBRC 16680 could represent an adaptive evolutionary step during the bacterial development, which slowly leads to the genetic drift of both strains.

Assuming a cluster driven switch, a holistic view of the HePS cluster distribution in AAB in general would be interesting. But which AAB contains a corresponding HePS cluster? And can thereby physiological similarities be discovered? Since little is known about HePS clusters in AAB in general, we performed a whole genome sequencing of potentially HePS-forming AAB. We sequenced the genomes of *A*. *aceti* TMW 2.1153, *A*. *persici* TMW 2.1084 and *N*. *chiangmaiensis* NBRC 101099, which are all potential HePS producers. The genomes were assembled to either one contig, in case of *A*. *aceti* TMW 2.1153 and *N*. *chiangmaiensis* NBRC 101099, or to two contigs, in case of *A*. *persici* TMW 2.1084, carrying an additional plasmid of 526,169 bp. In all strains, HePS forming gene clusters could be detected, whereas each strain

carried a whole *pol* gene clusters, *polABCDE*, associated with pellicle formation. Furthermore, in *A*. *aceti* TMW 2.1153, *gum*-like genes, similar to genes in the *gum* cluster of *X*. *campestris*, could be identified.

We used the whole genome sequencing data to make a cross-strain representation of existing HePS clusters in AAB and to get better insights into their biosynthesis pathways and the possible physiological relevancy of the particular HePS. It could be shown, that in comparison to the different *Acetobacter* strains, most of the *Komagataeibacter* strains harbor only a fragmented *pol*-cluster in their genome, and moreover no *polE* gene. An exception are *Ko*. *xyli*nus E25 and *Ko*. *europaeus* NBRC 3261, whose *gum*-like cluster is flanked by the *polAB* and *polCD* genes. Thus, it could be a bold hypothesis that the *polE* gene solely occurs in AAB strains, lacking a cellulose pellicle formation. For the cellulose-forming *Komagataeibacter* strains, no additional acetic acid protection would be required, which makes the *pol* cluster obsolete, whereas in *Acetobacter* strains, acetic acid resistance is ensured by a capsular HePS pellicle. This shows, by means of the example of both *K*. *baliensis* strains, an evolutionary drift of both genera, which manifests itself, among other things, through the altering utilization of different EPS clusters.

ZUSAMMENFASSUNG

Bakterien sind in der Lage eine Vielzahl an verschiedenen Exopolysacchariden (EPS) zu bilden, wobei es sich entweder um Homopolysaccharide (HoPS) oder um Heteropolysaccharide (HePS) handelt. Es gibt bereits eine große Bandbreite an gut etablierten HoPS auf dem industriellen Markt, wie beispielsweise Cellulose, Dextran und Levan. Jedoch geraten neben den gut etablierten HoPS auch HePS immer mehr in den Fokus des Interesses. Ihrer strukturell bedingten Fähigkeit bereits in geringen Konzentrationen stark viskosifizierend zu wirken, macht sie zum idealen Zusatzstoff wenn es um die Verdickung von Lebensmitteln und Kosmetikprodukten geht. Dies kann man exemplarisch am HePS Xanthan sehen, dass vom Bakterium *Xanthomonas (X.) campestris* gebildet wird, und welches das erste im industriellen Maßstab produzierte Biopolymer ist. Allerdings führt die energieintensive HePS Synthese dazu, dass HePS nur in geringen Mengen gebildet werden, was höhere Kosten für die Etablierung eines neuen HePS am Markt bedeutet. Nichtsdestotrotz können neue HePS kommerzialisiert werden, indem man sich auf bestimmte Marktnischen spezialisiert, wie beispielsweise die strukturelle Verbesserung von Lebensmitteln durch direkte Einbringung des HePS während der *in-situ* Fermentation. Um hierbei die Sicherheit des jeweiligen Lebensmittels zu gewährleisten, ist es wichtig lebensmitteltaugliche, apathogenen Stämme zu verwenden. Dies bringt uns zu den Essigsäurebakterien (ESB), eine Familie aus Lebensmittel assoziierte, apathogenen Bakterien. Viele ESB sind in der Lage verschiedenen EPS zu bilden, wobei Cellulose das wohl bekannteste EPS darstellt und in verschieden industriellen Sektoren Anwendung findet.
Kozakia (*K.*) *baliensis*i ist ein relative neues Mitglied der ESB und produziert unter anderem Levan aus Saccharose. Während der Kultivierung von *K. baliensis* (DSM 14400, NBRC 16680) auf Saccharose-freiem Medium, konnte festgestellt werden, dass beide Stämme immer noch ein muköse, wasserlösliche Substanz bilden. Daraus lässt sich schließen, dass beide *Kozakia* Stämme ein unbekanntes und bislang noch nicht charakterisiertes HePS bilden.

Die vorliegende Arbeit hat das Ziel, dieses potentielle HePS von *K. baliensis* (DSM 14400, NBRC 16680) zu beschreiben, beginnend bei seinem genomischen Ursprung und seiner Biosynthese, gefolgt von seiner strukturellen Aufklärung und dem zusätzlichen Versuch die physiologische Bedeutung des HePS zu ergründen. Eine ersten Zuckermonomeren Analyse des isolierten EPS sollte zunächst Klarheit darüber schaffen, ob es sich generell um ein HePS

handelt, um dann mittels vergleichender Genomik die dazugehörigen HePS Cluster zu identifizieren. Dafür wurden vorläufige HPLC Analysen des isolierten *K. baliensis* (DSM 14400, NBRC 16680) HePS durchgeführt, mit denen verifiziert werden konnte, dass sich bei beide EPS, jeweils um ein HePS handelt, welches sich aus Glucose, Galactose und Mannose zusammensetzen. Daraufhin wurde eine Genomsequenzierung beider Stämme, mittels PacBio single-molecule real-time (SMRT) sequencing, durchgeführt. Mit den erhaltenen Daten wurden de novo assemlies generiert, die bei beiden Stämmen zu zirkulären Genomen zusammengesetzt werden konnten. So erhielt man für *K. baliensis* DSM 14400 sieben contigs, die zu einem zirkuläres Chromosom (2,888,029 bp), sowie vier Plasmide und zwei partielle Plasmide zusammengesetzt werden konnten. Die Genomdaten von *K. baliensis* NBRC 16680 resultierten in sechs contigs, die ein zirkuläres Chromosom (2,807,246 bp), sowie vier Plasmide und zwei partielle Plasmide ergaben.

Durch vergleichende Analysen beider *K. baliensis* (DSM 14400, NBRC 16680) Genome, mit verwandten ESB, unter zusätzlichen Einbezug der vorhandenen Literatur über die EPS Biosynthese von *X. campestris*, konnten HePS Cluster in beiden Stämmen identifiziert werden. Wie zu erwarten war, wurden in beiden *Kozakia* Stämmen vollständige ORFs für Levansucrasen identifiziert. Zusätzlich wurde in *K. baliensis* DSM 14400 Plasmid-codierte Cellulose Synthase Gene und ein verkürztes Levansucrase Operon gefunden, die nicht in *K. baliensis* NBRC 16680 zu finden sind. Des Weiteren wurde in beiden *K. baliensis* Stämmen HePS-assoziierte Cluster identifiziert, zu denen ein 25 kb großes "*gum*-like cluster" gehört. Dieses *gum*-like Cluster ist ähnlich zu dem gut beschriebenen *gum* Cluster in *X. campestris*, welches für die Xanthan Synthese codiert. Das *gum*-like beider *Kozakia* Stämme umfasst insgesamt 19 ORFs, mit folgenden homologen *gum*-like Genen: *gumB, -C*, *-D*, *-E*, *-H*, *-J*, *-K*, und *–M*. Während die *gum*-like Cluster der beiden *K. baliensis* (DSM 14400, NBRC 16680) Stämme untereinander eine identisch Organisation aufweisen, ist ihr nächster Verwandter unter den ESB zu finden, bei *Gluconacetobacter (Ga.) diazotrophicus* und dem Acetan-Bildner *Komagataeibacter (Ko.) xylinus*. Durch den direkten Vergleich der *gum*-like Cluster beider *Kozakia* Stämme mit den genannten ESB und dem *gum*-Cluster von *X. campestris*, konnten den einzelnen Genen bestimmte Funktionen zugeschrieben werden. Dabei codieren bestimmte Gene im *Kozakia* HePS Cluster ‚*gumD*, *gumH* und *gumM*, für Glykosyltransferasen, die wichtig für den Einbau von Glucose und Mannose ins HePS sind, und ebenfalls im xanthan- und acetan-Cluster vorkommen. Zusätzlich konnten Gene im *K. baliensis gum*-like Clusters identifiziert werden, die während der HePS Polymerisation und dem Export eine Rolle spielen; *gumB,*

gumC, gumE, und ebenfalls in den anderen Clustern zu finden sind. Des Weiteren beinhalten beide *K*. *baliensis* *gum*-like Cluster ein Gen, *gumK*, welches für eine potentielle Glykosyltransferase codiert, die sowohl in Xanthan, wie auch in Acetan, den Einbau von Glucuronsäre katalysiert. Dies lässt ein zusätzliches Vorhandensein von Glucuronsäure im *K*. *baliensis* HePS vermuten, welche jedoch nicht mittels HPLC detektiert werden konnte. Neben diesen Ähnlichkeiten liegen jedoch auch deutliche Unterschiede zwischen den einzelnen Clustern vor. So beinhaltet das Acetan-Cluster von *Ko xylinus*, sowie das HePS Cluster von *Ga*. *diazotrophicus*, beispielsweise eine vermeintliches Acetyltransferasegen, *gumF*, welches für ein Enzym codiert, dass Acetyl-Reste in das jeweilige HePS einbaut. Dieses Gen ist nicht im *K*. *bali*ensis *gum*-like Cluster enthalten, was vermuten lässt, dass beide *K*. *baliensis* HePS keine zusätzlichen Acetyl-Reste aufweisen.

Darüber hinaus ist in beiden *K*. *baliensis* Clustern kein *aceR*-Gen vorhanden, welches für eine Rhamnosyl-Transferase codiert, und fester Bestandteil des Acetan Clusters von *Ko*. *xylinus* ist. Dies ist darauf zurückzuführen, das einzig Acetan Rhamnose beinhaltet, demzufolge ist nur hier das *aceR* Gen zu finden.

Wie man an den getroffenen funktionellen Zuordnungen der einzelnen Gene im *gum*-like Cluster von *K*. *bali*ensis erkennen kann, hat die vergleichende Genomik einen großen Teil zur Aufklärung der HePS Biosynthese in *K*. *bali*ensis beigetragen, die ähnlich zu der von Acetan und Xanthan erscheint. Um nun einen genauen Einblick in die Struktur des *K*. *baliensis* HePS zu erhalten und um gleichzeitig zu überprüfen, ob die getroffenen Aussagen verifiziert werden können, wurde eine detaillierte Strukturanalyse des *K*. *baliensis* NBRC 16680 HePS, unter Verwendung von HPAEC-PAD, Methylierungsanalysen und NMR, durchgeführt. Dabei wurde die Strukturbestimmung des HePS nur mit dem NBRC 16680 HePS durchgeführt, da eine möglichen Cellulose Produktion, über das in *K*. *baliensis* DSM 14400 vorhandenen Cellulose Synthase Operon, zu unklaren Ergebnissen führen würde.

Mittels der Methylierungsanalysen des *K*. *baliensis* NBRC 16680 HePS konnten 1,4-/1,3,4-verknüpfte Glucoseeinheiten, 1,2-verknüpfte Mannose und 1,4-verknüpfte Glucuronsäure Einheiten, detektiert werden. Zudem konnte eine potentielle Uronsäure, sowie eine 1,6-verknüpfte Galactose Einheit als HePS Bestandteile identifiziert werden. Mit den erhaltenen Daten konnte jedoch keine exakte Struktur des *K*. *baliensis* NBRC 16680 HePS bestimmt werden. Kombiniert man jedoch die erhaltenen Daten, mit den bereits getroffenen Aussagen über die Funktion der einzelnen Gene im *gum*-like von *Kozakia,* lässt sich eine hypothetische Struktur des HePS erstellen. So lässt sich ableiten, dass das *K*. *baliensis* NBRC 16680 HePS

wahrscheinlich einen (1→4)-verknüpften Acetan-ähnliches Rückgrat hat, welches aus 1,4- und 1,3,4-verknüpften Glucose Einheiten aufgebaut ist, mit Abzweigungen an Position O3. Diese Hypothese wird dadurch gestützt, dass die O4-verknüpften Glucopyranose Einheiten, 1,4-Glcp und 1,3,4-Glcp in annährend gleichen Mengen vorhanden sind. Des Weiteren wird angenommen, das die Seitenketten des HePS aus 1,2-verknüpfter Mannose, 1,4-verknüpfter Glucuronsäure und 1,6-verknüpfter Glucose bestehen. 1,6-verknüpfte Galactose, sowie die vermeindliche Uronsäure, werden als weitere Bestandteile in der Seitenkette des *Kozakia* HePS vermutet. Da jedoch weder Acetan, noch Xanthan, Galactose beinhalten, und dementsprechend die zugehörigen Gene in den jeweiligen Clustern fehlen, konnte keine genaue Zuordnung getroffen werden. Es wird vermutet, dass sich die vorhandenen Galactose Einheiten am Ende der Seitenketten des *K*. *baliensis* HePS befinden, als pendent zu den Rhamnose Einheiten in Acetan. In Übereinstimmung dazu, konnte keine Rhamnose Einheiten und Acetyl-Gruppen im *K*. *baliensis* HePS identifiziert werden, die hingegen Bestandteil von Acetan sind und durch entsprechende Gene codiert werden, die nicht im *gum*-like Cluster von *K*. *baliensis* vorkommen.

Um die strukturelle Konstanz des *K*. *baliensis* NBRC 16680 HePS zu überprüfen, wurde seine chemische Zusammensetzung in Medien mit unterschiedlichen Kohlestoffquellen und/oder Magnesium untersucht.

Dabei konnten abhängig vom Medium unterschiedliche Mengen an *K*. *baliensis* HePS gewonnen werden, die unterschiedliche rheologische Eigenschaften und Elutionsprofile zeigten. Medium Unabhängig, zeigten die isolierten *K*. *baliensis* HePS jedoch eine analoge Zusammensetzungen der einzelnen Monomere, mit gleichen Mengen an Glucose, Galactose und Mannose. Daraus ergibt sich die Vermutung, dass die gerichtete Biosynthese zu einer konstanten Zusammensetzung der Wiederholungseinheiten des HePS führt, jedoch die Länge des HePS durch Faktoren beeinflusst werden kann. Wobei besonders Magnesium im Medium, über eine mögliche Stimulierung von Polymerasen, zu einer längeren Hauptkette des *K*. *baliensis* NBRC 16680 HePS und somit zu einem größeren Molekül führen kann.

Über die Identifizierung eines HePS Defizienten *K*. *baliensis* Stamms, konnte zusätzlich gezeigt werden, dass die Untersuchten Abweichungen nicht das Resultat eines simultan gebildeten, zweiten EPS sind. Der mutierte Stamm, weißt eine Transposon-Insertion vor dem *gumD*-Gen im *gum*-like Cluster von *K*. *baliensis* NBRC 16680 auf, die zu einer morphologischen Veränderung der Kolonien führt, von HePS bildenden Wildtyp (S), zu nicht HePS bildenden

mutierten Kolonien (R). Zusätzlich wird durch diesen mutierten Stamm die essenzielle Rolle der *gum*-Gene, speziell des *gumD*-Gens, für die HePS Bildung von *K*. *baliensis* demonstriert. Betrachtet man *K*. *baliensis* DSM 14400 und *K*. *baliensis* NBRC 16680 nun im direkten Vergleich miteinander, wird hier ein klarer Unterschied deutlich, nur *K*. *baliensis* NBRC 16680 enthält das spezifische Transposon im Genom. In diesem Zusammenhang wollten wir untersuchen, ob die Transposon Insertion einen physiologischen Effekt auf *K*. *baliensis* NBRC 16680 hat und ob diese ein zufälliges oder gerichtetes Event darstellt, welches möglicherweise durch äußere Faktoren beeinflusst wird. Dafür wurde die Muatationsfrequenz von *K*. *baliensis* NBRC 16680 in unterschiedlichen Wachstumsmedien ermittelt. Es konnte gezeigt werden, dass die Mutationsfrequenz von Wildtyp zu mutierten Kolonien, im Medium mit supplementärem Ethanol und Essigsäure signifikant erhöht war, was eine bessere Anpassung des mutierten *K*. *baliensis* NBRC 16680 vermuten lässt. Dieser Verbindung kann jedoch nicht getroffen werden, ohne dabei ein weiteres HePS Cluster von *K*. *baliensis* einzubeziehen. *K*. *baliensis* (NBRC 16680, DSM 14400) beinhaltet zwei chromosomal codierte HePS Cluster, das bereits erwähnte *gum*-Cluster, welches für das sekretierte HePS codiert, und das *polABCD*E Cluster, welches für die Pellikel-Bildung während der statischer Kultivierung verantwortlich ist. Während einer statischen Kultivierung sind beide *K*. *baliensis* Phänotypen, R und S, in der Lage einen Perllikel auf der Oberfläche des Mediums zu bilden. Jedoch geht diese morphologische Umgestaltung mit einer Veränderung der Zuckerzusammensetzung des Pellikels einher, so dass auch dies eine Rolle bei der besseren Anpassung von *K*. *baliensis* NBRC 16680 R an Ethanol und Essigsäure spielen kann. Um einen besseren Einblick in den Aufbau und die physiologische Funktion des Pellikels zu erhalten, wurde eine Deletion des *polE*-Gens in *K*. *baliensis* NBRC 16680 R durchgeführt, da das *polE*-Gens eine zentrale Rolle bei der Pellikel-Bildung in anderen ESB spielt. Wie zu erwarten war konnte der genetisch veränderte *K*. *baliensis* NBRC 16680 R:Δ*polE* keinen festen Pellikel mehr bilden, begann jedoch HePS in das umgebende Medium zu sekretierten, welches eine ähnliche HePS Zusammensetzung hatte, wie das Pellikel HePS von *K*. *baliensis* NBRC 16680 R. Nichtsdestotrotz, zeigte weiterhin einzig *K*. *baliensis* NBRC 16680 R eine erhöhte Toleranz gegen Ethanol und Essigsäure. Eine *polE* Deletion führt folglich zu einer verminderten Toleranz gegenüber Ethanol und Essigsäure, was eindeutig die Bedeutung des *polE* Gens in diesem Prozess unterstreicht. So nehmen wir an, dass die erhöhte Toleranz bei *K*. *baliensis* NBRC 16680 R durch eine Umwandlung des HePS, von eine sekretierten HePS im Wildtyp zu einem kapsulären HePS im R Stamm, zustande kommt. Dabei könnte PolE eine essentielle Rolle bei der Bildung des kapsulären HePS spielen, wobei eine

polE Deletion zur sekretion des HePS führt. Wenn man nun noch den Wildtyp Stamm von *K*. *baliensis* DSM 14400 mit berücksichtigt, der generell eine hohe Toleranz gegen Ethanol und Essigsäure aufweist, könnte man die morphologische Veränderung von *K*. *baliensis* NBRC 16680, S zu R, als ein adaptiver Anpassungsmechanismus an wechselnde Habitate betrachten. So könnte diese Veränderung ein evolutionärer Schritt des Bakteriums hin zu anderen Umweltnischen sein, und stellt somit einen langsamen Drift dieser beiden *K*. *baliensis* Stämme, DSM 14400 und NBRC 16680, dar.

Wenn man nun von einem HePS Cluster beeinflusster Switch bei *K*. *baliensis* ausgeht, wäre ein holistischer Blick auf die HePS Cluster Verteilung innerhalb der ESB interessant. Doch welche ESB beinhalten HePS Cluster und weisen diese physiologische Ähnlichkeiten auf?

Bisher gibt es keinen weitreichenden Überblick über HePS Cluster in ESB im generellen. Um einen besseren Überblick zu erhalten, wurden Genomsequenzierungen von potentiellen HePS-Bildnern durchgeführt; *A*. *aceti* TMW 2.1153, *A*. *persici* TMW 2.1084 and *N*. *chiangmaiensis* NBRC 101099. Die Genome wurde nachfolgend entweder zu einem Genomen mit einem einzigen Chromosom, wie im Fall von *A*. *aceti* TMW 2.1153 und *N*. *chiangmaiensis* NBRC 101099, oder zu Genomen mit einem Chromosom und einem Plasmid, wie bei *A*. *persici* TMW 2.1084, zusammengefügt. In allen drei Stammen konnten HePS Cluster identifiziert werden, wobei alle komplette *pol* Cluster (*polABCDE*) beinhalten. Des weiteren konnte in *A*. *aceti* TMW 2.1153 ein *gum*-like Cluster, welches Ähnlichkeit zum Cluster von *X*. *campestris* aufweist, identifiziert werden. Um nun einen generellen Überblick zu erhalten, wurden die gesammelten Genomdaten für einen Kreuzvergleich mit vorhanden ESB Genomdaten verwendet, und HePS Cluster identifiziert. Dadurch konnte gezeigt werden, dass im Vergleich zu verschiedenen *Acetobacter* Stämmen, die meisten *Komagataeibacter* Stämme nur ein zerstückeltes *polABCD* Cluster im Genom aufweisen und darüber hinaus kein *polE* Gen besitzen. Einen Sonderfall bilden hierbei *Ko*. *xyli*nus E25 und *Ko*. *europaeus* NBRC 3261, deren *gum*-like Cluster von den *pol* Genen flankiert ist. Es wird vermutet, dass die Cellulose bildenden *Komagataeibacter* Stämme keinen zusätzlichen Schutz gegen Essigsäure benötigen, was ein *pol* Cluster obsolet macht, wohingegen bei *Acetobacter* Stämmen die Essigsäureresistenz durch einen kapsulären HePS Pellikel gewährleistet ist. Dies zeigt, wie im Beispiel der beiden *K*. *baliensis* Stämme eine evolutionären Drift der verschiedenen *Acetobateracea* Gattungen der sich unter anderem durch die unterschiedliche Verwendung verschiedener EPS Cluster manifestiert.

1. INTRODUCTION

Phylogeny and Physiology of Acetic Acid Bacteria including *Kozakia baliensis*

Acetic acid bacteria (AAB) are obligate aerobic, Gram- negative α-Proteobacteria, within the family of *Acetobacteracea*. They are a monophyletic group, currently sub-divided into ten genera: *Acetobacter*, *Gluconobacter*, *Gluconacetobacter*, *Granulibacter*, *Asaia*, *Swaminathania*, *Kozakia*, *Neoasaia*, *Acidomonas*, *Saccharibacter* (Yamada 2008). The genus *Gluconacetobacter* is further divided into two subclusters, dating back to phylogenetical, phenotypical and ecological differences of the *Gluconacetobacter liquefaciens* and *Gluconacetobacter xylinus* group, therefore classified under a separate new genus, named *Komagataeibacter (Ko.)* (Yamada 2012).

The eponymous characteristic trait of acetic acid bacteria is the production of acetic acid from ethanol (Ruiz 2000). Only *Asaia* (*As*.) *sp*. makes an exeption to the ruleand cannot produce acetic acid from ethanol (Cleenwerck 2008). Apart from acetic acid, these microorganisms produce additional metabolites, including gluconic acid, 5-ketogluconate and dihydroxyacetone (Charney 1987; Prust et al. 2005). These products occur from incomplete oxidation reactions, primarily performed by membrane-bound dehydrogenases, anchored in the cytoplasmic membrane with their active centers oriented into the periplasmic space (Hekmat 2003; Matsushita 2003). Multiple species of acetic acid bacteria are capable of incomplete oxidation of a wide range of carbohydrates and alcohols to aldehydes, ketones and organic acids (Deppenmeier 2002; Matsushita 2003). The oxidation products are finally secreted from the cells into the external environment, thereby lowering the extracellular pH, which inhibits the growth of competing microorganisms (Deppenmeier 2009). Especially *Acetobacter* (*A*.) and *Gluconacetobacter* (*Ga*.) species have a very high ethanol oxidation capacity, leading to accumulation of vast amounts of acetic acid outside their cells. Therefore, AAB have developed strain-specific mechanisms for acid resistance: *A*. *aceti* strain 1023, for example, detoxifies acetic acid by possessing a complete set of citric acid cycle (CAC) enzymes, including a succinyl-CoA:acetate CoA-transferase (SCACT) via the acetic acid resistance gene *aarC* (Mullins 2013). Further, an active export of acetic acid to the extracellular space via efflux pumps could be observed (Matsushita 2005).

AAB can be isolated from a variety of different sugary or alcohol-rich habitats like nectar, flowers, fruits, spoiled beverages and fermented foods. Some AAB from the genera

Acetobacter, *Gluconacetobacter* and *Swaminathania* are associated with plants and the promotion of plant growth by fixing atmospheric nitrogen (Pedraza, 2008). Further, some AAB establish symbiotic relationships with several insects, all relying on sugar-based diets (Crotti 2010), or work as pathogens, *Gluconobacter (G.) morbifer* sp. nov., in case of *Drosophila melanogaster* (Roh 2008).

Depending on their habitat, AAB utilize altering metabolic pathways. Hence, not all AAB oxidize their substrates completely. Based upon these altering pathways, AAB are classified into their three main genera: *Acetobacter* and *Gluconacetobacter* strains prefer alcohols as substrates and are commonly detected in alcoholic habitats, completely oxidizing the ethanol to acetic acid, which is further oxidized to CO_2 and H_2O (overoxidation of ethanol). *Gluconobacter*, *Saccharibacter (S.)* and *Neoasaia (N.)* mainly occur in sugar rich environments and are specialized in the incomplete oxidations of different sugars, while oxidation of acetic acid is absent (Deppenmeier 2009). *Asaia*, *Kozakia*, *Swaminathania* and *Granulibacter* strains are physiological "intermediates" of these two genera due to their ability to weakly overoxidice acetate (Yamada 2008). The genus *Kozakia* was subsequently proposed by Lisdiyanti (2002) as the sixth genus with a single species, *Kozakia* (*K*.) *baliensis*, which was isolated from palm brown sugar and ragi collected in Bali and Yogyakarta, Indonesia. *K*. *baliensis* weakly oxidizes acetate and lactate to carbon dioxide and water and produces acetic acid from ethanol. Although *Kozakia* is a plant connected bacteria, it does not assimilate ammoniacal nitrogen, and is unable to fix atmospheric nitrogen (Lisdiyanti 2002), unlike other plant associated AAB which act as growth promoting bacteria for the colonized plants (Pedraza 2008).

The core metabolism of AAB was chiefly highlighted by predictions from genome-sequencing of *Gluconobacter oxydans* 621H (Prust et al., 2005), whereas the detailed mechanisms for oxygen consumption and energy generation are not clearly understood yet. *G*. *oxydans* 621H oxidizes sugars and alcohols mainly in the periplasm via incomplete oxidation, whereas a small fraction is metabolized in the cytoplasm by the Entner-Doudoroff and pentose phosphate pathway (Deppenmeier and Ehrenreich 2009). The Embden-Meyerhof pathway is not functional and the citrate cycle incomplete, due to the absence of the succinate dehydrogenase (Prust et al. 2005). The incomplete oxidations are predominantly performed by membrane-bound dehydrogenases, which transfer the released electrons into the respiratory chain for reduction of oxygen and energy conservation. The respiratory chains of acetic acid bacteria, generally contain cytochrome c, ubiquinol and a terminal ubiquinol oxidase of the cytochrome o or cytochrome d type (Matsushita 1994). They differ concerning their electron carriers in their

cytoplasmic membranes; *Gluconobacter, Gluconacetobacter*, *Acidomonas, Kozakia* and *Asaia* use ubiquinones of the Q-10 type, whereas *Acetobacter* possesses ubiquinone of the Q-9 type as major ubiquinone (Franke 1999; Yamada 1997). The active centers of the membrane-bound dehydrogenases are oriented into the periplasmic space, enabling AAB to consume exceeding amounts of substrates in sugar, alcohol or acid rich environments. Nevertheless, as their incomplete pathways result in an ineffective energy generation, AAB typically convert only small amounts of the substrate to biomass.

AAB Exopolysaccharide Production and Biosynthesis

Acetic acid bacteria are well known to efficiently convert different carbon sources into a diverse range of extracellular polysaccharides. These EPS can be categorized by their monomer composition into homopolysaccharides (HoPS), which consist of only one type of sugar monomer, and heteropolysaccharides (HePS), which contain different sugar types. These EPSs could be either attached to the cell surface as capsular polysaccharide (CPS), or secreted into the external environment, which are referred as exopolysaccharides (EPS). Polysaccharides are essential for the survival of AAB, for instance in form of slime causing biofilm formation, as an extracellular matrix against dehydration, or in form of a pellicle, which leads to floating cultures of the surface of the media to increase aeration. The most prominent EPS derived from AAB is water-insoluble cellulose, a HoPS that is formed by *Acetobacter* and *Gluconacetobacter* strains during static cultivation, enabling a high aeration state of the self-organizing pellicle. HoPS are formed via a certain polymerizing enzyme that connects specific precursor substrates into a polymer, via a mostly synthase dependent pathway (Schmid 2015). The cellulose formation is performed by a cellulose synthase, catalyzing the polymerization of nucleotide sugar precursors (UDP-glucose) to a cellulose chain. The genes coding for the cellulose synthase subunits are located side by side in one operon, which differ among the AAB (Römling 2005; Wong et al. 1990). Besides cellulose, as a HoPS formed of β-(1→4)-linked glucosyl units, AAB are moreover capable of producing fructans (polymers consisting of β-fructose units), like levan, and α-(1→6)-linked glucans, like dextran (Naessens 2005). In case of the levan formation, a selective binding of sucrose to the responsible 6-β-D-fructosyltransferase (levansucrase) takes place, followed by a cleavage step which results in the release of glucose from the sucrose and the transfructosylation of the remaining fructose moiety to a growing polymer chain (Hernandez et al. 1995). Dextrans are fromed via an enzyme,

initially named dextran dextrinase (DDase) but now referred to as dextrin dextranase, which converts chains of α(1→4)-linked glucosyl residues to new chains of α(1→6)-linked units (Hehre 1951).

Besides the HoPS formation, different AAB strains - such as *A. tropicalis* (Moonmangmee 2002b), *A. aceti* (Moonmangmee 2002a), *Ga. diazotrophicus* (Serrato et al. 2013) and *Ko. xylinus* (Jansson et al. 1993) - also produce strain-specific HePS. Whereas during the HoPS formation, only a single glycosyltransferase-containing operon is involved, the HePS formation includes a higher number of different glycosyltransferases (GT) (Schmid et al. 2014). HePS are formed in the cytoplasm via the sequential addition of nucleoside diphosphate sugars, such as UDP-glucose, UDP-galactose, GDP-mannose and TDP-rhamnose or their derivatives (e. g. GDP-glucuronic acid) to growing repeating units. The over-all assembly of the repeating unit is subsequently initialized via a so-called priming glycosyltransferase (GT), which loads the first sugar nucleotide precursor to a C55 isoprenylphosphate lipid carrier anchored in the inner membrane (Schmid 2015). This first connected sugar monomer serves as acceptor of the next sugar nucleotide and so on, until the repeating unit is formed, which can contain up to four or five different sugar types. The final polymerization reaction occurs at the periplasmic side of the cytoplasmic membrane and requires the transfer of the undecaprenyl pyrophosphate-linked repeat unit across the membrane (Rehm 2010). In Gram-negative bacteria the polymerization of repeating units, constituting the final HePS, is mostly performed via a Wzx/Wzy- dependent pathway. The respective strains using this pathway, carry the genes for a flippase (Wzx), which catalyzes the transport across the membrane, and a polymerase (Wzy), that is involved in the assembly of the HePS (Schmid et al. 2014). The HePS biosynthesis requires one particular glycosyltransferase gene for each included sugar residue. These genes are encoded in organized clusters, that can reach up to 20 genes (Rehm 2010). The most prominent organization of such genes is known as the *gum*-cluster from *Xanthomonas (X.). campestris*, that is involved in the xanthan gum synthesis. The *gum*-cluster comprises all genes coding for GTs and enzymes that are necessary for the polymerization and secretion of xanthan (Becker et al. 1998; Vorhölter 2008). The xanthan synthesis follows a precisely prescribed procedure, in which GumD catalyzes the first step of the HePS synthesis by transferring the first sugar-1-phosphate to an undecaprenyl-phosphate-lipid carrier in the membrane (Chou 1997). This step is folowed by *gumM*, *gumH, gumK* and *gumI* encoded glycosyl-transferases, which sequentially transfer glucosyl-1-phosphate, mannosyl-1-phosphate, glucuronyl acid 1-phosphate and a final mannosyl-1-phosphate residue in a specific order (Pühler 2008; Barreras 2004). Once the

repeating unit is completed, the unit is transported through the membrane into the periplasm and further polymerized. In *X*. *campestris* the flippase (Wzx) and polymerase (Wzy) are coded by the *gumE* (Wzy) and the *gumJ* (Wzx) gene (Tao et al. 2012).
Clusters similar to the *gum*-cluster of *X*. *campestris* could also be found in different AAB. Particularly structurally alike are the HePS cluster of *Ga*. *diazotrophicus* PAI5 (Bertalan et al. 2009) and the acetan cluster of .*Ko*. *xylinus* E25 (Griffin 1994). Both clusters harbor several *gum*-like genes (*gumB, gumC, gumD, gumE, gumF, gumH, gumK &gumM*), suggesting a xanthan-like biosynthesis pathways.

HePS Application

Bacterial exopolysaccharides provide an alternative to polymers of plant origin. Bacteria efficiently convert several carbon sources into a diverse range of polysaccharides with unique chemical and physical properties. Beyond HoPS, especially HePS have unique properties; with their complex, mostly branched structure, they can cause drastic viscosity increases of aqueous solutions already in low concentrations. Several bacterial polymers are produced commerciallyin large-scale fermentations. A prominent example is the commercially important and widely used polysaccharide xanthan, produced by *X*. *campestris* (Baird 1983). Industrial applications of xanthan are very diverse and include areas like foods, cosmetics, as well as oil recovery, in which xanthan is always used as an additive (Nwodo et al. 2012; Rosalam and England 2006). Due to their mostly energy-intensive biosynthesis, HePS are generally formed only in small amounts, making newly discovered HePS more expensive for the industrial production. However, newly discovered microbial HePS can be commercialized when they are used in specific market niches, e.g. directly in situ produced during food fermentation by a deliberately added appropriate starter strain (Raspor 2008). To guarantee the safety of the respective products, it is important to use a non-pathogenic, natural and traditionally food-associated bacteria as a starter culture for the EPS production. One group of food-grade bacteria are acetic acid bacteria, which are main actors during the production of several foods and beverages , for example vinegar, kombucha or kefir (Cleenwerck 2008). They are generally considered as non-pathogenic, however, it should be mentioned that in some specific, hospital-associated cases, AAB were isolated from clinical samples from patients suffering form chronic diseases and/or indwelling devices (Alauzet 2010). As already mentioned, acetic acid bacteria are producers of commercially used homopolysaccharides, such as cellulose and levan.

Additional, they produce different kinds of heteropolysaccharides. Furthermore, some AAB carry HePS and HoPS clusters and produce different types of EPSs simultaneously. The simultaneous production of different kinds of EPS could lead to significantly higher (generally desired) rheological effects, via the combination of different qualities of high and low molecular weight EPSs. Only a few multiple EPS producing bacteria among *Acetobacteraceae* have been identified so far (Couso 1987; Jansson et al. 1993; Kornmann et al. 2003), however, these new candidates could offer great commercial potential.

Thesis Outline

Kozakia baliensis is well known to produce high molecular weight levans from sucrose, which significantly improve the quality of breads (Jakob 2012c; Jakob 2013). Parenthetically analyzed EPS of *K. baliensis* DSM 14400, isolated from sucrose deficient NaG medium, revealed that it consists of several sugar monomers (Jakob 2012b), indicating that both *Kozakia* strains additionally produce a new class of a so far not characterized HePS. The additional formed HePS is a mucous, water-soluble substances, which could exhibit promising properties for diverse (food) biotechnological applications.

The basic aim of the underlying thesis was to gain knowledge of this new, unique HePS formed by *K. baliensis*. Whereas a precise characterization of the *K. baliensis* HePS can only be achieved with an exact genetic description of the HePS clusters, which leads to better insights of the HePS formation and its structural composition. In order to obtain a complete understanding of the novel *K. baliensis* HePS, an investigation of its structure and its physiological significance for the bacterium is essential.

In particular, this thesis comprehends a holistic approach to reveal
(1) the genetic background of the *K. baliensis* produced HePS
(2) the structure of the particular *K. baliensis* HePS
(3) the possible ecological role of the HePS.

The aim was to gain knowledge on the genetic determinants of the *K. baliensis* DSM 14400 and NBRC 16680 HePS formation, received from the whole-genome sequencing data, and to utilize the created knowledge to simplify the complex structure determination. HePS forming

clusters should be identified, whereby the predicted functions of the particular enzymes could reveal the later structure of the *K*. *baliensis* HePS.

We further speculated that a detailed genetic characterization of the *K*. *baliensis* derived HePS could give a better insight into the HePS formation of of acetic acid bacteria in general.

In addition, further HePS forming AAB should be sequenced in order to obtain a comprehensive overview of the *pol-* and *gum*-like genes via comparative analysis of the different strains, in particular, *A*. *aceti* TMW 2.1153, *A*. *persici* TMW 2.1084 and *N*. *chiangmaiensis* NBRC 101099.

Since the structure of a HePS is decisive for his rheological properties, the *K*. *baliensis* NBRC 16680 structure should be analyzed via HPLC, NMR spectroscopy and AF4-MALS-RI. Additionally, structural changes that can be caused by addition of different carbohydrates or magnesium to the medium should be investigated. The received structural data should be further compared with rheological data of the *K*. *baliensis* NBRC 16680 HePS, also with a special focus of a possible effect triggered by magnesium.

For the general characterization of the *K*. *baliensis* HePS, its physiological function should be elucidated. In particular, the role of the HePS with respect to the ethanol and acetic acid tolerance of *K*. *baliensis* should be determined. Since interactions of the HePS habituated assertiveness are assumed to be related with the pellicle formation of *K*. *baliensis*, both should always be considered in context to each other. The physiological role of the HePS and the *K*. *baliensis* pellicle should be clarified via a non-HePS producing and a pellicle deficient mutant strain of *K*. *baliensis* NBRC 16680, using an AAB specific deletion system. Further, via a *polE*-gene knockout, the role of the *polE* gene for the pellicle formation in *K*. *baliensis* should also be elucidated.

On this way, the obtained results should contribute to the gain of knowledge regarding *K*. *baliensis* produced EPS and its physiological relevance in particular and to a better understanding of the genetic organization and HePS formation of AAB in general. Further, the obtained results regarding the specific characterization of involved enzymes can be used for the knowledge-based optimization and engineering of *K*. *baliensis* produced HePS.

2. MATERIAL UND METHODS

2.1 Strains, media and growth conditions

All components used for the cultivation of microorganisms were either autoclaved or sterilized (by filtration) before usage. Sugars were autoclaved separately.

2.1.1 Acetic acid bacteria

A total number of 30 strains of the genera *Gluconobacter* (*G.*), *Gluconacetobacter* (*Ga.*), *Acetobacter* (*A.*), *Asaia*, *Neoasaia* (*N.*) and *Kozakia* (*K.*) were used in the beginning of this study. The bacteria were cultivated aerobically at 30 °C in sodium-gluconate medium (NaG) (20 g/L sodium gluconate, 3 g/L yeast extract, 2 g/L peptone, 3 g/L glycerol, 10 g/L mannitol, optional 20 g/L agar, pH adjusted to 6.0), either in liquid cultures (200 rpm), static cultures or on NaG-plates (20 g/l agar). The media were supplemented, depending on the experimental approach, with fructose (10 g/L), glucose (10 g/L), or magnesium (200 mg/L).
All acetic acid bacteria (AAB) strains used in this work stem from the Technische Mikrobiologie Weihenstephan (TMW) culture collection and are sorted via a TMW-number or, if available, a DSM or NBRC number.

2.1.2 *Escherichia* strains

Escherichia (*E.*) *coli* strains DH5α and Top 10 were cultivated aerobically (180 rpm) at 37 °C in Luria-Bertani (LB) medium (Sambrook 1989) (10 g/L peptone; 5 g/L yeast extract; 5 g/L NaCl). To select recombinant plasmid harboring *E. coli* cells, the LB medium was supplemented with 50 µg/mL kanamycin (Km; AppliChem). Occasionally an additional blue/with screening was performed, by adding 30 mg/mL 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and 100 mM Isopropylthio-β-galactoside (IPTG) to LB media plates.

Table 1 Overview of the used *E. coli* strains.

***E. coli* strain**	**Genotype**	**References**
DH5A	F$^-$, *endA1*, *hsdR17* (rk-mk-), *supE44*, *thi1*, *recA1*, *gyrA* (Nalr), *relA1*, D (*lacZYAargF*), U169, F80*lac*ZDM15	Hanahan (1983)
TOP10	F$^-$, *mcrA*Δ (*mrr-hsd*RMS-*mcr*BC), φ80*lac*ZΔM15, Δ*lacX74*, *recA1*, *ara*D139, Δ(*araleu*)7697, *galU*, *galK*, rpsL (Str^R), *endA1*, *nupG*	Invitrogen, USA

2.1.3 Cryo-conservation and strain verification

Each AAB strain was cultured on NaG plates, directly from the Technische Mikrobiologie Weihenstephan (TMW) culture collection. Afterwards, 10 ml NaG medium was inoculated with a single colony of each strain, respectively, followed by an overnight incubation at 30° C, 200 rpm. The cells were harvested by centrifugation (5000 g, 5 min), suspended in 1.6 mL NaG medium and aliquoted to 800 µL samples. The cell suspension was added with 800 µl of glycerol and the mixture was stored at -80° C. From glycerol cultures, strains were reconstituted for usage by streaking them on mMRS1 or NBB (Döhler, Darmstadt, Germany) agar.

Selected strains were further analyzed using 16S rRNA PCR as phylogenetic marker gene in order to check whether the respective strain was assigned correctly. For the detection of the 16S rRNA gene a degenerated primer pair was used; 616V (5´- AGAGTTTGATYMTGGCTCAG-3´) and 630R (5´- CAKAAAGGAGGTGATCC-3´), followed by sanger sequencing by GATC Biotech (Konstanz, Germany) of the particular PCR product.

2.2 HePS production and isolation

For the heteropolysaccharide (HePS) isolation from liquid media, a single colony of an HePS producing strain was used to inoculate 10 mL of NaG medium. The inoculated strain was incubated for 16 h on a rotary shaker (30°C, 200 rpm) and subsequently used as the starting culture for a following, 10 mL, 24 h incubation. After washing the cells in 1 mL NaG medium, 500 mL of different NaG media with varying carbon sources (10 g/L mannitol (NaG), 10 g/L glucose (NaG-g), 10 g/L fructose and 10 g/L glucose (NaG-fg), 10 g/L glucose, 10 g/L fructose, and 200 mg/L magnesium (NaG-fgm) were inoculated with 500 µL of the 24 h culture, respectively. After aerobic cultivation of the main cultures for 32h at 30°C on a rotary shaker (200 rpm), cells were removed by centrifugation (5500 rpm, 10 min) and the supernatant was mixed with two volumes of cold ethanol (4°C) and stored for 2 days at 4°C to precipitate HePSs (method described by Tieking (2005)). The precipitated HePSs were collected via centrifugation (13.000 rpm, 30 min) and dissolved in ddH_2O, followed by a dialysis step (MWCO 14 kDa) in ddH2O for at least 48 h at 4°C. Finally, the purified HePSs were lyophilized and quantified by weighing.

To obtain large amounts of pellicle EPS, pellicle producing strains were cultured in unmodified NaG medium in cell culture flasks, to ensure a large surface for the oxygen supply. Briefly, 10% of the seed culture was inoculated to 30 mL NaG medium and incubated statically at 30◦C

for 14 days. The pellicle EPS was purified from the culture and cells were separated from EPS via ultra-sonification (10 min) and mechanical disruption, related to the method of Ali et al. (2011b). The culture was centrifuged (10 min, 10.000 g) and the supernatant was saved in another flask. The cell pellet was washed 2 times with 10 mM HEPES buffer (pH 7), and suspended in the same buffer. The suspension was again ultra-sonicated for 10 min, followed by a centrifugation step for 10 min, 13.000 g. The resulting supernatant was combined, with the present supernatant from the firs centrifugation, and EPS was precipitated with cold ethanol (2:1, v/v) and kept overnight at 4°C. This step was repeated three times, followed by a dialysis step (MWCO 14 kDa) of the recovered (centrifugation) and in ddH_2O re-dissolved EPS.

2.3 Structural analysis of EPSs

2.3.1 Determination of the HePS monosaccharide compositions & configuration

To analyze the neutral sugar composition of the isolated HePS, high performance liquid chromatography (HPLC) was used. The purified polysaccharide samples were hydrolyzed with 10,5 % (v/v) of perchloric acid over 7 h at 100°C, followed by a centrifugation step (4°C, 10 min, 13000 g) for removal of possible impurities, such as proteins. The supernatant was analyzed using a Rezex RPM column (Phenomenex, Germany) coupled to a refractive index (RI) detector (Gynkotek, Germany) corresponding to the method of Kaditzky (2008). Sugar monomers were identified according to their retention time using suitable monosaccharide standards (glucose (Glc), galactose (Gal), mannose (Man), rhamnose (Rha)). The mobile phase was water, with a flow rate of 0.6 mL/min. HPLC runs were controlled with Chromeleon Software (Dionex, Germany).

To obtain further information about other HePS sugar components such as uronic acids, the polysaccharides were hydrolyzed by a combination of methanolysis and TFA hydrolysis and analyzed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) (De Ruiter et al. 1992). Briefly, hydrogen chloride in methanol (1.25 M, 2mL) was added to 5-10 mg sample and methanolysis was performed for 16 h at 80°C. An aliquot (40 µL) was evaporated and subsequently hydrolyzed with 2 M TFA for 1 h at 121°C. After evaporation, the samples were redissolved in water and analyzed for the released monosaccharides by HPAEC-PAD on an ICS-5000 system (Thermo Scientific Dionex, CA, USA) using a CarboPac PA-20 column at 25°C. A flow rate of 0.4 mL/min and a gradient composed of (A) bidistilled water, (B) 0.1 M sodium hydroxide, and (C) 0.1 M sodium

hydroxide + 0.2 M sodium acetate were used: Before every run, the column was rinsed with 100% B for 10 min and equilibrated for 20 min with 90% A and 10% B. After injection, the following gradient was applied: 0-1.5 min, from 90% A and 10% B to 96% A and 4% B; 1.5-22 min, isocratic, 96% A and 4% B; 22-32 min from 96% A and 4% B to 100% B; 32-42 min, isocratic, 100% C.

The absolute configuration of the neutral sugars was determined after hydrolysis of the polysaccharides with 2 M trifluoroacetic acid TFA for 30 min at 121°C. Subsequently, the acid was evaporated and the sugars were heated with 150 µM of (*R*)-2-octanol and 5 µL of TFA overnight (Leontein et al. 1978). The solvent was removed and the sample was silylated by using 80 µL of *N*,*O*-bis(trimethylsilyl)trifluoroacetamide and 20 µL of pyridine. The silylated sugar derivatives were analyzed by GC-MS (GC-2010 Plus and GC-MS-QP2010 Ultra, Shimadzu, Japan) equipped with an Rxi-5Sil MS column (30 m x 0.25 mm i.d., 0.25 µm film thickness, Restek, Germany) using the following conditions: initial column temperature, 150°C; ramped at 1°C/min to 200°C; ramped at 15°C/min to 300°C. Injection temperature was 275°C and split injection was used at a split ratio of 10:1. Helium was used as carrier gas at 40 cm/sec and the transfer line was held at 275°C.

2.3.2 Structural analysis via NMR spectroscopy and methylation analysis

The glycosidic linkages were analyzed by methylation analysis as described previously (Nunes et al. 2008; Wefers and Bunzel 2015). Briefly, the polysaccharides were dissolved in 2 mL of dimethyl sulfoxide and freshly ground sodium hydroxide (100 mg) was added. The mixture was incubated for 90 min in an ultrasonic bath and 90 min at room temperature. Methyl iodide (1 mL) was added and the mixture was sonicated for 30 min and incubated for 30 min at room temperature. The methylated polysaccharides were extracted into dichloromethane (3 mL), and the organic layer was washed with 0.1 M sodium thiosulfate (5 mL) and twice with water. The solvent was evaporated and the residue was dried overnight in a vacuum oven at 40°C. After the addition of 2 M TFA (2 mL), the residue was incubated at 121°C for 90 min. The TFA was removed by evaporation, and $NaBD_4$ (20 mg) in aqueous NH_3 solution (2 M) was added. The partially methylated monosaccharides were reduced for 1 h and reduction was terminated by adding glacial acetic acid (100 µL). While cooling on ice, 450 µL of 1-methylimidazole and 3 mL of acetic anhydride were added, and the sample was incubated for 30 min at room temperature. Water (3 mL) was added, and the partially methylated alditol acetates (PMAA) were extracted with 5 mL of dichloromethane. The organic phase was washed with water (5

mL, three times) and the residual water was removed by freezing overnight at -18°C. The PMAAs were analyzed on a GC-MS system (GC-2010 Plus and GCMS-QP2010 SE instruments, Shimadzu, Japan) equipped with an Rxi-5Sil MS column (30 m x 0.25 mm i.d., 0.25 µm film thickness, Restek). The following conditions were used: Initial column temperature, 140°C, held for 2 min; ramped at 1°C/min to 180°C, held for 5 min; ramped at 10°C/min to 300°C, held for 5 min. Helium was used as carrier gas at 40 cm/sec. Split injection with a split ratio of 30:1 was used, and the injection temperature was 250°C. The transfer line was held at 275°C and electron impact spectra were recorded at 70 eV. Semiquantitative analysis of the PMAAs was performed by using a GC-FID system (GC-2010 Plus, Shimadzu) equipped with a DB-5MS column (30 m x 0.25 mm i.d., 0.25 µm film thickness, Agilent Technologies, CA). The same conditions as described for the GC-MS analyses were used, with the exception of a reduced split ratio of 10:1. The FID temperature was 240°C and nitrogen was used as makeup gas. Molar response factors were used to calculate the portions of the PMAAs (Sweet et al. 1975).

Carboxyl reduction of the uronic acids was performed as described previously (Pettolino et al. 2012). The samples were dissolved in 1 mL of bidistilled water, followed by the addition of MES (200 µL, 0.2 M, pH 4.75) and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide-metho-*p*-toluene sulfonate (400 µL, 500 mg/mL). The solution was mixed and incubated for 3 h at room temperature. While cooling on ice, 1 mL of 4 M imidazole (pH 7.0) was added. Subsequently, 1 mL of aqueous $NaBD_4$ solution (70 mg/mL) was added and the sample was incubated for 3 h at room temperature. The reaction was terminated by the addition of glacial acetic acid (500 µL) and the sample was dialyzed for 24 h against water (Molecular Weight Cut Off: 3500 Da). After freeze drying, the carboxyl-reduced samples were methylated, hydrolyzed, reduced, and acetylated as described above.

NMR spectroscopic analysis was carried out on a Bruker DRX-500 spectrometer equipped with a 5mm probe (Bruker, Germany). The HePS derived from NaG-fgm was dissolved in D_2O (0.8 % w/v) and spectra were acquired at 350 K. The HDO signal was suppressed by using the zgpr pulse sequence. Dimethyl sulfoxide was used for spectrum calibration (2.71 ppm, (Gottlieb et al. 1997)).

2.3.3 Separation of isolated HePS and size determinations via AF4-MALS_RI

Asymmetric flow field flow fractionation coupled to multi-angle light scattering (AF4-MALS; Wyatt Technology) was used to separate HePS samples and to subsequently analyze their size

distributions. Isolated HePS samples were first dissolved in ddH_2O by vigorous vortexing (0.1 mg/mL) and afterwards centrifuged at 14000 g for 15 min to remove undissolved particles, which hindered AF4 separations in previous experiments (data not shown). For sample separations 100 µl of the obtained supernatants were injected into the AF4-channel, respectively. The flow conditions and materials/equipment for separation of HePS samples were identical to those previously described by Ua-Arak (2015). HePS separations were analyzed using ASTRA 6.1 Software (Wyatt Technology), while the integrated "particle mode" and the model "random-coil" allowed evaluation of separated fractions regarding their root-mean square (rms) radii size distributions. High performance size exclusion chromatography (HPSEC) was performed on an HPLC system (L-7100 pump, L-7490 RI detector, Merck/Hitachi, Germany) equipped with a TSKgel G5000PWXL column (300 x 7.8 mm, 10 µm particle size) and a TSKgel G3000PWXL column (300 x 7.8 mm, 7 µm particle size) in series. Elution was carried out with bidistilled water at a flow rate of 0.6 mL/min and 60°C.

2.3.4 Rheological measurements

For rheological measurements of the isolated HePS, serial dilutions were prepared in ddH_2O (1,5%, 1%, 0,75%, 0,5%, 0,25%, w/v). The HePS dilutions were stirred for 30 min at room temperature and heated twice for 20 minutes at 80°C. Steady shear rheological data were obtained with a Rheometer (Physica MCR 501 Rheometer (Anton Paar, Austria)) using a double gap geometry (DG 26.7-SS, Anton Paar, Austria) at a constant temperature of 20°C. The shear rate profile moved along a logarithmic scale beginning with a starting sequence including a shear rate from 0.1 to 1,000 s^{-1} followed by a reverse sequence from 1,000 to 0.1 s^{-1}. Per decade, five values were recorded over a period of 10 seconds. The actual measurement was preceded by an adaptation phase of 30 seconds.

To describe the steady shear rheological properties of the samples, the data were fitted to the well-known power law model (Eq. 1) which is used extensively to describe the flow properties of non- newtonian liquids (Barnes 1989).

$$\sigma = K\gamma^n \quad (1)$$

where σ is the shear stress (Pa), γ is the shear rate (s-1), K is consistency index (Pa sn), and n is the flow behavior index (dimensionless). The steady shear rheological measurements were conducted in biological triplicates. The reported results were expressed as an average of the three measurements.

2.4 Genomics

2.4.1 Genome sequencing

For genome sequencing, high molecular weight DNA was isolated from six acetic acid bacteria strains, following the instructions of the Quiagen Genomic DNA Kit (Qiagen, Hilden, Germany). The genomic DNA of *A. aceti* TMW 2.1153, *A. cerevisiae* TMW 2.1084, *G. albidus* TMW 2.1191, *K. baliensis* DSM 14400, *K. baliensis* NBRC 16680 and *N. chiangmaiensis* NBRC 101099 were submitted to GATC Biotech (Germany) for PacBio single-molecule real-time (SMRT™) sequencing (Eid 2009; McCarthy 2010). A single library was prepared for each of the six strains and an insert size of 8–12 kb was selected for library creation, resulting in at least 200 Mb raw data from 1 SMRT cell (1×120 min movies) applying P4-C2 chemistry.

2.4.2 Genome assembly

Raw data were assembled via a hierarchical genome-assembly process, version 3 (HGAP3), whereas in the first step, the longest reads are used for assembling of a contiguous genome. After this preassembly process, shorter reads improve the accuracy of the long reads. Continuative, the preassembled reads are combined by the Celera Assembler to a long continuous contig, whereas the Celera Assembler assumes the genome to be linear, so the assembly can be slightly larger than the reference because the ends of the single contig that cover a genome can have overlaps. A final algorithm called Quiver polishes the assembly for final consensus accuracies, by taking into account all underlying data (Chin et al. 2013). The resulting assemblies were confirmed to specific quality criteria including; N50 read length, mean read score (0.85 or higher), pre-assembly yield (0.4 or higher), number of contigs, coverage and average consensus concordance (higher than 99%). The polished assemblies were split into contigs using BioPerl (http://www.bioperl.org) and further tested for their redundancy using NCBI BLAST (Altschul 1990; Camacho 2009). Moreover, the accuracy of the genome assembly was evaluated with BridgeMapper (RS_BridgeMapper), which is part of the SMRT-Analysis software. In order to receive circularized genomes, contigs were checked for overlapping ends with the dotplot tool Gepard (Krumsiek 2007). In addition, they were observed with SMRT-View 2.30 (Pacific Biosciences, Menlo Park, USA) for noticeable changes in the coverage and mapping quality (polishing), while a decrease in mapping quality at both ends indicated overlapping ends. Contigs with overlapping ends were further circularized by introducing an in silico break into the contig, followed by a reassembly and cirularization step using minimus2 (AMOS, http://amos.sourceforge.net). The resulting

circularized contigs were tested with Gepard for their accuracy followed by an additionally alignment (NCBI-BLAST) against the original contigs, in order to prevent a loss of the initial sequence information. The circularized contigs of one bacterial chromosome, as well those contigs where a circularization was not possible, were merged together in a final fasta-file and were further utilized as a reference for SMRT-Analysis resequencing, using the RS-Resequencing_1 protocol. A continuous repetition of the resequencing process took place, until the consensus accordance was at a value 100 %. Each resequencing job was examined with respect to quality criteria mentioned above, using SMRT-Analysis statistics and SMRT-View. The consensus sequence of each final genome was stored as fasta-file, which served as input for following genome analysis applications, including annotation.

2.4.3 Genome annotation and submission

Initial ORF predictions and annotations of the circularized genomes were accomplished automatically using the program RAST (Rapid Annotations using Subsystems Technology), a SEED-based, prokaryotic genome annotation service (Aziz 2008; Overbeek et al. 2014).
Further, genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (Angiuoli 2008). For NCBI genome submission a bioproject (PRJNA311264) was created, including six single biosamples for each of the obtained AAB genome sequences, respectively. The submission procedure was done as described online (http://www.ncbi.nlm.nih.gov/genbank/genomesubmit) and genomes were submitted without RAST annotations. Past submission locus tags and protein IDs were requested from the NCBI or assigned automatically. EPS cluster annotations were further verified via homology searches against GeneBank/EMBL using the function BLASTP (Altschul 1990; Camacho 2009). Whole genomic sequence data of the six sequenced AAB strains (see 2.4.1) have been deposited in GeneBank.

2.5 Genetics

2.5.1 General techniques

For general genomic DNA isolation the E.Z.N.A. Bacterial DNA Kit (Omega Bio-tek, Norcross, USA) was used. Plasmid DNA was purified using the GeneJET Plasmid Miniprep Kit (ThermoFisher, Germany). Preparative DNA isolations from agarose gels were performed with the peqGOLD Gel extraction Kit (peqlab, Germany), PCR products were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA). Restriction endonuclease

digestions, dephosphorylation with shrimp alkaline phosphatase (SAP) and ligations with T4 DNA ligase were performed as recommended by the suppliers (Fermentas, Germany). Unknown DNA sequences were amplified using Taq-DNA- Polymerase (Qbiogene, USA) or Phire Hot Start II DNA Polymerase (ThermoFisher, Germany). For amplification of known DNA sequences as well as 16S rRNA genes the proofreading Phusion High-Fidelity DNA Polymerase (Finnzymes, Finland) was used. When specific primer sets were put in the reaction mixture, PCR cycle conditions were routinely performed according to the manufacturers instructions. Oligonucleotide primers for PCR reactions were obtained from MWG Biotech AG (Germany). PCR products were sequenced by GATC Biotech (Germany). Cloning experiments were designed and documented using Clone Manager™ 9 Professional Edition.

2.5.2 Deletion system

For plasmids preparation, the GeneJET Plasmid Miniprep Kit (Thermo Fisher scientific, Waltham, USA) was used. Genomic DNA from *K. baliensis* NBRC 16680 R (Brandt 2016) was extracted with the E.Z.N.A. Bacterial DNA Kit (Omega Biotek, Norcross, USA) and DNA purification was done with the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA). Restriction enzymes, DNA ligase, and alkaline phosphatase (FastAP) were obtained from Fermentas (Waltham, USA). PCRs were performed according to the Phusion High-Fidelity DNA Polymerase manuals from New England Biolabs (Frankfurt, Germany). For construction of the deletion vector, a fusion PCR technique was used to ligate the PCR products of flanking regions according to a long flanking homology (LFH) protocol (Peters 2013; Wach 1996). The length of the homology sequences were 20 bp. Primers are listed in Table 2. An enzyme-free cloning technique (Tillett 1999) was used for the further construction of the deletion vector, with the pKOS6b plasmid as basis. *K. baliensis* NBRC 16680 R was transformed via electroporation (Creaven et al. 1994; Hall et al. 1992; Krajewski 2010; Mostafa et al. 2002) with Gene Pulser Xcell™ Electroporation Systems from Bio Rad (München, Germany). Therefore, cells were inoculated to an OD_{600} of 0.3 in NaG medium and finally grown to an OD_{600} of 0.9. The culture was centrifuged at 5000 g, at 4 °C for 10 min, and washed three times in 1 mM HEPES buffer (pH 7). Cells were resuspended in 1 mM HEPES buffer, supplemented with ¼ volume of glycerin and shock frozen in 50 µL aliquots. The electroporation took place in cuvettes with 2 mm electrode distance from Bio Rad (München, Germany). The electroporation was carried out under constant conditions: 2.5 kV, 25 µF, and 400 Ω. Fresh enriched NaG medium (450 mM mannitol, 15 g/L yeast extract, 15 mM $CaCl_2$, 10 mM $MgSO_4$ and 6 mM glycerin) was added immediately after the pulse. The treated cells were incubated

on a rotary shaker over 14 h and subsequently plated on NaG plates containing 50 µg/mL kanamycin for the first selection step with an incubation time of 48 h. The obtained colonies were saved for cryo-conservation and transferred to 5-fluorocytosin (FC, 60 µg/mL) NaG plates. Phire Hot start DNA polymerase (Thermo Fisher scientific; Waltham, USA) was used for colony PCR reactions, to screen for mutants or to confirm integration of the deletion vector into the genome. PCR products were sequenced via sanger sequencing by GATC Biotech (Konstanz, Germany).

Table 2 Primers used for the construction and verification of the *polE* deletion system of *K. baliensis* NBRC 16680 R Restriction sites in bold. Primer flanking homology sequences are underlined.

Primer	Sequence (5`-> 3`)	Primer location [bp]	Product size
G4F_Fw	ACGACGCGCATGAGTAAACC	1.471.189 – 1.471.208	
TE_Rv	CAACCCGTGATTTGAACTCC	mobile element	
P1_*polE*_KpnI_Fw	AAGCT**GGTACC**GCACTTTACCGAAGCGAAGGGATG	2.545.632 – 2.545.666	
P2_*polE*_Rv	CATCCGGGCGGAAGCCGCCTTTCGCGTGGAAGATCTG	2.546.552 – 2.546.570	957 bp
P3_*polE*_Fw	CAGATCTTCCACGCGAAAGGCGGCTTCCGCCCGGATG	2.547.627 – 2.547.645	
P4_*polE*_XbaI_Rv	CGA**TCTAGA**CTCCAAAGGCGAGCCGGTTTATAATG	2.548.552 – 2.548.586	963 bp
pK18MCS-Fw	TTCCGGCTCGTATGTTGTG	pKOS6b	
pK18MCS-Rv	CAGCTGGCAATTCCGGTTC	pKOS6b	
CL_*polE*_Fw	TGCCCGCATCTGAACATCGC	2.545.569 – 2.545.588	
CL_*polE*_Rv	CGAGCCGGTTTATAATGACG	2.548.549 – 2.548.568	
CL2_*polE*_Fw	CTGGTCCCAGATTTGGATTC	2.545.442 – 2.545.461	179 bp
P3P4_*polE*_Rv	AACCTATTGGGCTGTCGTTG	2.547.969 – 2.547.950	1471 bp

2.5.3 Identification of a mobile element insertion in *K. baliensis* NBRC 16680

For the genetic characterization of a mutated, rough colony strain of *K. baliensis* NBRC 16680, genomic DNA of the mutant strain (*K. baliensis* NBRC 16680 R) was isolated following the instructions of the E.Z.N.A. Bacterial DNA Kit (Omega Biotek, Norcross, USA). DNA sequences of specific EPS clusters from both strains were amplified using Taq-DNA-Polymerase (Qbiogene, USA) or KAPA HiFi PCR Polymerase (Peqlab, Erlangen, Germany). Primers used for PCR reactions are listed in Table 4. Resulting PCR fragments were compared between both strains and further restriction maps of the respective PCR fragments of the corresponding EPS cluster were created, using specific restriction endonuclease combinations (Table 3). Restriction endonuclease (RE) digestions of the received PCR products were performed as recommended by the suppliers (Fermentas, St. Leon-Roth, Germany). To genotype a detected polymorphism of the fourth, divergent segment of the *gum*-cluster of *K. baliensis* NBRC 16680 R, primer walks were performed with several primers (Table 4),

followed by a PCR reaction including a primer set with extensions, to exclude an incorrect primer binding. In order to enable sequencing of the PCR products of segment 4 of the *gum*-cluster of *K*. *baliensis* NBRC 16680 R, restriction enzymes were tested which cut at a single position, exclusively in the PCR fragments of *K*. *baliensis* NBRC 16680 R. Following restriction enzymes were used: HpaI, ApaI, DraI, SalI, XhaI; SacI, AgeI, EcoRI, EcoRV, HindIII, NheI, and SmaI. After a successful restriction with HpaI the two resulting fragment could be sequenced via sanger sequencing by GATC Biotech (Konstanz, Germany). Preparative DNA isolations from agarose gels were performed with the peqGOLD Gel extraction Kit (Peqlab Erlangen, Germany), PCR products were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA). An insertion of a transposable element with an associated loop formation was examined by restriction with a T7-Endonuclease I (New England BioLabs, Beverly, MA) and further PCR reactions with mobile element specific primers Table 4.

2.5.4 Computational analysis

Gene sequences were generally edited and analyzed in terms of ORFs, RE digestion, dyad symmetries and energy calculations for potential hairpin structures using Clone Manager 5 (Sci Ed Central, USA). Protein alignments were created using CLUSTALW (EMBL-EBI, UK).

Table 3 PCR segments and corresponding restriction sites of the *kps*-, *pol*- and *gum*-cluster of *K. baliensis* NBRC 16680

Cluster segments	*Segment size [bp]*	*Restriction enzyme*	*Fragments [bp]*	*Fragments size [bp]*
kps-cluster				
Segment 1	4416	EcoRV	1 – 153	153
625.648 – 630.064 bp		SmaI	154 – 1960	1807
			1961 – 2710	750
			2711 – 2958	248
			2959 – 4305	1347
			4306 - 4416	111
Segment 2	4660	EcoRV	1 – 370	370
630.044 – 634.704 bp			371 – 1327	957
			1328 – 1628	301
			1629 – 2192	564
			2193 – 2267	75
			2268 – 3289	1022
			3290 – 4485	1196
			4486 – 4632	147
			4633 - 4660	28
Segment 3	4422	EcoRV	1 – 1082	1082
634.685 – 639.107 bp		HindIII	1083 – 1286	204
			1287 – 1640	354
			1641 – 3128	1488
			3129 – 3923	795
			3924 – 4401	478
			4401 - 24422	21
pol-cluster				
Segment 1	4351	AscI	1 – 968	969
2.543.246 – 2.547.597 bp		BglII	970 – 1273	304
		BspHI	1274 – 2667	1394
			2668 – 2771	104
			2772 – 3308	537
			3309 – 4126	818
			4127 – 4292	166
			4293 - 4351	59
gum-cluster				
Segment 1	4001	HindIII	1 – 106	106
1456282 – 1460282 bp		NruI	107 – 2090	1984
			2091 – 3054	964
			3055 - 4001	947
Segment 2	4009	HindIII	1 – 605	605
1460262 – 1464270 bp		NruI	606 – 716	111
			717 – 3272	2556
			3273 - 4009	737
Segment 3	3310	HindIII	1 – 158	158
1464251 – 1467560 bp		NruI	159 – 1196	1038
			1197 – 1778	582
			1779 – 2003	225
			2004 – 3245	1242
			3246 - 3310	65
Segment 4	4703	NruI	1 – 593	593
1467546 – 1472248 bp		AflII	594 – 653	60
		EcoRV	654 – 1554	901
			1555 – 1954	400
			1955 – 4477	2523
			4478 - 4703	226
Segment 5	4016	NruI	1 – 1201	1201
1472249 – 1476264 bp		EcoRV	1202 – 1617	416
			1618 – 2241	624
			2242 - 4016	1775
Segment 6	2617	EcoRV	1 – 1616	1616
1476238 – 1478854 bp			1617 – 2537	921
			2538 - 2617	80

Table 4 Primers used for the identification of genotypic aberrations and mobile element insertion in *K. baliensis* NBRC 16680 R

Primer	Sequence (5`-> 3`)	Primer location	Product size	Segment
Kps_S1_Fw	GGCAGTGTTTCCGGCTTGAG	625.648 – 625.468 bp	4416 bp	segment 1
Kps_S1_Rv	CGTAAACGGCGTGCCCTAAC	630.044 -630.064 bp		segment 1
Kps_S2_Fw	CGTTAGGGCACGCCGTTTAC	630.044 -630.064 bp	4660 bp	segment 2
Kps_S2_Rv	CCACCGTGAGCCGGATAATC	634.684 - 634.704 bp		segment 2
Kps_S3_Fw	GATTATCCGGCTCACGGTGG	634.685 – 634.705 bp	4422 bp	segment 3
Kps_S3_Rv	GCTTCGATAACCGCAACACC	639.687 - 639.107 bp		segment 3
Pol_S1_Fw	TACGCCGCGTCGAAGATAC	2.543.246 - 2.543.266 bp	4351 bp	segment 1
Pol_S1_Rv	TTATCTCGGTTGGCGAGAAG	2.547.577 - 2.547.597 bp		segment 1
Gum_S1_Fw	GACCGAGCGTCCAGTTATTC	1.456.282 – 1.456.302 bp	4001 bp	segment 1
Gum_S1_Rv	GGATGCTGCGCACTAATGGC	1.460.262 – 1.460.282 bp		segment 1
Gum_S2_Fw	CGCCATTAGTGCGCAGCATC	1.460.262 – 1.460.282 bp	4009 bp	segment 2
Gum_S2_Rv	CGATGCTGGTGCGGGTAATG	1.464.250 – 1.464.270 bp		segment 2
Gum_S3_Fw	CATTACCCGCACCAGCATCG	1.464.551 – 1.464.571 bp	3310 bp	segment 3
Gum_S3_Rv	CGGTCTTCAATCCTGTTCTG	1.467.540 – 1.467.560 bp		segment 3
Gum_S4_Fw	CAGGATTGAAGACCGTACTC	1.467.246 – 1.467.266 bp	4703 bp	segment 4
Gum_S4_Rv	AAACGGGTCCTTTCGAGTCC	1.472.228 – 1.472.248 bp		segment 4
Gum_S5_Fw	TCTTCCGGCAACCGCGTATC	1.472.249 – 1.472.269 bp	4016 bp	segment 5
Gum_S5_Rv	GATCGGCGTCGTTCCAACTG	1.476.244 – 1.476.264 bp		segment 5
Gum_S6_Fw	TCGCGTGCAGTTGGAACGAC	1.476.238 – 1.476.258 bp	2617 bp	segment 6
Gum_S6_Rv	ACATCACCAGCACGTAGAAC	1.478.834 – 1.478.854 bp		segment 6
Primer	**Sequence (5`-> 3`)**	**Primer location**	**Product size**	**Corresponding Primer**
TE_Fw	ACCAGATGGCCAACAGCAAC	mobile element		Gum_S4_Rv
TE_Rv	GATTGAGCGATCATCGTCTG	mobile element		G4F_Fw
TE_A_Fw	GGCCGTCCGACCATTCAATC	77.284 – 77.303 bp	198 bp	TE_Rv
TE_B_Fw	AAACCGCCTTGGCCGTTCTG	1.430.163 – 1.430.144 bp	147 bp	TE_Rv
TE_C_Fw	ACGCCACATAATCCGCATAG	2.279.157 – 2.279.176 bp	250 bp	TE_Rv
G4B_Fw	AATGCCTGAGGGAACATA	1.468.226 – 1.468.243 bp		Gum_S4_Rv
G4C_Fw	ACGCGACATTGGAGTCTTTC	1.469.281 – 1.469.300 bp		Gum_S4_Rv
G4D_Fw	TATCGGTCGTCCCATAGAAC	1.469.724 – 1.469.743 bp		Gum_S4_Rv
G4E_Fw	GTTGCGCGTCATCTAAGTTC	1.470.646 – 1.470.665 bp	1603 bp	Gum_S4_Rv
G4F_Fw	ACGACGCGCATGAGTAAACC	1.471.189 – 1.471.208 bp	1060 bp	Gum_S4_Rv
G4G_Fw	TTATCGCATTGCGCGTGCTC	1.471.598 – 1.471.617 bp	561 bp	Gum_S4_Rv
G4E_ex_Fw	GGCGATGTCAAATCGGGTTG CGCGTCATCTAAGTTC	1.470.646 – 1.470.665 bp	1635 bp	Gum_S4_ex_Rv
Gum_S4_ex_Rv	CTGGCCGCTTAATCGAAACG GGTCCTTTCGAGTCC	1.472.228 – 1.472.248 bp		
Extension_Fw	GGCGATGTCAAATCGG		1635 bp	Extension_Rv
Extension_Rv	CTGGCCGCTTAATCG			

2.6 Growth behavior analysis

2.6.1 Mutation frequency of *K. balie*nsis NBRC 16680 in different media

Growth dynamics were studied during cultivation of *K. bali*ensis NBRC 16680 in NaG medium, containing different supplements. The carbon source in NaG media, mannitol, was replaced respectively by 10 g/L sucrose (NaG-s), 10 g/L sucrose, 10 g/L fructose and 10 g/L glucose (NaG-fgs) or 10 g/L fructose, 10 g/L glucose and 200 mg/L magnesium (NaG-fgm). Furthermore, stress conditions were tested, via addition of 3% ethanol (NaG-EtOH), or 0,4% acetic acid (NaG-AA), to unmodified NaG-medium. *K. baliensis* NBRC 16680 was first cultivated overnight in 10 ml of unmodified NaG medium at 30 °C. About 1×10^8 CFU/mL seed culture were transferred to 10 ml of modified NaG medium, respectively. The flasks were incubated at 30 °C with rotary shaking at 200 rpm for 48 h and samples adducted at 0, 24 and 48 h. Cell numbers of wildetype and rough colonies (*ΔgumD*) were counted. To describe the spontaneous mutation rates, the mutation frequency was calculated (Eq. 1) (Foster 2006).

$f = r/\mathrm{N}$ (1)

where *r* is the observed number of mutants in a culture (cfu/mL), N is the number of total cells (cfu/mL), *f* is the mutant fraction or frequency (dimensionless).
For targeting the transposon insertion side, random colony PCRs of respective rough colonies were carried out with Phire Hot start DNA polymerase (Thermo Fisher scientific; Waltham, USA). A primer set of a genomic primer (G4F_Fw) and a primer, targeting the mobile element (TE_Rv) were used; primers are listed in Table 4. PCR products were subsequently sequenced via sanger sequencing by GATC Biotech (Konstanz, Germany).

2.6.2 Growth dynamics of different *K. baliensis* strains in acetic acid and ethanol

*K. bali*ensis DSM 14400, NBRC 16680, the *ΔgumD* mutant (Brandt 2016) and the *ΔpolE* mutant (see 2.5.2) were grown on NaG agar plates, with either ethanol or acetic acid, in different concentrations. The ethanol containing plates ranged between 1- 10% of ethanol, and the acetic acid plates, between 0.1% - 1% of acetic acid. Each strain was scattered onto NaG plates directly from cryo cultures and incubated for 3 days at 30°C.

Furthermore, a static cultivation in NaG medium with 3% ethanol and 0,6% acetic acid was carried out. The *K. baliensis* NBRC 16680 strain, *ΔgumD* mutant (Brandt 2016) and the *ΔpolE* mutant (see 2.5.2) were grown as seed cultures in unmodified NaG medium, overnight. Cultures were inoculated with an OD_{600} 0.3 and cultivated up to 0.9, respectively. For static cultures,

300 µl of the seed culture were inoculated into 3 ml NaG medium and cultivated statically at 30 °C. Cells were harvested by centrifugation at 6000 g and cell pellets were dried overnight at 120 °C. The dry weight was measured each day, over a time span of 7 days.

2.6.3 Microscopy

Microorganisms were microscopically investigated using an Axiostar Plus microscope (Zeiss, Germany).

3. RESULTS (THESIS PUBLICATIONS)

PART I **PAGES 38 - 50**	**Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis***

HePS are in the focus of attention, concerning their unique properties. Their distinct characteristics are eligible for the food and cosmetic industry as thickening additives, wherefore HePS forming bacteria receive high attention. However, HePS can be also directly produced in situ during food fermentation by an intentionally added HePS producing starter strains. Hence, our superordinated goal was to identify a new promising biotechnologically relevant HePS, produced by an AAB, a group of non-pathogenic, food-grade bacteria. *Kozakia* (*K*.) *baliensi*s is a relatively new member of AAB, which produces high amounts of a new, uncharacterized EPS, from mannitol and glycerol as (main) carbon sources. This study aims to identify the structure, as well as the genetic background of two selected *K*. *baliensis* EPSs (DSM 14400, NBRC 16680). For this purpose, a double-track method was used, based on the combination of structural analysis via HPLC, together with whole genome sequencing of both selected *K*. *baliensis* strains. Via whole genome sequencing of *K*. *baliensis* DSM 14400 and NBRC 16680, circularized genomes could be established and typical EPS forming clusters were identified in both strains. Beside HoPS forming clusters, like complete ORFs coding for levansucrase genes in both strains and a cellulose synthase cluster in *K*. *baliensis* DSM 14400, also HePS forming clusters were tracked. In addition to the *polABCDE* cluster, which is responsible for the pellicle formation, a new *gum*-like cluster, related to the well-characterized *gum* cluster coding for the xanthan synthesis in *X*. *campestris,* was identified. The identified *gum*-like cluster showed as well similarity to clusters from other AAB such as *Ga*. *diazotrophicus* and the acetan cluster of *Ko*. *xylinus*. Further, comparative genomics got adjuvant insights into the biosynthesis of this novel *gum*-like HePS, and allowed to establish a hypothetical composition of the HePS. The HPLC analysis revealed that the secreted *K*. *baliensis* HePS is composed of glucose, galactose and mannose, respectively for each strain (DSM 14400, NBRC 16680). This is in agreement with the predicted sugar monomer composition based on the in silico genome analysis of the respective clusters. The accuracy of the in silico genome analysis was moreover confirmed by the discovery of a mutant strain of *K*. *baliensis* NBRC 16680 with a deficient HePS production. PCR analysis and following sanger sequencing unveiled a transposon insertion in front of the

gumD gene, located in the *gum*-like cluster of the mutant strain, which indicated the essential role of *gumD* and of the associated *gum* genes for the HePS production of *K*. *baliensis* NBRC 16680. Thus it can be shown, that the genomic approach applied in this work is enormously time saving and efficiently supports future chemical analyses such as NMR analysis, which is depict in the following section, for final elucidation of these complex *K*. *baliensis* HePS structures.

Authors contributions: Julia U. Brandt was responsible for the independent planning and conducting the experiments. Julia U. Brandt evaluated the data and wrote the main text of the manuscript. Frank Jakob was the project leader responsible for supervision. Rudi Vogel conducted the conceptual setup and supervision of the study including infrastructural demands and final editing of the manuscript. Andi Geissler and Jürgen Behr were involved in the analysis of the whole genome sequencing data.

Brandt *et al. Microb Cell Fact (2016) 15:170*
DOI 10.1186/s12934-016-0572-x

Microbial Cell Factories

RESEARCH Open Access

Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*

Julia U. Brandt, Frank Jakob*, Jürgen Behr, Andreas J. Geissler and Rudi F. Vogel

Abstract

Background: Acetic acid bacteria (AAB) are well known producers of commercially used exopolysaccharides, such as cellulose and levan. *Kozakia (K.) baliensis* is a relatively new member of AAB, which produces ultra-high molecular weight levan from sucrose. Throughout cultivation of two *K. baliensis* strains (DSM 14400, NBRC 16680) on sucrose-deficient media, we found that both strains still produce high amounts of mucous, water-soluble substances from mannitol and glycerol as (main) carbon sources. This indicated that both *Kozakia* strains additionally produce new classes of so far not characterized EPS.

Results: By whole genome sequencing of both strains, circularized genomes could be established and typical EPS forming clusters were identified. As expected, complete ORFs coding for levansucrases could be detected in both *Kozakia* strains. In *K. baliensis* DSM 14400 plasmid encoded cellulose synthase genes and fragments of truncated levansucrase operons could be assigned in contrast to *K. baliensis* NBRC 16680. Additionally, both *K. baliensis* strains harbor identical *gum*-like clusters, which are related to the well characterized *gum* cluster coding for xanthan synthesis in *Xanthomanas campestris* and show highest similarity with *gum*-like heteropolysaccharide (HePS) clusters from other acetic acid bacteria such as *Gluconacetobacter diazotrophicus* and *Komagataeibacter xylinus*. A mutant strain of *K. baliensis* NBRC 16680 lacking EPS production on sucrose-deficient media exhibited a transposon insertion in front of the *gumD* gene of its *gum*-like cluster in contrast to the wildtype strain, which indicated the essential role of *gumD* and of the associated *gum* genes for production of these new EPS. The EPS secreted by *K. baliensis* are composed of glucose, galactose and mannose, respectively, which is in agreement with the predicted sugar monomer composition derived from in silico genome analysis of the respective *gum*-like clusters.

Conclusions: By comparative sugar monomer and genome analysis, the polymeric substances secreted by *K. baliensis* can be considered as unique HePS. Via genome sequencing of *K. baliensis* DSM 14400 + NBRC 16680 we got first insights into the biosynthesis of these novel HePS, which is related to xanthan and acetan biosynthesis. Consequently, the present study provides the basis for establishment of *K. baliensis* strains as novel microbial cell factories for biotechnologically relevant, unique polysaccharides.

Keywords: Acetic acid bacteria, *Kozakia baliensis*, Exopolysaccharides, Genome analysis, Gum-like cluster, Heteropolysaccharides

Background

The production of exopolysaccharides (EPSs) is a common attribute of many bacteria. On the basis of their monomer composition, EPSs are divided into two groups, homo- (HoPS) and heteropolysaccharides (HePS). These EPSs can be used as an extracellular matrix, in form of slime for biofilm formation and protection, or in form of a pellicle, which leads to floating cultures of the surface of the media to increase aeration. Especially HePS have unique properties, since their complex, mostly branched structures are responsible for drastic viscosity increases of aqueous solutions already in low concentrations. This characteristic of HePS is exploited in the food and cosmetic industry (e.g. in sauces, dressings, tooth paste, lotions etc.) always on an empirical basis. Therefore,

*Correspondence: frank.jakob@wzw.tum.de
Technische Universität München, Lehrstuhl für Technische Mikrobiologie, Gregor-Mendel-Straße 4, 85354 Freising, Germany

EPS forming bacteria receive high attention for biotechnological applications, particularly *Xanthomonas (X.) campestris*, which produces the commercially important and widely used polysaccharide xanthan [1]. Industrial applications of xanthan are widely diversified and include areas like, foods, cosmetics as well as oil recovery [2, 3].

As an alternative to the addition of EPS as ingredients of food, EPS can also be produced in situ upon food fermentation by deliberately added appropriate starter strains [4]. In the respective products, the safety of the used starter culture and its metabolic products are of great importance and limit in situ EPS production to non-pathogenic, naturally and traditionally food-associated bacteria (e.g. in yoghurt, sourdough, kefir or kombucha). One group of non-pathogenic, food-grade bacteria comprises acetic acid bacteria (AAB), which have important roles in food and beverage production, for example vinegar, kombucha or kefir [5].

AAB are well known for their ability to produce large amounts of EPSs, either homopolysaccharides, like dextrans, levans and cellulose, or different kinds of heteropolysaccharides [6], such as in the case of *Acetobacter (A.) tropicalis* [7], *A. aceti* [8], *Gluconacetobacter (Ga.) diazotrophicus* [9] and *Komagataeibacter xylinus* [10].

Heteropolysaccharides are formed in the cytoplasm via the sequential addition of nucleoside diphosphate sugars to growing repeating units. The assembly of the repeating unit is subsequently initialized via a so-called priming glycosyltransferase (GT), which loads the first sugar nucleotide to a C55 isoprenylphosphate lipid carrier anchored in the inner membrane [11]. This first sugar connected to the C55 carrier serves as acceptor of the next sugar nucleotide and so on, till the repeating unit is formed. The biosynthesis of these repeating units requires different kinds of GTs that are genetically encoded in organized clusters, which can reach up to 20 genes [12]. One distinct organization of such genes is known as the *gum*-cluster from *X. campestris*, involved in xanthan synthesis. This cluster includes all genes coding for GTs and enzymes necessary for the polymerization and secretion of xanthan [13, 14]. In Gram negative bacteria the secretion and polymerization of HePS producing bacteria is mostly performed via a Wzx/Wzy-dependent pathway, whereas Wzx acts as a flippase, catalyzing the transport across the membrane, and Wzy is involved in the assembly of the HePS [15]. In *X. campestris* these two enzymes are coded by the *gumE* (Wzy) and the *gumJ* (Wzx) gene [16].

Furthermore, a few AAB strains produce different types of EPSs simultaneously. This simultaneous production of different kinds of EPS could lead to significantly higher (mostly desired) rheological effects, via the combination of different qualities of high and low molecular weight EPSs. Only a few multiple EPS producing bacteria among *Acetobacteracea*e have been identified so far [10, 17, 18], whereas these new candidates could offer great commercial potential (e.g. as EPS producing starter cultures for food fermentations) due to their specialization on EPS production.

Kozakia (K.) baliensis is a relatively new member of the family of *Acetobacteraceae* and is already well known to produce high molecular weight levans from sucrose, which significantly improve the quality of breads [19] [20]. Throughout cultivation of two *K. baliensis* strains (DSM 14400, NBRC 16680) on sucrose deficient media, we found that both strains still produce high amounts of mucous, water-soluble substances, which could exhibit promising properties for diverse (food) biotechnological applications.

Therefore, we wanted to detect the basic sugar composition of these polymeric substances via sugar monomer analysis and identify responsible biosynthesis clusters and putative transferases catalyzing the incorporation of specific sugars via whole genome sequencing of both *Kozakia* strains. In this way, first insights should be got into (i) the basic composition of these EPS and (ii) their respective biosynthesis routes.

Methods

Strains, media and growth conditions

Kozakia strains (DSM 14400^T, NBRC 16680) were screened for their ability to produce mucous substances on modified sodium-gluconate medium (NaG) agar (without sucrose) over a time span of 72 h at 30 °C. Both *K. baliensis* strains were generally cultivated aerobically at 30 °C in liquid NaG media (20 g/L sodium gluconate, 3 g/L yeast extract, 2 g/L peptone, 3 g/L glycerol, 10 g/L mannitol, pH adjusted to 6.0).

General molecular techniques

For genome sequencing, genomic DNA was isolated following the instructions of the Qiagen Genomic DNA Kit (Qiagen, Hilden, Germany). The genomic DNA of *K. baliensis* strains DSM 14400 and NBRC 16680 were submitted to GATC Biotech (Germany) for PacBio single-molecule real-time (SMRT) sequencing, respectively. A single library was prepared for both strains, which were run on one SMRT cell, respectively. All generated sequences were assembled with a hierarchical genome-assembly process version 3 (HGAP3), including an assembly with the Celera Assembler and assembly polishing with Quiver [21]. Initial ORF predictions and annotations were accomplished automatically using the program RAST, a SEED-based, prokaryotic genome annotation service [22]. Annotations were corrected via the NCBI Prokaryotic Genome Annotation Pipeline. EPS cluster annotations were performed

via homology searches against GeneBank/EMBL using the function BLASTP. Whole genomic sequence data of both *K. baliensis* strains (DSM 14400, NBRC 16680) have been deposited in GeneBank.

For genetic characterization of a mutant strain of *K. baliensis* NBRC 16680, genomic DNA of the mutant strain was isolated following the instructions of the E.Z.N.A. Bacterial DNA Kit (Omega Biotek, Norcross, USA). DNA sequences were amplified using Taq-DNA-Polymerase (Qbiogene, USA) or KAPA HiFi PCR Polymerase (Peqlab, Erlangen, Germany). Primers used for PCR reactions are listed in Additional file 1: Table S1. Restriction endonuclease digestions were performed as recommended by the suppliers (Fermentas, St. Leon-Roth, Germany). Following restriction enzymes were used: HpaI, ApaI, DraI, SalI, XhaI; SacI, AgeI, EcoRI, EcoRV, HindIII, NheI, and SmaI. PCR products were sequenced via sanger sequencing by GATC Biotech (Konstanz, Germany). Preparative DNA isolations from agarose gels were performed with the peqGOLD Gel extraction Kit (Peqlab Erlangen, Germany), PCR products were purified using the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA).

EPS production and isolation

To isolate EPS samples from liquid media, the respective strains were grown in 10 mL of NaG media and incubated for 32 h on a rotary shaker (200 rpm) at 30 °C. After cell removal, EPS in the supernatant was precipitated with cold ethanol (2:1, v/v) and kept overnight at 4 °C. This step was repeated three times, followed by a dialysis step (MWCO 14 kDa) of the recovered (centrifugation) and in ddH_2O re-dissolved EPS. Finally, the purified HePSs were lyophilized and quantified by weighing.

Determination of sugar monomers in isolated EPS samples

The monosaccharide composition of isolated EPS was investigated via high performance liquid chromatography (HPLC). The purified polysaccharide samples were hydrolyzed with 15 % of perchloric acid (70 %) over 7 h at 100 °C, followed by a centrifugation step (4 °C, 10 min, 13,000*g*) for removal of possible impurities such as proteins. The supernatant was analyzed using a Rezex RPM column (Phenomenex, Germany) coupled to a refractive index (RI) detector (Gynkotek, Germany) according to the method of [23]. Sugar monomers were identified according to their retention time using suitable monosaccharide standards (D-glucose, D-galactose, D-mannose, D-rhamnose). The mobile phase was water, with a flow rate of 0.6 mL/min.

Comparison of pellicle and EPS production between *K. baliensis* NBRC 16680 wildtype and mutant strain

To investigate surface pellicle production of an EPS deficient mutant strain of *K. baliensis* NBRC 16680, 5 mL of static cultures were inoculated with 50 µl of overnight cultures (wildtype and mutant, respectively) and cultivated statically at 30 °C for 3 days. Pellicle production was observed and documented macroscopically. Furthermore, EPS production in shaking cultures was investigated via cultivation of the mutant strain of *K. baliensis* NBRC 16680 over 32 h in NaG media. Isolation and quantification of EPS was performed according to the described EPS isolation (see above).

Results

Analysis of EPS composition and production

For EPS extraction, *K. baliensis* strains (DSM 14400, NBRC 16680) were grown for 32 h in 10 mL of NaG liquid medium without sucrose. After freeze-drying of the isolated EPS, the amount was measured by weight, which resulted in about 1.87 ± 0.04 g/L (DSM 14400) and 1.71 ± 0.05 g/L (NBRC 16680) of EPS for each strain. Both isolated EPS showed good solubility in water after freeze-drying. Concentrations of ~5 g/L (0.5 % aqueous solution) resulted in drastic viscosity increases. In the HPLC pattern of the perchloric acid hydrolysate of both *K. baliensis* EPS, three peaks could be observed (Fig. 1). The retention times of the detected monomers were consistent with the retention times of the standards D-glucose (Glc), D-galactose (Gal) and D-mannose (Man).

Global genome properties of *K. baliensis* DSM 14400 and NBRC 16680

Both *K. baliensis* genomes were sequenced via PacBio single-molecule real-time (SMRT) sequencing. De novo assembly was carried out by using the hierarchical genome assembly process (HGAP) method [21]. For *K. baliensis* DSM 14400 seven contigs were generated upon sequencing, which were further assembled to one finished genome. The genome consists of one circularized chromosome (2,888,029 bp) and four circularized plasmids designated as pKB14400_1 (253,508 bp), pKB14400_2 (102.455 bp), pKB14400_3 (176,911 bp) and pKB14400_4 (18,625 bp), as well as two partial plasmids [pKB14400_5 (35.463 bp), pKB14400_6 (30.533 bp)]. Genome sequencing of *K. baliensis* NBRC 16680 resulted in six contigs, which were assembled to one chromosome (2,807,246 bp) and three plasmids: pKB16680_1 (135,727 bp), pKB16680_2 (80,099 bp) and pKB16680_3 (36,022 bp), as well as two partial plasmids [pKB16680_4 (14.635 bp), pKB16680_5 (11.848 bp)]. The average GC content of the DSM 14400 strain is 57.4 %, and 57.7 % for the NBRC 16680 strain. Both *Kozakia* chromosomes show high conformity, the average nucleotide identities are 98 % within a query coverage of 94 %. The plasmid identity is low up to unique plasmids in *K. baliensis* DSM 14400 (Fig. 2).

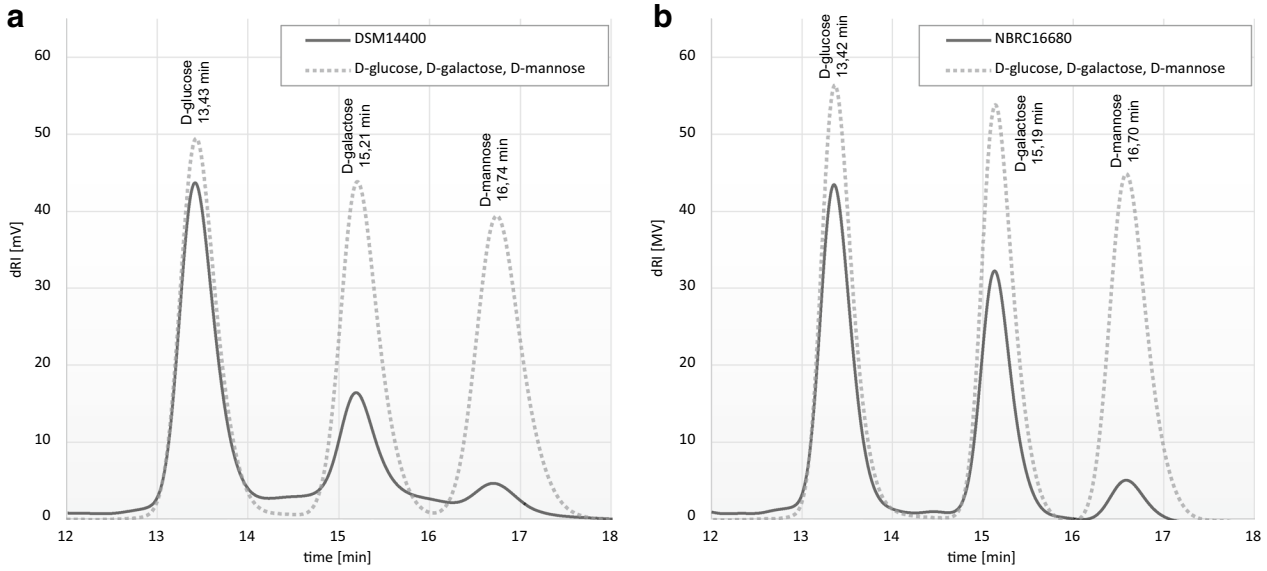
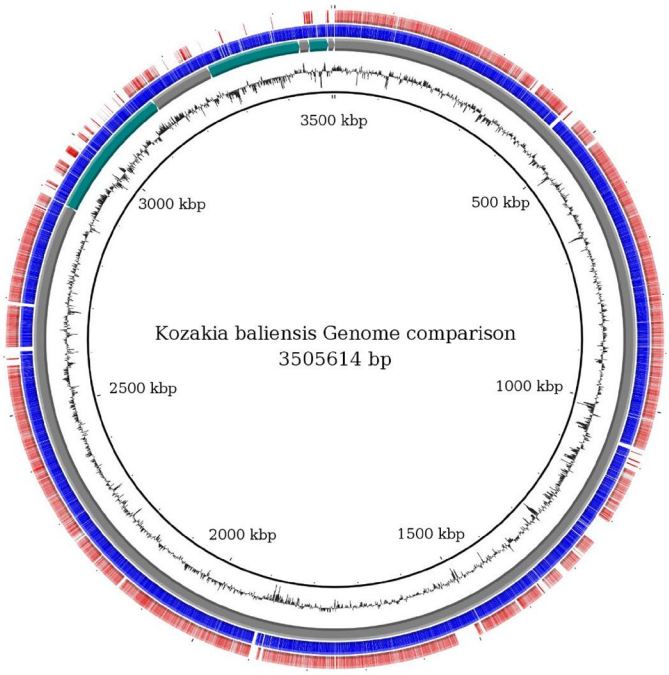


Fig. 1 Monomer composition of the HePS isolated from *K. baliensis* strains DSM 14400 and NBRC 16680. HPLC profiles of the perchloric acid hydrolysates of HePS from *K. baliensis* DSM 14400 (**a**) and *K. baliensis* NBRC 16680 (**b**) in comparison with a 25 mM standard mix (*dotted line*) composed of D-glucose (13,43 min), D-galactose (15,21 min) and D-mannose (16,74 min). The retention times of the detected sugar monomers in the respective acid hydrolysates were consistent with the retention times of the standards

Clusters coding for putative HePS biosynthesis

In both *Kozakia* genomes typical clusters for putative HePS production were identified, including a HePS gene cluster of 25 kb ("*gum*-like cluster", Fig. 3) and a "*pol*-cluster" (Fig. 5b), which comprises the genes *polABCDE* and was previously shown to be involved in pellicle formation in *Acetobacter tropicalis* via the biosynthesis of capsular HePS [24]. The genetic organizations of the *gum*-like clusters of both *K. baliensis* strains are depicted in Fig. 3a and exemplarily compared to *gum*-like clusters of *Komagataeibacter (Ko.) xylinus* E25 (Fig. 3b) (formerly *Ga. xylinus*), *Ga. diazotrophicus* PA15 (Fig. 3c) and *X. campestris* (Fig. 3d). While the *gum*-like clusters of *K. baliensis* share identical genetic organizations among each other, they are next related to *gum*-like clusters from the AAB strains *Ko. xylinus* and *Ga. diazotrophicus*. On the contrary, the genetic organization of the gum-cluster of *X. campestris* differs remarkably from those of the depicted AAB strains (Fig. 3), while the gum proteins of *X. campestris* exhibit principal homology to those of the depicted AAB strains (Table 1). The *gum*-like clusters of both *Kozakia* strains involve in total 19 ORFs, which are mainly designated as glycosyltransferases with unknown function or hypothetical proteins (11 genes), and eight proteins, which show homology to the well-characterized *gum*-proteins catalyzing xanthan biosynthesis in *X. campestris.* These homologous gum-like genes are *gumB, -C, -D, -E, -H, -J, -K,* and *–M* (Fig. 3, Table 2). The *gumD* gene is described to catalyze the first step of the HePS synthesis, by transferring the first sugar-1-phosphate to an undecaprenyl-phosphate-lipid carrier in the membrane [25]. *gumH, gumK* and *gumM* encode glycosyltransferases, which are involved in the sequential transfer of mannosyl-1-phosphate, glucosyl-1-phosphate and glucuronyl acid 1-phosphate residues (GlcA), from activated sugar nucleotides, including UDP-glucose, UDP-glucuronic acid and GDP-mannose, respectively [26, 27]. *GumE* and *gumJ* are assigned to have a function during the polymerization and translocation of the repeating units [26, 28]. *GumB* and *gumC* share sequence similarities to the *Escherichia* (*E.*) *coli* proteins Wza and Wzc and therefore could be involved in the export of the HePS [29, 30]. Not all genes of the *gum*-operon from *X. campestris* are present in *K. baliensis* (*gumF, -G, -I* and *–L*, Fig. 3). These genes are *inter alia* associated with the incorporation of acetyl- and pyruvyl-residues into the xanthan backbone [13].

Via comparison of the genetic organizations and protein homologies of the gum-like clusters (Fig. 3, Table 1), the gum-like clusters of *K. baliensis* are most related to that of *Ga. diazotrophicus*, which produces a HePS composed of Glc, Gal, Man in approximate ratios of 6:3:1 [31]. This is in agreement with the detected sugar monomers in the isolated HePS of *K. baliensis* DSM 14400 and NBRC 16680 (Fig. 1). Though, there are different genes in *K. baliensis*/*Ga. diazotrophicus* gum clusters (e. g "*gumF*" putatively incorporating acetyl residues, Fig. 3b). Furthermore, a homology comparison between the previously



Strain	BioSample	Accession Numbers	Genome coverage	contigs	G + C content	Total number of ORFs	ORFs with assigned function	Ribosomal RNAs	tRNAs
K. baliensis DSM14400	SAMN04396914	CP014674 – CP014680	65	7	57,4 %	3366	3290	23S (5 copies) 16S (5 copies) 5S (5 copies)	57
K. baliensis NBRC16680	SAMN04396915	CP014681 – CP014686	72	6	57,7 %	2936	2860	23S (5 copies) 16S (5 copies) 5S (5 copies)	57

Fig. 2 Genome comparison and general features of *K. baliensis* strains DSM 14400 and NBRC 16680. Starting from inside: *circle 1* shows the general genomic position in kilobases; *circle 2* depicts the varying G+ C-content of *K. baliensis* DSM 14400 at different genetic loci; *circle 3* is composed of the seven contigs of *K. baliensis* DSM 14400 [main chromosome and additional (partial) plasmids]; *circle 4* reflects the coding density of *K. baliensis* DSM 14400; *circle 5* shows the blast identities (*red*) of *K. baliensis* NBRC 16680 in comparison to *K. baliensis* DSM 14400 (note the low identity in plasmid regions)

described acetan biosynthesis cluster from *Ko. xylinus* E25 and the gum-like clusters of *K. baliensis* revealed several "*ace-genes*" to be homologous to the gum-like genes of *Kozakia* (Fig. 3c, Table 1).

The *gum*-like clusters of both *Kozakia* strains are separated from genes, which encode enzymes necessary for the synthesis of specific activated nucleotide precursors, including UDP-glucose, UDP-galactose and GDP-mannose. These genes are located at different genomic positions in both *K. baliensis* strains. These sequences were at first automatically annotated in the course of genome annotations and afterwards assigned to a proposed biosynthesis pathway for activated nucleotide sugars, which again was reconstructed from schemes involving essential enzymes for the respective biosynthesis of activated sugars [26, 18] (Fig. 4).

Enzymes and clusters coding for putative HoPS biosynthesis

In both strains complete ORFs coding for levansucrases of the glycoside hydrolase 68 family (GH 68) could be detected. Both *K. baliensis* strains possess one chromosomally encoded levansucrase gene, respectively, which share identical sequences and exhibit highest similarities to *Ko. xylinus* levansucrase (AB034152) (Fig. 5a). *K. baliensis* DSM 14400 additionally harbors a plasmid-encoded levansucrase, which shares highest similarity to *Ga. diazotrophicus* levansucrase *lsdA* gene, but seems to be inactive (interrupted) due to the insertion of a mobile element in the N-terminal domain. Similarly to *Ga. diazotrophicus*, a (partial) type II dependent secretion operon is associated with the interrupted levansucrase in

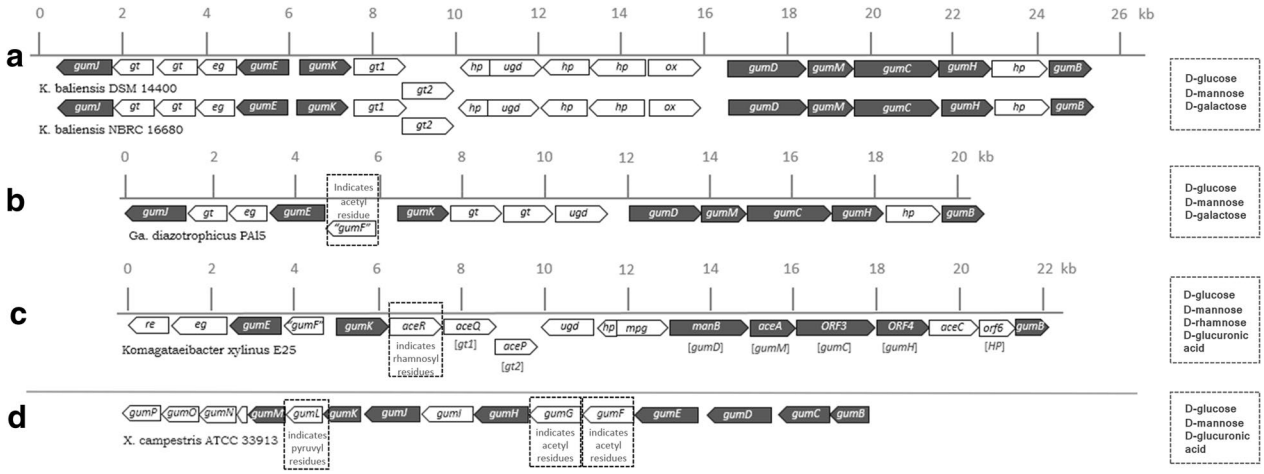


Fig. 3 Genetic organization of HePS biosynthesis encoding *gum*-like clusters. The *gum*-like clusters of *K. baliensis* strains DSM 14400 + NBRC 16680 are depicted in (**a**). The cluster has an overall size of ~25 kb and involves 19 genes, including glycosyltransferases (*gt*), hypothetical proteins (*hp*), and eight *gum* like genes (*gumB, -C, -D, -E, -H, -J, -K, and –M*), which are marked in *grey*. Furthermore, the cluster contains a putative endoglucanase (*e.g.*), oxidoreductase (*ox*) and a UDP-glucose dehydrogenase (*ugd*). **b**, **c** show the related *gum* like clusters of the AAB strains *Ga. diazotrophicus* PAl5 and *Komagataeibacter xylinus* E25. The *Ga. diazotrophicus* cluster exhibits, in comparison to both *K. baliensis* clusters, an additional *gumF* gene, that could putatively incorporate acetyl-residues at specific positions into the related HePS. The so called acetan cluster of *Ko. xylinus* harbors—besides an additional *gumF* gene—a rhamnosyl transferase, coded by *aceR*, as well as a mannose-phosphate-guanyltransferase (*mpg*). The nomenclature for the acetan cluster in (**c**) is based on Griffin AM, Morris VJ and Gasson MJ [44], while brackets under the particular genes mark the homologous *gum* genes. In (**d**) the *gum* cluster of *X. campestris,* which consists of *gumB, -C, -D, -E, -F, -G, -H, -I, -J, -K, -L, -M, -N and –P*, is depicted. The *dotted squares* in the particular *gum*-clusters of B, C and D mark genes, which alter between the specific clusters (relative to both *K. baliensis* gum-like clusters). The corresponding monomer compositions of the respective HePSs are shown at the *right*. The putative functions of the corresponding annotated genes are as follows (derived from Pühler et al. [26] and Griffin et al. [45]): *aceA* UDP-glucosyltransferase, *aceC* GDP-mannosyltransferase, *aceP* glucosyltransferase, *aceQ* glucosyltransferase, *aceR* rhamnosyltransferase, *gumJ* export protein, *gumE* polymerization or export protein, *gumK* catalyzes the addition of glucuronic acid, *gumD* catalyzes the addition of glucose-1-phosphate, *gumM* catalyzes the addition of glucose in β-1,4-position, *gumC* polymerization and export protein, *gumG* acetyl transferase, *gumL* pyruvyl transferase, *gumH* catalyzes the addition of internal mannose, *gumI* β-mannosyltransferase, *gumB* polymerization and export protein, RE dTDP-4-dehydrorhamnose 3,5-epimerase, *manB* Mannose-1-phosphate guanylyltransferase, *tp* transporter

K. baliensis DSM 14400. Though, no levanase gene (*lsdB*) could be detected downstream of the (interrupted) *lsdA* gene of *K. baliensis* (Fig. 5c).

Moreover, we identified a cellulose synthase operon on plasmid 3 of *K. baliensis* DSM 14400 (pKB14400_3), including genes encoding the three cellulose synthase subunits A, B and C, as well as a diguanylate cyclase (DGC) and phosphodiesterase (Fig. 5d).

Genetic characterization of a HePS deficient mutant of *K. baliensis* NBRC 16680

We identified a spontaneous mutant of *K. baliensis* NBRC 16680, which exhibited an altered rough colony morphology on solid NaG agar and was not able to secrete HePS in shaking cultures in contrast to the wildtype strain (Fig. 6a, b). Nevertheless, both types were still able to form a pellicle, floating on the media surface of a static culture (Fig. 6c), which was shown to be dependent of functional Pol proteins in *A. tropicalis* [24]. To identify possible mutations in the respective *pol* (Fig. 5b) and *gum*-like clusters of *K. baliensis* NBRC 16680, PCR screenings covering these genomic regions were performed, respectively. No mutations were observed in the *pol*-clusters of the mutant strain. The 25 kb *gum*-like cluster was divided into six segments (approximate sizes of about 4–5 kb), which were amplified via PCR reactions (primers listed in Additional file 1: Table S1). In segment four, which includes two hypothetical proteins (1,466,953–1,467,825 bp, 1,469,289–1,467,916), an oxidoreductase (1,469,488–1,470,672 bp) and a part of the *gumD* gene (1,471,468–1,472,248 bp), a larger PCR product as expected (6100 bp) was observed in case of the mutant strain. Sanger sequencing of the larger PCR amplicon yielded no positive results, possibly due to a transposon insertion, leading to a hairpin loop formation, which could hinder sequencing [32]. Therefore, several restriction enzymes were tested to perform a single cut in the larger PCR product of the rough strain for possible interruption of the loop structure of a putative transposon. Restriction of the larger PCR product with HpaI allowed sequencing of the obtained restriction fragments (possibly because of disruption of an energy-rich

Table 1 Homology comparison between Gum proteins of *K. baliensis* DSM 14400 and Gum proteins of HePS biosynthesis related bacteria

Kozakia baliensis DSM 14400	Kozakia baliensis NBRC 16680	Gluconacetobacter diazotrophicus PAI 5	Komagataeibacter xylinus E25	Xanthomonas campestris str. ATCC 33913
Protein	Protein sequence identities & query cover	Protein sequence identities & query cover	Protein sequence identities & query cover	Protein sequence identities & query cover
GumJ A0U89_05160;502aa	498/502 (99 %); 100 % A0U90_06730 "GumJ"	322/510 (63 %); 98 % GDI2535 "GumJ"		102/456 (22 %); 96 % XCC2446 GumJ
GumE A0U89_05140;416aa	410/416 (99 %); 100 % A0U90_06750 "GumE"	226/393 (58 %); 94 % GDI2538 "GumE"	203/411 (49 %); 98 % H845_814 "GumE"	96/332 (29 %); 74 % XCC2451 GumE
GumK A0U89_05135;373aa	370/373 (99 %); 100 % A0U90_06755 "GumK"	244/369 (66 %); 98 % GDI2542 "GumK"	244/369 (66 %); 98 % H845_816 "GumK"	162/369 (44 %); 91 % XCC2445 GumK
GumD A0U89_05095;420aa	414/420 (99 %); 100 % A0U90_06795 "GumD"	239/420 (57 %); 100 % GDI2547 "GumD"	234/474 (49 %); 94 % H845_822 "AceA" (Griffin,1994)	159/486 (33 %); 96 % XCC2452 GumD
GumM A0U89_05090;271aa	268/271 (99 %); 100 % A0U90_06800 "GumM"	175/253 (69 %); 92 % GDI2548 "GumM"	137/248 (55 %); 91 % H845_823 "ORF3" (Griffin,1994)	66/230 (29 %); 85 % XCC2443 GumM
GumC A0U89_05085;722aa	716/722 (99 %); 100 % A0U90_06805 "GumC"	351/708 (50 %); 97 % GDI_2549 "GumC"	319/709 (45 %); 95 % H845_824 "ORF4" (Griffin,1994)	111/446 (25 %); 95 % XCC2453 GumC
GumH A0U89_05080;341aa	332/341 (99 %); 100 % A0U90_06810 "GumH"	223/370 (60 %); 98 % GDI2550 "GumH"	190/341 (56 %); 99 % H845_825 "AceC" (Griffin,1994)	162/372 (44 %); 97 % XCC2448 GumH
GumB A0U89_05070;189aa	189/189 (99 %); 100 % A0U90_06820 "GumB"	107/179 (60 %); 94 % GDI2552 "GumB"	102/183 (56 %); 95 % H845_828 "GumB"	50/145 (34 %); 66 % XCC2454 GumB

stem loop structure) with primers G4F_Fw and P4.2_Rv (Fig. 7). In this way, a transposon insertion in front of the *gumD* gene of the rough strain could be identified, as well as the transposon itself. Via a further PCR reaction with specific transposon specific primers in both directions (TE_Fw & TE_Rv), the exact location of the transposon insertion could be identified. The transposon is located in front of the *gumD* gene, while a short region of the starting sequence of *gumD* had duplicated (direct repeat, DR). The transposon insertion possibly leads to an interruption of the promotor region of these gene and, subsequently, a total inactivation of HePS production and secretion (Fig. 6b). The identical nucleotide sequence of the mobile element could be detected thrice in the genome of *K. baliensis* NBRC 16680, each chromosomally located: 77,364–78,676 bp, 1,430,134–1,428,821 bp, 2,279,289—2,280,601 bp. The transposon shows furthermore high similarity to parts of a putative transposon of *G. oxydans* 621H (GOX1325) [Identities: 771/827(93 %), Query: (98 %)].

Discussion

Acetic acid bacteria are well known to produce relatively large quantities of EPS. These can be either HoPS or HePS, but only a few AAB are noted to produce both kinds of the EPS simultaneously. Strains of *Ko. xylinus* (formerly *Ga. xylinus*) produce for example a water insoluble cellulose, levan as well as a water soluble HePS named acetan [17, 10, 18]. *K. baliensis* is a new candidate within the family of AAB that was investigated in this study to produce as well multiple EPSs.

Characterization of novel HePS from *K. baliensis* via comparative genome and sugar monomer analyses

Kozakia (K.) baliensis DSM 14400 is known to produce high amounts of high molecular weight levan with sucrose as the main carbon source [19, 20]. Besides the well-studied levan, we elucidated the identity and the genetic background of so far not characterized HePS from two *K. baliensis* strains in this work, via sugar monomer analysis of isolated EPS and genome sequencing/EPS biosynthesis cluster annotation of/in the respective strains. Both EPS consist of glucose, galactose and mannose. Via comparison of the genomes of *K. baliensis* with related AAB and comprehension of known literature about EPS biosynthesis in AAB and the commercial xanthan producer *X. campestris*, we were able to get new insights into HePS biosynthesis in *K. baliensis*. The identified *gum*-like clusters from *K. baliensis* are highly similar to other HePS producing AAB, like *Ga. diazotrophicus* PA15 (Fig. 3b) and *Ko. xylinus* E25 (Fig. 3c). *Ga. diazotrophicus* harbors – in comparison to *Kozakia*—the most related HePS biosynthesis cluster, which also includes several *gum*-like genes (*gumB, gumC, gumD, gumE, gumH, gumJ, gumK* &

Table 2 Overview of *gum*-genes and their corresponding predicted protein functions involved in xanthan biosynthesis in *X. campestris* [26]

gene	Predicted Protein functions for *Xanthomonas campestris* (overview from Pühler, 2008)	*Kozakia baliensis* DSM 14400 & NBRC 16680	*Gluconacetobacter diazotrophicus* PAI 5	*Komagataeibacter xylinus* E25
	xanthan composition D–glucosyl-, D–mannosyl- and D–glucuronyl acid residues	HePS composition D–glucosyl-, D–mannosyl- and D–galactosyl residues	HePS composition D–glucosyl-, D–mannosyl- and D–galactosyl residues	HePS composition D–glucosyl-, D–mannosyl-, D-rhamnosyl- and D–glucuronyl acid residues
gumB	polymerization and export of xanthan (similarity to Wza protein of *E. coli*)	*gumB*	*gumB*	*gumB*
gumC	export of xanthan (similarity to Wzc protein of *E. coli*)	*gumC*	*gumC*	*gumC*
gumD	Transfers a glucosyl-1-phosphate residue from UDP-glucose to an undecaprenylphosphate lipid carrier	*gumD*	*gumD*	*gumD*
gumE	polymerase function (similaritys to Wzy protein of *E. coli*)	*gumE*	*gumE*	*gumE*
gumF	acetylates a specific mannose moiety of xanthan		*gumF*	*gumF*
gumG	acetylates a specific mannose moiety of xanthan			
gumH	transfers a mannosyl residue from GDP-mannose to a lipid-linked disaccharide by establishing a β-1,3-glycosidic link	*gumH*	*gumH*	*gumH*
gumI	transfers a mannosyl residue from GDP-mannose to a lipid-linked tetrasaccharide by establishing a β-1,2-glycosidic link			
gumJ	polymerase function (similarity to Wzx protein of *E. coli*)	*gumJ*	*gumJ*	
gumK	adds a glucuronyl residue from UDP-glucuronate to a lipid-linked glucose–glucose–mannose trisaccharide	*gumK*	*gumK*	*gumK*
gumL	adds a pyruvyl group to the last-added mannose			
gumM	transfers a glucosyl residue frum UDP-glucose to an undecaprenyl–diphosphate–glucose, resulting in a β-1,4-linked glucose-disaccharide	*gumM*	*gumM*	*gumM*

The presence of homologous *gum*-genes in *K. baliensis*, *Ga. diazotrophicus* and *Ko. xylinus* is indicated by white arrows

gumM, derived from the well characterized xanthan biosynthesis) and produces a HePS that consists of D-glucose, D-galactose and D-mannose as well [31]. *Ko. xylinus* produces "acetan", which consists of D-glucose, D-rhamnose, D-mannose and D-glucuronic acid [10, 33, 34]. Via comparison between both *K. baliensis* clusters and the above mentioned AAB clusters with the well characterized *gum*-cluster of *X. campestris*, genes could be identified, which code for enzymes that are at least necessary for incorporation of some of the HPLC analyzed monomers glucose and mannose into the HePS of *K. baliensis* DSM 14400/NBRC 16680 (Table 2). Though it has to considered, that some further present sugar transferases with unknown function could catalyze the incorporation of these sugar monomers at certain positions with certain linkage types. Since the HePS from *X. campestris* and *Ko. xylinus* do not contain galactose [17], one or more of the homologous glycosyltransferases solely present in *K. baliensis* and *Ga. diazotrophicus gum*-like clusters could function as specific galactosyltransferases (Fig. 3). Additionally, further differences among the compared AAB *gum*-like clusters could be identified (Fig. 3). The *Ko. xylinus* E25 "acetan" cluster is flanked by a mannose-1-phosphate guanylyltransferase *manB*, which is separately located from the HePS cluster in both *K. baliensis* genomes. In contrast to *K. baliensis*, a *gumJ* gene, necessary for the translocation of the HePS across the membrane [11], is missing in the acetan cluster. However, this could be functionally replaced by the exclusively in the acetan cluster present *aceC* gene, which contains a MATE Wzx-like domain, that could act as a flippase (see also "Background" section), assisting the membrane translocation of acetan. Via homology comparison between the *aceP* and *aceQ* gene of the *Ko. xylinus* E25 cluster with the glycosyltransferase 1 and 2 genes of both *K. baliensis* strains (Fig. 3), these could be identified as additional glucosyltransferases, as shown for *Ko. xylinus* by Ishida T, Sugano Y and Shoda M [35]. Furthermore, the cluster from *Ko. xylinus* involves a rhamnosyltransferase gene (*aceR*) as well as a *gumF* like gene (Fig. 3), coding for a protein that incorporates acetyl residues, which both are not included in the respective *K. bali*ensis clusters. This is in agreement with the structural analysis

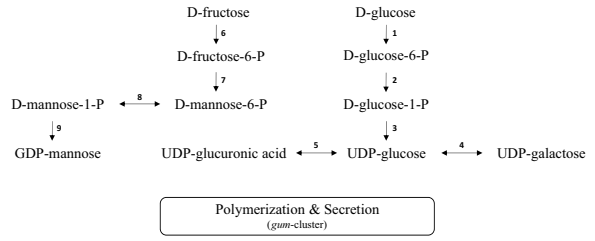


Fig. 4 Schematic representation of the proposed nucleotide sugar biosynthesis related to EPS production in *K. baliensis*. Starting with the phosphorylation of fructose to fructose-6-phosphate (6) or glucose to glucose-6-phosphate (1), these intermediates can be converted into mannose-6-phosphate (7) or glucose-1-phosphate (2), respectively. Mannose-6-phosphate can be further converted into mannose-1-phosphate (8) and finally into GDP-mannose. UGP (3) catalyzes the synthesis of UDP-glucose from glucose-1-phosphate. UDP-glucose can be further isomerized to UDP-galactose (4) or UDP-glucuronic acid (5). The proposed pathway for the biosynthesis of activated nucleotide sugar precursors is based on publications from Kornmann et al. [18] and Pühler et al. [26]. The corresponding genomic locations of these respective genes are listed in Additional file 2: Table S2: *gk*-gene coding for a Glucokinase (EC 2.7.1.2); 2: *pgm*-gene coding for a Phosphoglucomutase (EC 5.4.2.2); 3: *ugp*-gene coding for an UDP-glucose-1-phosphate uridylyltransferase (EC 2.7.7.9); 4: *galE*-gene coding for an UDP-glucose-4-epimerase (EC 5.1.3.2); 5: *ugd*-gene coding for an UDP-glucose dehydrogenase (EC 1.1.1.22); 6: *fk*-gene coding for a Fructokinase (EC 2.7.1.4); 7: *mpi*-gene coding for a Mannose-6-phosphate isomerase (EC 5.3.1.8); 8: *pmm*-gene coding for a Phosphomannomutase (EC 5.4.2.8); 9: *mpg*-gene coding for a Mannose-1-phosphate guanyltransferase (EC 2.7.7.22)

of both *K. baliensis* HePS, which contain no rhamnose, while the missing of acetyl-groups in both *Kozakia* HePS has still to be investigated. On the contrary, the most similar *gum*-like cluster of *Ga. diazotrophicus* (in comparison to *Kozakia*) also includes the putative acetyltransferase *gumF*, indicating additional acetyl residues. This shows that - although the *gum*-like HePS clusters of *K. baliensis* and *Ga. diazotrophicus* are similarly organized, most of the respective encoded proteins exhibit relatively high homology among each other and HPLC analysis revealed these EPS to be composed of the same sugar monomers—the differences in specific gene sets could lead to specific structures and properties of these EPS. These observations nicely demonstrate the high value of the sequencing based function prediction regarding the investigated gum-like clusters. The detection of a *gumK* gene in both *K. baliensis* clusters for example suggests the existence of an additional glucuronic acid (GlcA) residue, which could not be detected via HPLC sugar monomer analysis, while preliminary NMR analyses confirmed this prediction (data in progress). This shows that genomic analysis, as an interposed first step, can successfully support the prediction of complex HePS structures.

The active expression of the HePS clusters from *K. baliensis* was most likely proven via a mutant (R-strain) in an essential HePS-formation associated gene (*gumD*, Fig. 6). This mutant strain of *K. baliensis* NBRC 16680

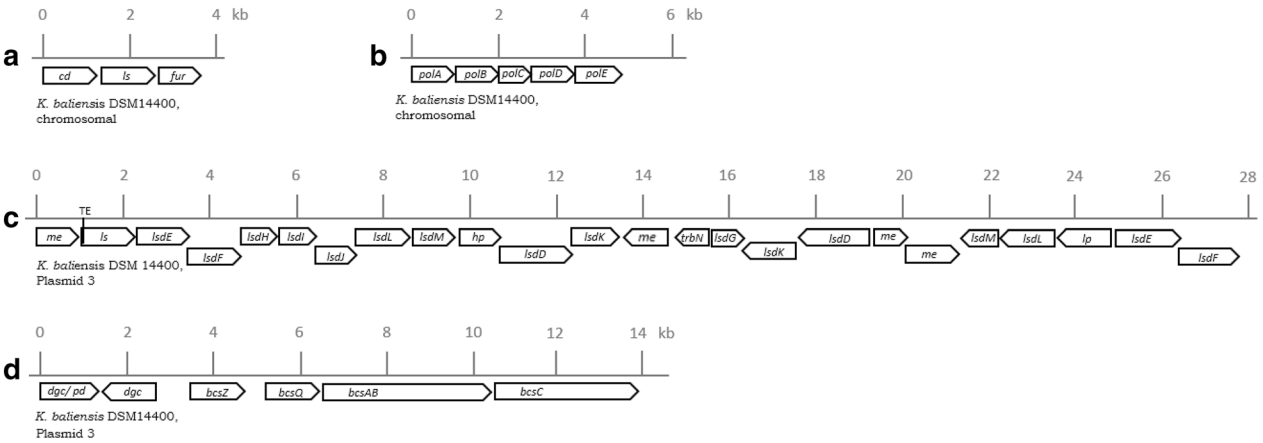


Fig. 5 Overview of additional EPS forming enzymes and clusters in *K. baliensis*. Figure (**a**) shows the *levansucrase* genes (*ls*) (EC 2.4.1.10) of *K. baliensis* DSM 14400 and NBRC 16680, flanked by a cysteine desulfurase (EC 2.8.1.7) (*cd*) and a ferric uptake regulation protein (*fur*), respectively. In (**b**) the genetic organization of the chromosomally located *pol* cluster of *K. baliensis* DSM 14400 is exemplarily depicted. The pol cluster contains 5 genes (*polABCDE)*, which are functionally designated as *polA*: dTDP-glucose 4,6-dehydratase (EC 4.2.1.46), *polB*: glucose-1-phosphate thymidylyltransferase (EC 2.7.7.24), *polC*: dTDP-4-dehydrorhamnose 3,5-epimerase (EC 5.1.3.13), *polD*: dTDP-4-dehydrorhamnose reductase (EC 1.1.1.133), *polE*: alpha-L-Rha alpha-1,3-L-rhamnosyltransferase (EC 2.4.1.-). **c** Shows an additional plasmid located levansucrase gene (EC 2.4.1.10) of *K. baliensis* DSM 14400. It shares highest similarity to *Ga. diazotrophicus* levansucrase *lsdA* and is flanked by a (partial) type II dependent secretion operon [45]). In (**d**) is the genetic organization of the *K. baliensis* DSM 14400 cellulose synthase operon illustrated. The operon consists of six genes: *dgc/pd* diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s), *dgc* diguanylate cyclase, *bcsZ* endoglucanase precursor (EC 3.2.1.4), *bcsQ* NTPase, *bcsAB* cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12), *bcsC* cellulose synthase operon protein C

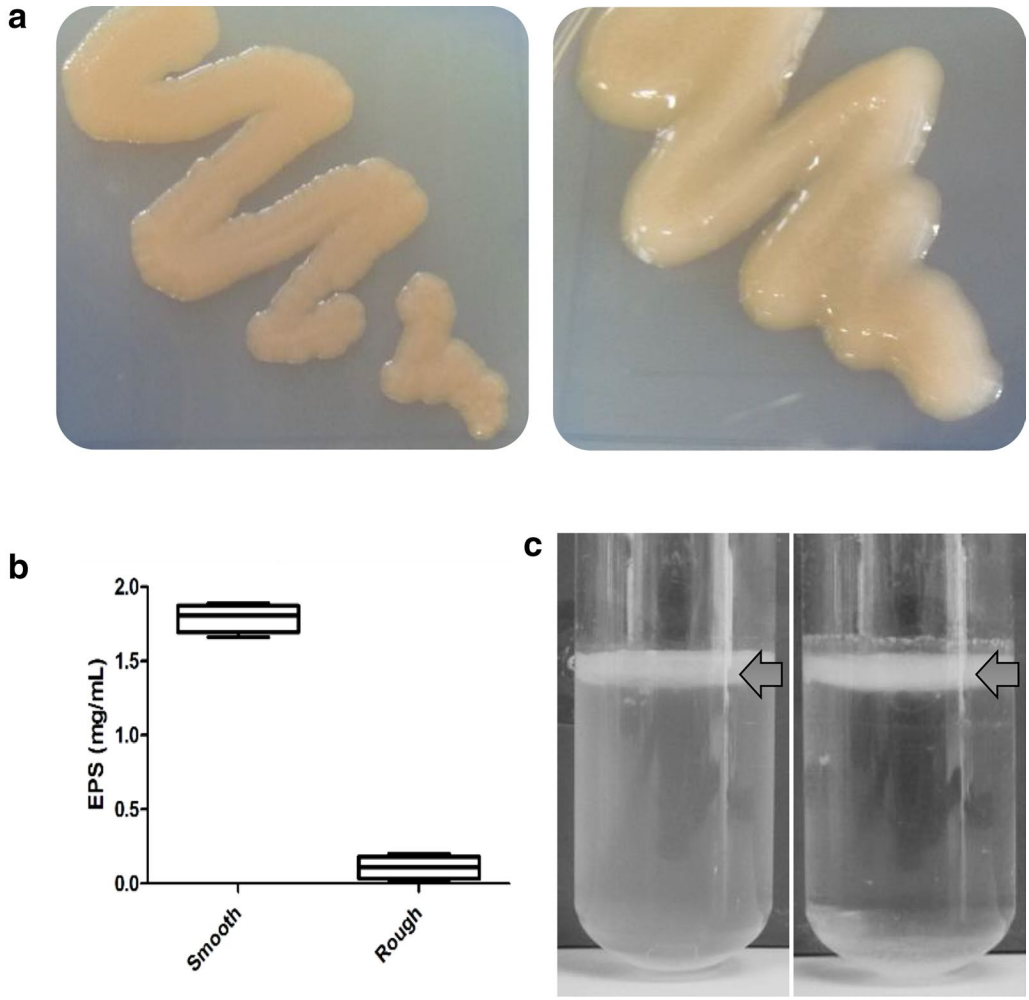


Fig. 6 Comparison of *K. baliensis* NBRC 16680 wildtype and a HePS deficient mutant strain. **a** Phenotyps of *K. baliensis* NBRC 16680 rough (*left*, mutant) and smooth strain (*right*, wildtype); **b** isolatable polymeric substances from supernatants of *K. baliensis* NBRC 16680 wildtype strain (*left*) and *K. baliensis* NBRC 16680 mutant strain (*right*); **c** growth behavior of *K. baliensis* NBRC 16680 wildtype (*left*) and mutant strain (*right*) in static culture. The mutant strain exhibits a transposon insertion in its *gum*-like cluster (Fig. 7)

is not the result of an aberration in *polABCDE* cluster (Fig. 5b) as demonstrated for *A. tropicalis* [24]. Amplification and sequencing of the *pol* clusters of the R strain showed no aberration in comparison to the S (wildtype) strain. Since both the R- and the S-strain are still able to form a surface pellicle, the synthesis of secreted HePS by *K. baliensis* is unlinked with pellicle formation [24]. Via PCR screening of the gum-like cluster of *K. baliensis* a transposon insertion in front of the *gumD* gene was found (Fig. 7). This gene catalyzes the first step of HePS synthesis in *X. campestris*, by transferring the a glucosyl-1-phosphate residue from UDP-glucose to an undecaprenyl phosphate residue. A mutation in this gene leads to a lack in xanthan synthesis by *X. campestris*, confirming the essential function of this protein during xanthan synthesis [28]. Therefore, a similar function can be assumed in the case of *K. baliensis* NBRC 16680.

Multiple EPS production by *K. baliensis* and its possible role in the environment

In addition to the *gum*-like and *pol* clusters, we were able to identify chromosomally encoded levansucrases in both investigated *Kozakia* strains. Further HoPS forming enzymes were detected on plasmid 3 of *K. baliensis* DSM 14400 [a cellulose synthase operon and a levansucrase flanked by a type II secretion operon with high similarity to a *Ga. diazotrophicus* levansucrase (GDI_RS02220)], both of which are framed by mobile elements. In contrast to the cellulose synthase operon, including all necessary elements for cellulose formation [36], the levansucrase

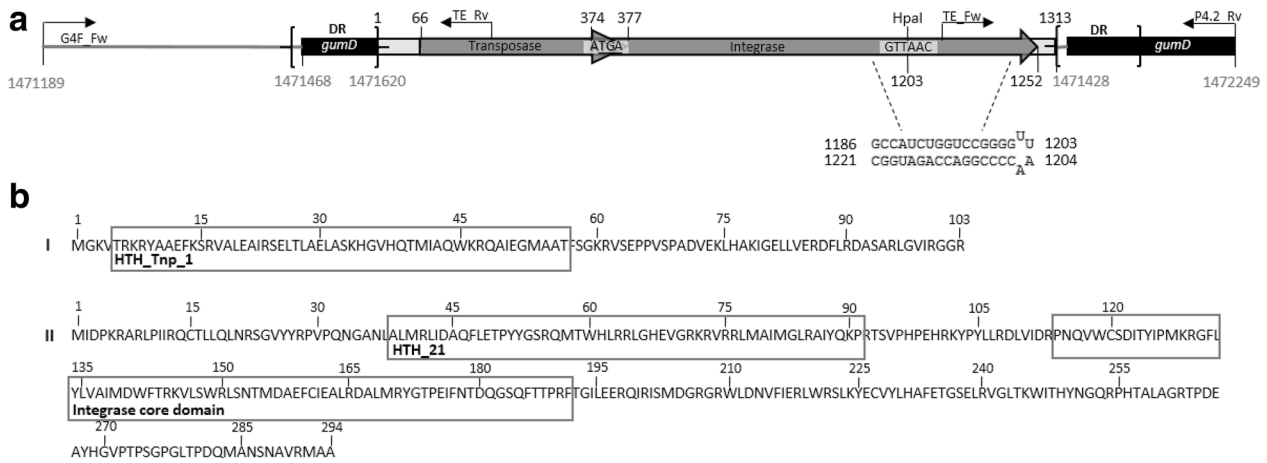


Fig. 7 Schematic representation of the transposon insertion at the *gumD* locus in the HePS biosynthesis deficient mutant strain of *K. baliensis* NBRC 16680. **a** Transposon insertion locus at the *gumD* gene. The first 279 bp represent the upstream region of the *gumD* gene (1,471,189–1,471,468 bp). The transposon insertion caused a duplication in the insertion region [direct repeat sequence (DR)] comprising 40 bp of the upstream chromosomal region of the *gumD* gene, as well as a part of the *gumD* sequence itself (1,471,428–1,471,620 bp) (the DRs are highlighted by *black brackets*, respectively). The mobile element has a total length of 1313 bp, which includes an ORF coding for a putative transposase (66–377), an ORF coding for a putative integrase (374–1252) and flanking regions (1–66; 1253–1313). The nucleotide sequence of the mobile element (1313 bp) is identically present at three further genomic regions of the *K. baliensis* NBRC 16680 chromosome (77,364–78,676 bp, 1,430,134–1,428,821 bp, 2,279,289–2,280,601 bp). At positions 1186–1221 bp of the mobile element a putative stem loop structure (dyad symmetry: 29.8 kcal/mol) could be predicted, which contains the restriction site for HpaI (5′-GTTAAC-3′; further explanation in text). The depicted genomic area was amplified with primers G4F_Fw and P4.2_Rev. **b** In I the protein sequence of a putative transposase with a predicted HTH_Tnp1 helix turn helix (bit score: 41.95, E-value: 5.57e-07) is depicted; II shows the protein sequence of a putative integrase and its predicted functional domains: HTH-like helix-turn-helix domain (bit score: 57.18, E-value: 1.44e-11) and integrase core domain (bit score: 104.65, E-value: 1.42e-28)

on plasmid 3 is interrupted by a transposon insertion and accordingly inactive. This collection of genes connected to polysaccharide formation therefore appears to provide an evolutionary and ecological advantage and possibly leads to a high physiological adaption of this strain to its environment. The additional production of other/further EPS like cellulose may lead to further advantages in changing environments. The production of cellulose is mostly known from *Komagataeibacter* and *Acetobacte*r, which prefer alcohol-enriched environments. These bacteria grow in cellulose surface pellicles that-apart from floating on surfaces for sufficient oxygen supply due to their strictly aerobic metabolism—most likely function as a barrier and protect them against osmotic stress caused by alcohol [37]. It seems that by randomly up taking EPS sequences from related AAB, *K. baliensis* collects EPS forming genes that are not necessarily active, like the levansucrase on plasmid 3 of the DSM 14400 strain, but may be activated in times of environmental changes. Related bacteria (especially regarding EPS synthesis as demonstrated in this study) such as *Ga. diazotrophicus* use these EPS for the protection against abiotic or biotic factors like desiccation and osmotic stress [38]. *Ga. diazotrophicus* is a nitrogen-fixing, endophytic bacterium, known to symbiotically colonize plants like sugar-cane [39], rice [40], as well pineapple [41] via production of levan from sucrose [42]. Because of the fact, that both *K. baliensis* genomes do not contain any nitrogen fixing genes in contrast to *Ga. diazotrophicus*, it is possible that instead of playing a role as an essential nitrogen fixing endosymbiont, *K. baliensis* uses EPS such as HePS or levan to colonize plants without positive influence on the respective plants. The association of *K. baliensis* with plant material is furthermore indicated by endoglucanases in the gum-like clusters (Fig. 3a), which could play a role during plant cell wall degradation [43]. Some other AAB strains are well known to cause fruit infections like the so-called pink disease in pineapples that results in pinkish discolorations [4]. Moreover, *X. campestris* uses xanthan as a virulence factor during infection of diverse plants. Chou et al. demonstrated that disruption of *gumD* leads to a reduced virulence of *X. campestris* in case of black rot in *Brassica oleracea*. This implicates the importance of specific EPS during microbe-plant interactions, while any involvement of EPS from *K. baliensis* in plant pathogenicity/colonization still remains to be demonstrated.

Conclusions

Via a comparative genomic approach we could get new valuable insights into the biosynthesis of novel HePS produced by *K. baliensis*, which is related to the biosynthesis of the biotechnologically widely used HePS

xanthan. Though, the properties of these novel HePS might remarkably differ from xanthan due to their different sugar/acid compositions pointing out the uniqueness of these novel HePS. The genomic approach applied in this work is enormously time saving and efficiently supports future chemical analyses such as NMR for final elucidation of these complex HePS structures. The obtained data can be used for the knowledge-based optimization and engineering of HePS production by *K. baliensis* via specific characterization of enzymes involved (identification of specific enzyme functions via activity assays and generation of deletion mutants) and transcriptional/proteomic studies.

Additional files

Additional file 1: Table S1. Primers used for transposon identification in *K. baliensis* NBRC16680.

Additional file 2: Table S2. Enzymes related to the EPS nucleotide sugar biosynthesis in *K. baliensis* (DSM 14400, NBRC 16680) and their genetic location.

Abbreviations

AAB: acetic acid bacteria; BLASTP: basic local alignment search tool; DGC: diguanylate cyclase; DSM: Deutsche Sammlung von Mikroorganismen; EPS: exopolysaccharide; Gal: galactose; GH 68: glycoside hydrolase 68 family; Glc: glucose; GT: glycosyltransferases; gum: protein involved in HePS biosynthesis (derived from *X. campestris*); HePS: heteropolysaccharide; HGAP: hierarchical genome-assembly process; HoPS: homopolysaccharide; HPLC: high-performance liquid chromatography; K.: *Kozakia*; Ko.: *Komagataeibacter*; Man: mannose; MWCO: molecular weight cut-off; NaG: modified sodium-gluconate medium; NBRC: Biological Research Centre NITE, Japan; ORF: open reading frame; Pol: protein involved in pellicle formation; RI: refractive index; SMRT: single-molecule real-time sequencing; X.: *Xanthomonas*.

Authors' contributions

JUB, FJ and RFV were involved in planning the experimental setup and writing the manuscript. JUB performed the experimental work, evaluated data and wrote the main text of the manuscript. JUB, AJG and JB were involved in genome annotations. All authors read and approved the final manuscript.

Acknowledgements

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

Genome sequencing data were deposited at the National Center for Biotechnology Information (NCBI). The corresponding accession numbers are depicted in Fig. 2.

Funding

Part of this work was funded by the German Federal Ministry for Economic Affairs and Energy via the German Federation of Industrial Research Associations (AiF) (FEI); project number AiF 18071 N. Open access publishing was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program.

Received: 9 June 2016 Accepted: 28 September 2016
Published online: 04 October 2016

References

1. Baird JK, Sandford PA, Cottrell IW. Industrial applications of some new microbial polysaccharides. Nat Biotechnol. 1983;1(9):778–83.
2. Nwodo UU, Green E, Okoh AI. Bacterial exopolysaccharides: functionality and prospects. Int J Mol Sci. 2012;13:14002–15.
3. Rosalam S, England R. Review of xanthan gum production from unmodified starches by *Xanthomonas comprestris* sp. Enzym Microbial Technol. 2006;39:197–207.
4. Raspor P, Goranovic D. Biotechnological applications of acetic acid bacteria. Crit Rev Biotechnol. 2008;28:101–24.
5. Cleenwerck I, De Vos P. Polyphasic taxonomy of acetic acid bacteria: an overview of the currently applied methodology. Int J Food Microbiol. 2008;125:2–14.
6. Deppenmeier U, Ehrenreich A. Physiology of acetic acid bacteria in light of the genome sequence of Gluconobacter oxydans. J Mol Microbiol Biotechnol. 2009;16:69–80.
7. Moonmangmee S, Toyama H, Adachi O, Theeragool G, Lotong N, Matsushita K. Purification and characterization of a novel polysaccharide involved in the pellicle produced by a thermotolerant *Acetobacter* strain. Biosci Biotechnol Biochem. 2002;66(4):777–83.
8. Moonmangmee S, Kawabata K, Tanaka S, Toyama H, Adachi O, Matsushita K. A novel polysaccharide involved in the pellicle formation of *Acetobacter* aceti. J Biosci Bioeng. 2002;2:192–200.
9. Serrato RV, Meneses CH, Vidal MS, Santana-Filho AP, Iacomini M, Sassaki GL, Baldani JI. Structural studies of an exopolysaccharide produced by Gluconacetobacter diazotrophicus Pal5. Carbohydrat Polym. 2013;98(1):1153–9.
10. Jansson PE, Lindberg J, Wimalasiri KS, Dankert MA. Structural studies of acetan, an exopolysaccharide elaborated by *Acetobacter* xylinum. Carbohydrat Res. 1993;245(2):303–10.
11. Schmid J, Sieber V. Enzymatic transformations involved in the biosynthesis of microbial exo-polysaccharides based on the assembly of repeat units. ChemBioChem. 2015;16(8):1141–7.
12. Rehm BHA. Bacterial polymers: biosynthesis, modifications and applications. Nat Rev Microbiol. 2010;8:578–92.
13. Becker A, Katzen F, Pühler A, Ielpi L. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Appl Microbiol Biotechnol. 1998;50:145–52.
14. Vorhölter FJ, Schneiker S, Goesmann A, Krause L, Bekel T, et al. The genome of *Xanthomonas campestris* pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J Biotechnol. 2008;134(1):33–45.
15. Schmid J, Koenig S, Rühmann B, Rütering M, Sieber V. Biosynthese und Genomik mikrobieller polysaccharide. Biospektrum. 2014;20(3):288–90.
16. Tao F, Wang X, Ma C, Yang C, Tang H, Gai Z, Xu P. Genome sequence of Xanthomonas campestris JX, an industrially productive strain for Xanthan gum. J Bacteriol. 2012;17:4755–6.
17. Couso RO, Ielpi L, Dankert MA. A xanthan-gum-like polysaccharide from *Acetobacter* xylinum. Microbiology. 1987;133:2123–35.
18. Kornmann H, Duboc P, Marison I, von Stockar U. Influence of nutritional factors on the nature, yield, and composition of Exopolysaccharides produced by *Gluconacetobacter xylinus* I-2281. Appl Environ Microbiol. 2003;69:6091–8.
19. Jakob F, Meißner D, Vogel RF. Comparison of novel GH 68 levansucrases of levan-overproducing Gluconobacter species. Acetic Acid Bact. 2012;1:e2.
20. Jakob F, Pfaff A, Novoa-Carballal R, Rübsam H, Becker T, Vogel RF. Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function. Carbohydr Polym. 2013;92:1234–42.
21. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Meth. 2013;10(6):563–9.
22. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res. 2014;42:206–14.

23. Kaditzky SV, Rudi F. Optimization of exopolysaccharide yields in sourdoughs fermented by *Lactobacilli*. Eur Food Res Technol. 2008;2008(228):291–9.
24. Deeraksa A, Moonmangmee S, Toyama H, Yamada M, Adachi O, Matsushita K. Characterization and spontaneous mutation of a novel gene, polE, involved in pellicle formation in *Acetobacter* tropicalis SKU1100. Microbiology. 2005;151:4111–20.
25. Chou F, Chou H, Lin Y, Yang B, Lin N, Weng S, Tseng Y. The *Xanthomonas campestris* gumD Gene Required for synthesis of xanthan gum is involved in normal pigmentation and virulence in causing black rot. Biochem Biophys Res Commun. 1997;233:265–9.
26. Pühler A, Vorhölter FJ, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Patschkowski T, Rückert C, et al. The genome of *Xanthomonas campestris* pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. J Biotechnol. 2008;134:33–45.
27. Barreras M, Abdian PL, Ielpi L. Functional characterization of GumK, a membrane-associated beta-glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis. Glycobiology. 2004;14:233–41.
28. Ielmini N, Katzen F, Ferreiro DU, Oddo CG, Vero M, Becker A, Pu A. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol. 1998;7:1607–17.
29. Vojnov AA, Zorreguieta A, Dow JM, Daniels MJ, Dankert MA. Evidence for a role for the GumB and GumC gene products in the formation of xanthan from its pentasaccharide repeating unit by *Xanthomonas campestris*. Microbiology. 1998;144(6):1487–93.
30. Dong C, Beis K, Nesper J, Brunkan-Lamontagne AL, Clarke BR, Whitfield C, Naismith JH. Wza the translocon for *E. coli* capsular polysaccharides defines a new class of membrane protein. Nature. 2006;444:226–9.
31. Serrato RV, Meneses CH, Vidal MS, Santana-Filho AP, Iacomini M, Sassaki GL, Baldani JI. Structural studies of an exopolysaccharide produced by Gluconacetobacter diazotrophicus Pal5. Carbohydr Polym. 2013;98:1153–9.
32. Castillo-Lizardo M, Henneke G, Viguera E. Replication slippage of the thermophilic DNA polymerases B and D from the Euryarchaeota Pyrococcus abyssi. Front Microbiol. 2014;5:1–10.
33. Abdian PL, Lellouch AC, Gautier C, Ielpi L, Geremia RA. Identification of essential amino acids in the bacterial alpha-mannosyltransferase AceA. J Biol Chem. 2000;275:40568–75.
34. Geremia RA, Roux M, Ferreiro DU, Dauphin-Dubois R, Lellouch AC, Ielpi L. Expression and biochemical characterisation of recombinant AceA, a bacterial alpha-mannosyltransferase. Mol Gen Genet. 1999;261:933–40.
35. Ishida T, Sugano Y, Shoda M. Novel glycosyltransferase genes involved in the acetan biosynthesis of *Acetobacter* xylinum. Biochem Biophys Res Commun. 2002;295:230–5.
36. Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, Benziman M, Gelfand DH, Meade JH. Emerick aW: genetic organization of the cellulose synthase operon in *Acetobacter* xylinum. Proc Natl Acad Sci USA. 1990;87:8130–4.
37. Perumpuli PABN, Watanabe T, Toyama H. Pellicle of thermotolerant *Acetobacter* pasteurianus strains: characterization of the polysaccharides and of the induction patterns. J Biosci Bioeng. 2014;2:134–8.
38. Velazquez-Hernandez ML, Baizabal-Aguirre VM, Cruz-Vazquez F, Trejo-Contreras MJ, Fuentes-Ramırez LE, Bravo-Patino A, Cajero-Juarez M, Chavez-Moctezuma MP, Valdez-Alarco´n JJ. Gluconacetobacter diazotrophicus levansucrase is involved in tolerance to NaCl, sucrose and desiccation, and in biofilm formation. Arch Microbiol. 2010;193:137–49.
39. Cavalcante VA, Dobereiner J. A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. Plant Soil. 1998;108:23–31.
40. Muthukumarasamy R, Kang UG, Park KD, Jeon WT, Park CY, Cho YS, Kwon SW, Song J, Roh DH, Revathi G. Enumeration, isolation and identification of diazotrophs from Korean wetland rice varieties grown with long-term application of N and compost and their short-term inoculation effect on rice plants. J Appl Microbiol. 2007;102:981–91.
41. Tapia-Hernández A, Bustillos-Cristales MR, Jiménez-Salgado T, Caballero-Mellado J, Fuentes-Ramírez LE. Natural endophytic occurrence of *Acetobacter* diazotrophicus in pineapple plants. Microbial Ecol. 2000;39:49–55.
42. Hernandez L, Arrieta J, Menendez C, Vazquez R, Coego A, Suarez V, Selman G, Petit-Glatron MF, Chambert R. Isolation and enzymic properties of levansucrase secreted by *Acetobacter* diazotrophicus SRT4, a bacterium associated with sugar cane. Biochem J. 1995;309:113–8.
43. Adriano-Anayal M, Salvador-Figueroa M, Ocampo JA, García-Romera I. Plant cell-wall degrading hydrolytic enzymes of Gluconacetobacter diazotrophicus. Symbiosis. 2005;40:151–6.
44. Griffin AM, Morris VJ, Gasson MJ. Genetic analysis of the acetan biosynthetic pathway in *Acetobacter xylinum*. Int J Biol Macromol. 1994;16:287–9.
45. Arrieta JG, Sotolongo M, Mene C, Alfonso D, Trujillo LE, Soto M, Ram R. A type II protein secretory pathway required for levansucrase secretion by gluconacetobacter diazotrophicus. J Bacteriol. 2004;186:5031–9.

PART II PAGES 53 - 62	Characterization of an acetan-like heteropolysaccharide produced by *Kozakia baliensis* NBRC 16680.

Previous findings showed that *K*. *baliensis* DSM 14400 and NBRC 16680 produce a HePS composed of glucose, galactose and mannose, respectively. In order to elucidate the exact structural components and macromolecular properties of this HePS, we carried out methylation analysis, NMR spectroscopy, rheological measurements and size determinations. Since *K*. *baliensis* DSM 14400 harbors an additional cellulose forming cluster, potentially forming cellulose, which is difficult to separate from the remaining HePS, the exact structure elucidation was carried out solely with *K*. *baliensis* NBRC 16680. Based on the in part I presented genomic date of *K*. *baliensis* NBRC 16680, showing that the biosynthesis of the *K*. *baliensis* HePS is similar to the one from acetan and xanthan, the obtained data were used as the basis for the structure determination. The HePS cluster of *K*. *baliensis* NBRC 16680 was compared with the *gum*-cluster of *X*. *campestris* and the acetan cluster of *Komagataeibacter (K.) xylinus* E25, resulting in a predicted arrangement of the individual sugar monomers, which together with the NMR data finally lead to a verisimilar HePS structure. In accordance to this, the *K*. *baliensis* NBRC 16680 HePS is composed of 1,4- and 1,3,4-substituted glucose units, suggesting a (1→4)-linked acetan-like backbone with ramifications at position *O*3. Further, we identified 1,2-substituted mannose, 1,4-substituted glucuronic acid and 1,6-substituted glucose units as possible side chain components, which would be in concordance with the acetan structure. In addition, deviating from the acetan structure, 1,6-substituted galactose units as well as a putative uronic acid were detected, which are as well assume to be part of the side chains, marking this HePS as a unique EPS that differs from known AAB EPS.

Since the focus of this work lies in the identification of a new food applicable HePS, it is important to desire constant properties of the *K*. *baliensis* HePS, to obtain reproducibility during the food production processes. To address this, the chemical composition of the HePS produced by *K*. *baliensis* NBRC 16680 in media with different initial carbon sources or supplemented magnesium was investigated, resulting in different HePS amounts, respectively possessing different rheological properties and elution profiles during AF4 and HPSEC separations. However, the *K*. *baliensis* NBRC 16680 HePS derived from the different media showed an analogue monomer composition, with similar amounts of glucose, galactose, and mannose. We presume that this is the result of a differentiated regulation process of biosynthetic enzymes involved in HePS polymerization. Especially magnesium in the medium may result in a longer

main chain of the *K. baliensis* NBRC 16680 HePS, resulting in larger molecules with a constant monomer composition.

Authors contributions: Julia U. Brandt was responsible for the independent planning and conducting the experiments. Julia U. Brandt and Daniel Wefers performed the main experimental work. Julia U. Brandt evaluated the data and wrote the main text of the manuscript. Frank Jakob was the project leader responsible for supervision. Rudi Vogel and Mirko Bunzel conducted the conceptual setup and supervision of the study including infrastructural demands and final editing of the manuscript.

International Journal of Biological Macromolecules 106 (2018) 248–257

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: www.elsevier.com/locate/ijbiomac

Characterization of an acetan-like heteropolysaccharide produced by *Kozakia baliensis* NBRC 16680

Julia U. Brandt[a], Frank Jakob[a,*], Daniel Wefers[b], Mirko Bunzel[b], Rudi F. Vogel[a]

[a] Technische Universität München, Lehrstuhl für Technische Mikrobiologie, 85354 Freising, Germany
[b] Karlsruhe Institute of Technology (KIT), Institute of Applied Biosciences, Department of Food Chemistry and Phytochemistry, 76131 Karlsruhe, Germany

ARTICLE INFO

Article history:
Received 23 May 2017
Received in revised form 20 July 2017
Accepted 2 August 2017
Available online 5 August 2017

Keywords:
Kozakia baliensis
Heteropolysaccharide (HePS)
Acetan-like

ABSTRACT

Kozakia baliensis NBRC 16680 produces a heteropolysaccharide (HePS), whose biosynthesis is similar to the biosynthesis of the exopolysaccharide acetan. To elucidate structural components and macromolecular properties of this HePS, we carried out methylation analysis, NMR spectroscopy, rheological measurements and size determinations. In accordance with acetan, the HePS are composed of 1,4-substituted Glc*p*, 1,3,4-substituted Glc*p*, 1,2-substituted Man*p*, 1,4-substituted Glc*p*A, and 1,6-substituted Glc*p* units. In contrast to acetan, rhamnose and acetylation of side chains were not detected. Furthermore, a putative, unidentified uronic acid and 1,6-substituted Gal*p* units were found to be HePS constituents, both of which could not be correlated with the responsible HePS biosynthesis in contrast to the other present structural elements. Depending on the initial carbon source, *K. baliensis* HePS were produced in different amounts and exhibited different rheological properties and elution profiles during AF4 and HPSEC separations. In conclusion, we propose an acetan-like HePS with slight molecular weight variations depending on the production conditions.

1. Introduction

The production of exopolysaccharides (EPS) is a common attribute of bacteria. With respect to their cellular attachment, EPS are divided into two groups: cell surface attached capsular polysaccharides (CPS) and secreted exopolysaccharides (EPS). Both kinds of EPS are of increasing interest from a biological point of view, as well as for industrial applications. EPS play an important role in bacteria–host interaction, biofilm formation [1], and protection against desiccation or resistance to antimicrobial agents [2,3]. The industry is taking advantage of the unique rheological properties of diverse bacterial EPS, which can be usually recovered in higher purity than plant polysaccharides. The addition of EPS strongly affects the rheology and texture of aqueous solutions, resulting in drastic viscosity increases, which makes them a useful ingredient for e.g. the cosmetic or food industry.

Acetic acid bacteria (AAB) are Gram-negative obligate aerobes belonging to the α-proteobacteria subdivision and are well known to produce large amounts of EPS, either homopolysaccharides (HoPS), such as dextrans, levans [4–6], and cellulose, or different kinds of heteropolysaccharides (HePS), such as acetan [7] and gluconacetan [8]. The biosynthetic pathways and chemical structures of AAB derived HePS are related to the biotechnologically widely used HePS xanthan, which is formed by *Xanthomonas* (*X.*) *campestris* [7,9]. Acetan, gluconacetan, and xanthan are complex HePS, which are composed of a β-(1 →4)-linked glucopyranosyl backbone. This linear backbone is substituted with a tri- or pentasaccharide side chain at position *O*3 of every second glucose unit. The side chain structure is dependent on the actual HePS. The side chains of both xanthan and acetan are composed of a mannose unit which is substituted with a 1,4-substituted glucuronic acid unit at position *O*2. In case of xanthan, one terminal mannose unit is attached to the glucuronic acid unit, whereas the side chain of acetan contains two further 1,6-substituted glucose units and one terminal rhamnose unit [10].

Recently, we have shown that two *Kozakia (K.) baliensis* strains (DSM 14400, NBRC 16680) not only produce high molecular weight levan from sucrose, but also form large amounts of a soluble unique HePS composed of glucose, galactose, and mannose [9]. The underlying HePS forming biosynthesis cluster is related to both the *gum*-cluster of *X. campestris* [11], and the acetan cluster of *Komagataeibacter (K.) xylinus* E25 [12]. Since *K. baliensis* is a non-

* Corresponding author at: Technische Universität München, Lehrstuhl für Technische Mikrobiologie, Gregor-Mendel-Straße 4, 85354, Freising, Germany.
E-mail address: frank.jakob@wzw.tum.de (F. Jakob).

http://dx.doi.org/10.1016/j.ijbiomac.2017.08.022
0141-8130/

pathogenic, naturally and traditionally food-associated bacterium, these new HePS are potentially promising hydrocolloids for novel food formulations.

For the applicability of EPS in foods, either directly produced during natural/traditional food fermentations or added as isolated substances, constant properties of the particular polysaccharides are desired to obtain reproducible food production processes. Moreover, the EPS yields are usually optimized by varying production media, mainly the carbon sources. In general, inconsistent results have been published concerning the influence of carbon sources on the EPS yields and compositions. Kornmann et al. [13] reported that nutritional factors have a strong influence on the yield of gluconacetan produced by *Komagataeibacter* (*Ko.*) *xylinus*, whereas the HePS compositions were constant independent of the production conditions. Contrary, a report on EPS produced by *Lactobacillus delbrueckii* subsp. *bulgaricus* suggests that the initial carbon sources affect the final EPS composition [14]. However, hardly any information is available on the influence of the carbon source on the monosaccharide composition of AAB derived HePS.

Furthermore, knowledge about the structures of HePS formed by AAB is scarce. Hypothetical main structures of HePS were only proposed for a few AAB strains, such as *Komagataeibacter (Ko.) xylinus* [7] and *Gluconacetobacter (Ga.) diazotrophicus* [15]. In addition, putative pellicle HePS structures were described for a small number of *Acetobacter* strains, such as *Acetobacter (A.) tropicalis* [16] and *A. aceti* [17].

In the present study, we propose a hypothetical structure of the *K. baliensis* HePS based on the determination of the glycosidic linkages by methylation analysis and NMR spectroscopy, the annotation of gene/protein functions in the respective HePS biosynthesis cluster, and the comparison to known literature about HePS from AAB. In addition, the chemical composition of the HePS produced by *K. baliensis* NBRC 16680 in media with different initial carbon sources was investigated. In order to get further insights into the macromolecular properties of HePS recovered from *K. baliensis* NBRC 16680, steady shear rheological measurements as well as molecular size determinations were carried out.

2. Material and methods

2.1. Strains, media and growth conditions

Kozakia baliensis NBRC 16680 (NBRC; National Institute of Technology and Evaluation (NITE) Biological Resource Center, Japan) was generally pre-incubated aerobically at 30 °C in liquid unmodified NaG media (20 g/L sodium gluconate, 3 g/L yeast extract, 2 g/L peptone, 3 g/L glycerol, 10 g/L mannitol, pH adjusted to 6.0). For subsequent EPS production experiments, the NaG media were modified with different carbon sources. Mannitol was replaced by either 10 g/L glucose (NaG-g), 10 g/L fructose and 10 g/L glucose (NaG-fg), or 10 g/L fructose, 10 g/L glucose and 200 mg/L magnesium (NaG-fgm), respectively. For preparation of the EPS production media, cells from pre-cultures were centrifuged, washed twice with peptone water (3.6 g/L KH_2PO_4, 7.2 g/L $NaH_2PO_4 \times 2H_2O$, 4.3 g/L NaCl, 1 g/L pepton) and resuspended in 1 mL of either modified or non-modified liquid NaG media. Main cultures were performed in 500 mL Erlenmeyer flasks with 50 mL of modified NaG media, inoculated with 500 μL of the pre-cultures and kept at 30 °C in a rotary shaker (200 rpm) for 32 h. Subsequently, cells were removed and the EPS containing supernatants were precipitated with cold ethanol (2:1, v/v) and kept overnight at 4 °C. This step was repeated three times, the HePS were recovered by centrifugation, re-dissolved in ddH_2O, and dialyzed (MWCO 14 kDa). Finally, the purified HePS were lyophilized and quantified by weighing.

2.2. Neutral monosaccharide analyses

To analyze the neutral sugar composition of the isolated HePS, high performance liquid chromatography (HPLC) was used. The purified polysaccharide samples were hydrolyzed with 10.5% (v/v) of perchloric acid over 7 h at 100 °C, followed by a centrifugation step (4 °C, 10 min, 13,000*g*) to remove possible impurities, such as proteins. The supernatant was analyzed using a Rezex RPM column (Phenomenex, Germany) coupled to a refractive index (RI) detector (Gynkotek, Germany) corresponding to the method of Kaditzky [18]. Sugar monomers were identified according to their retention times using suitable monosaccharide standards (glucose (Glc), galactose (Gal), mannose (Man), rhamnose (Rha)). The mobile phase was water, the flow rate was 0.6 mL/min.

The absolute configuration of the neutral sugars was determined after hydrolysis of the polysaccharides with 2 M trifluoroacetic acid (TFA) for 30 min at 121 °C. Subsequently, the acid was evaporated, and the sugars were heated overnight with 150 μL of (*R*)-2-octanol and 5 μL of TFA [19]. The solvent was removed and the sample was silylated by using 80 μL of *N*,*O*-bis(trimethylsilyl)trifluoroacetamide and 20 μL of pyridine. The silylated sugar derivatives were analyzed by GC–MS (GC-2010 Plus and GC–MS-QP2010 Ultra, Shimadzu, Japan) equipped with an Rxi-5Sil MS column (30 m x 0.25 mm i.d., 0.25 μm film thickness, Restek, Germany) using the following conditions: initial column temperature, 150 °C; ramped at 1 °C/min to 200 °C; ramped at 15 °C/min to 300 °C. Injection temperature was 275 °C, and split injection was used at a split ratio of 10:1. Helium was used as carrier gas at 40 cm/sec, and the transfer line was held at 275 °C.

2.3. Structural analysis

To obtain further information about other HePS sugar components such as uronic acids, the polysaccharides were hydrolyzed by a combination of methanolysis and TFA hydrolysis and analyzed by high performance anion exchange chromatography with pulsed amperometric detection (HPAEC-PAD) [20]. Briefly, hydrogen chloride in methanol (1.25 M, 500 μL) was added to 200 μg sample, and methanolysis was performed for 16 h at 80 °C. The acid was evaporated, and the residue was hydrolyzed with 500 μL of 2 M TFA for 1 h at 121 °C. After evaporation, the samples were redissolved in water and analyzed for the released monosaccharides by HPAEC-PAD on an ICS-5000 system (Thermo Scientific Dionex, CA, USA) using a CarboPac PA-20 column at 25 °C. A flow rate of 0.4 mL/min and a gradient composed of (A) bidistilled water, (B) 0.1 M sodium hydroxide, and (C) 0.1 M sodium hydroxide +0.2 M sodium acetate were used: Before every run, the column was rinsed with 100% B for 10 min and equilibrated for 20 min with 90% A and 10% B. After injection, the following gradient was applied: 0–1.5 min, from 90% A and 10% B to 96% A and 4% B; 1.5–22 min, isocratic, 96% A and 4% B; 22–32 min from 96% A and 4% B to 100% B; 32–42 min, isocratic, 100% C.

The glycosidic linkages of NaG and NaG-fgm were analyzed by methylation analysis as described previously [21,22]. Briefly, the polysaccharides were dissolved in 2 mL of dimethyl sulfoxide and freshly ground sodium hydroxide (100 mg) was added. The mixture was incubated for 90 min in an ultrasonic bath and 90 min at room temperature. Methyl iodide (1 mL) was added, and the mixture was sonicated for 30 min and incubated for 30 min at room temperature. The methylated polysaccharides were extracted into dichloromethane (3 mL), and the organic layer was washed with 0.1 M sodium thiosulfate (5 mL) and twice with water. The solvent was evaporated and the residue was dried overnight in a vacuum oven at 40 °C. After the addition of 2 M TFA (2 mL), the residue was incubated at 121 °C for 90 min. TFA was removed by evaporation, and $NaBD_4$ (20 mg) in aqueous NH_3 solution (2 M) was

added. The partially methylated monosaccharides were reduced for 1 h, and reduction was terminated by adding glacial acetic acid (100 μL). While cooling on ice, 450 μL of 1-methylimidazole and 3 mL of acetic anhydride were added, and the sample was incubated for 30 min at room temperature. Water (3 mL) was added, and the partially methylated alditol acetates (PMAA) were extracted with 5 mL of dichloromethane. The organic phase was washed with water (5 mL, three times) and the residual water was removed by freezing overnight at −18 °C. The PMAAs were analyzed on a GC–MS system (GC-2010 Plus and GCMS-QP2010 SE instruments, Shimadzu, Japan) equipped with an Rxi-5Sil MS column (30 m x 0.25 mm i.d., 0.25 μm film thickness, Restek). The following conditions were used: Initial column temperature, 140 °C, held for 2 min; ramped at 1 °C/min to 180 °C, held for 5 min; ramped at 10 °C/min to 300 °C, held for 5 min. Helium was used as carrier gas at 40 cm/sec. Split injection with a split ratio of 30:1 was used, and the injection temperature was 250 °C. The transfer line was held at 275 °C, and electron impact spectra were recorded at 70 eV. Semiquantitative analysis of the PMAAs was performed by using a GC-FID system (GC-2010 Plus, Shimadzu) equipped with a DB-5MS column (30 m × 0.25 mm i.d., 0.25 μm film thickness, Agilent Technologies, CA). The same conditions as described for the GC–MS analyses were used, with the exception of a reduced split ratio of 10:1. The FID temperature was 240 °C and nitrogen was used as makeup gas. Molar response factors were used to calculate the portions of the PMAAs [23].

Carboxyl reduction of the uronic acids was performed as described previously [24]. The samples were dissolved in 1 mL of bidistilled water, followed by the addition of MES (200 μL, 0.2 M, pH 4.75) and 1-cyclohexyl-3-(2-morpholinoethyl) carbodiimide-metho-*p*-toluene sulfonate (400 μL, 500 mg/mL). The solution was mixed and incubated for 3 h at room temperature. While cooling on ice, 1 mL of 4 M imidazole (pH 7.0) was added. Subsequently, 1 mL of aqueous $NaBD_4$ solution (70 mg/mL) was added and the sample was incubated for 3 h at room temperature. The reaction was terminated by the addition of glacial acetic acid (500 μL), and the sample was dialyzed for 24 h against water (Molecular Weight Cut Off: 3500 Da). After freeze drying, the carboxyl-reduced samples were methylated, hydrolyzed, reduced, and acetylated as described above.

NMR spectroscopic analysis was carried out on a Bruker DRX-500 spectrometer equipped with a 5 mm probe (Bruker, Germany). The HePS derived from NaG-fgm were dissolved in D_2O (0.8% w/v), and spectra were acquired at 350 K. The HDO signal was suppressed by using the zgpr pulse sequence. Dimethyl sulfoxide was used for spectrum calibration (2.71 ppm [25]).

2.4. Rheological measurements

For rheological measurements of the isolated HePS, serial dilutions were prepared in ddH_2O (1.5%, 1%, 0.75%, 0.5%, 0.25%, w/v). The HePS dilutions were stirred for 30 min at room temperature and heated twice for 20 min at 80 °C. Steady shear rheological data were obtained with a Rheometer (Physica MCR 501 Rheometer (Anton Paar, Austria)) using a double gap geometry (DG 26.7-SS, Anton Paar, Austria) at a constant temperature of 20 °C. The shear rate profile moved along a logarithmic scale beginning with a starting sequence including a shear rate from 0.1 to 1000 s^{-1} followed by a reverse sequence from 1000 to 0.1 s^{-1}. Per decade, five values were recorded over a period of 10 s. The actual measurement was preceded by an adaptation phase of 30 s.

To describe the steady shear rheological properties of the samples, the data were fitted to the well-known power law model (Eq. (1)), which is extensively used to describe the flow properties of non-newtonian liquids [26].

$$\sigma = K\gamma^n \quad (1)$$

σ refers to the shear stress (Pa), γ is the shear rate (s^{-1}), K is consistency index (Pa sn), and n is the flow behavior index (dimensionless). The steady shear rheological measurements were conducted in biological triplicates. The reported results were expressed as an average of the three measurements.

2.5. Separation of isolated HePS and size determinations

Asymmetric flow field flow fractionation coupled to multi-angle light scattering (AF4-MALS; Wyatt Technology) was used to separate HePS samples and to subsequently analyze their size distributions. Isolated HePS samples were dissolved in ddH_2O by vigorous vortexing (0.1 mg/mL) and centrifuged at 14,000*g* for 15 min to remove undissolved particles, which hindered AF4 separations in previous experiments (data not shown). For sample separations, 100 μL of the obtained supernatants were injected into the AF4-channel, respectively. The flow conditions and materials/equipment for separation of HePS samples were identical to those previously described [27]. HePS separations were analyzed using ASTRA 6.1 Software (Wyatt Technology), whereas the integrated "particle mode" and the model "random-coil" allowed evaluation of separated fractions regarding their root-mean square (rms) radii size distributions. High performance size exclusion chromatography (HPSEC) was performed on an HPLC system (L-7100 pump, L-7490 RI detector, Merck/Hitachi, Germany) equipped with a TSKgel G5000PWXL column (300 × 7.8 mm, 10 μm particle size) and a TSKgel G3000PWXL column (300 × 7.8 mm, 7 μm particle size) in series. Elution was carried out with bidistilled water at a flow rate of 0.6 mL/min and 60 °C.

3. Results

3.1. Influence of the carbon source on the HePS yield and sugar monomer composition

Four fermentations with *K. baliensis* NBRC 16680 were performed in basic NaG media with different carbon sources: 10 g/L mannitol (NaG), 10 g/L glucose (NaG-g), 10 g/L fructose and 10 g/L glucose (NaG-fg), 10 g/L glucose, 10 g/L fructose, and 200 mg/L magnesium (NaG-fgm). During cultivation of *K. baliensis* NBRC 16680 in these media, different growth patterns were observed (Fig. 1A). In addition, varying pH-values were observed over the course of fermentation depending on the carbon source (Fig. 1B). *K. baliensis* NBRC 16680 showed the fastest growth in NaG-fgm media with about 4×10^9 CFU/mL after 8 h, whereas the cell number in NaG media at this time point was 6×10^8 CFU/mL. The cultivation in NaG-g and NaG-fg media showed intermediary growth curves, with about 1×10^9 CFU/mL at 8 h. However, after 32 h of incubation, *K. baliensis* NBRC 16680 reached comparable cell numbers ($\sim 1 \times 10^{10}$ CFU/mL) in all four media. Concomitant HePS production was measured after 32 h via dry weight determination of the isolated HePS (Fig. 1C). In unmodified NaG media, NBRC 16680 showed the lowest HePS production, with amounts of 1.73 ± 0.03 g/L. In both glucose and fructose containing media (NaG-fg, NaG-fgm) the HePS yields doubled compared to the NaG media (3.87 ± 0.01 g/L, 3.93 ± 0.2 g/L) (Fig. 1C). Addition of glucose to the NaG medium (NaG-g) also results in an increase compared to the NaG media, with a yield of 2.64 ± 0.1 g/L HePS. However, in contrast to the d-fructose containing media, the amount of NaG-g media derived HePS was significantly lower, even though the NaG-fg and NaG-g media showed the same growth pattern. Nevertheless, it should be noted that the NaG-fg and NaG-fgm medium contain twice the

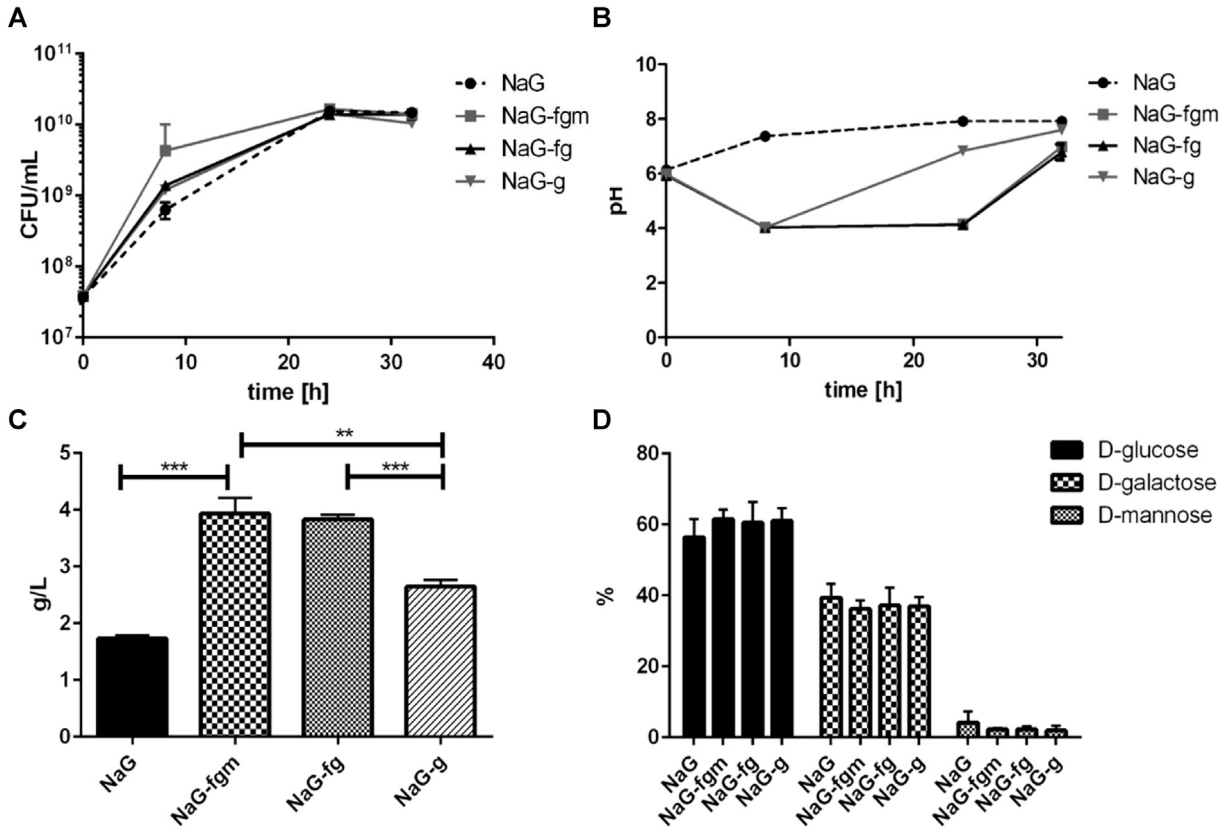


Fig. 1. Growth of *K. baliensis* NBRC 16680 in different media and influence of the provided carbon source on HePS amount and composition. (A) Growth curve and (B) pH-values during growth of *K. baliensis* NBRC 16680 in media with different carbon sources. NaG (●), NaG-fgm (■), NaG-fg (▲), NaG-g (▼). (C) Amount of precipitated HePS from *K. baliensis* NBRC 16680 during growth on different carbon sources: d-mannitol (NaG, black), d-glucose (NaG-g, striped), d-glucose & d-fructose (NaG-fg, dotted) and d-glucose, d-fructose and magnesium (NaG-fgm, chess pattern). (D) Neutral sugar composition of the isolated HePS for *K. baliensis* NBRC 16680, isolated from the media as described above. The portions of the monomer constituting the particular HePS are represented as percentage. Asterisks centered over the error bars indicate the relative level of the p-value. In general, *means $p < 0.05$, **means $p < 0.005$ and ***means $p < 0.0005$.

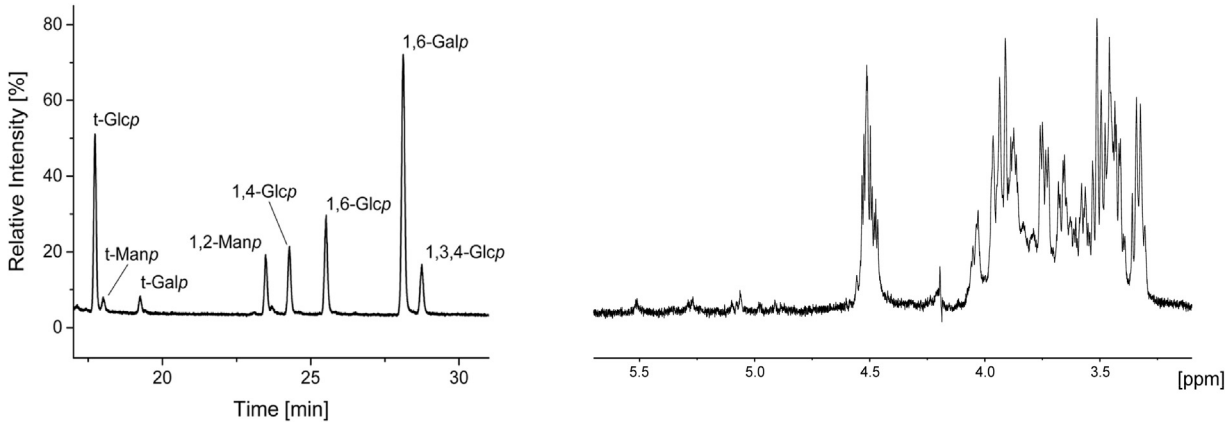


Fig. 2. (A) Total ion current chromatogram of the PMAAs resulting from methylation analysis of *K. baliensis* HePS derived from NaG-fgm medium. t = terminal, Glc = glucose, Gal = galactose, Man = mannose, *p* = pyranose. Numbers indicate the substituted positions of a sugar unit. (B) ^{1}H NMR spectrum of *K. baliensis* HePS derived from NaG-fgm medium in D_2O at 350 K. The HDO resonance at 4.2 ppm was suppressed by presaturation.

amount of sugars, which could generally contribute to an increased HePS production.

The neutral sugar composition of the isolated HePS samples was determined via HPLC analysis after acidic hydrolysis. HPLC analysis of the perchloric acid hydrolysate of *K. baliensis* NBRC 16680 HePS revealed three peaks. The retention times of the detected monomers were consistent with the retention times of the standards glucose (Glc), galactose (Gal), and mannose (Man). GC–MS analysis of the absolute configuration after chiral derivatization confirmed the presence of d-glucose, d-galactose, and d-mannose. A comparable neutral monosaccharide distribution was detected in all three biological replicates of HePS from *K. baliensis* NBRC 16680 grown in NaG media (Fig. 1D). The other three production media (NaG-g, NaG-fg, NaG-fgm) also yielded comparable neutral sugar compositions (Fig. 1D). For further structural and rheological analyses (sections 3.2 + 3.3), NaG and NaG-fgm derived HePS were selected, as they were produced in significantly different amounts (Fig. 1C).

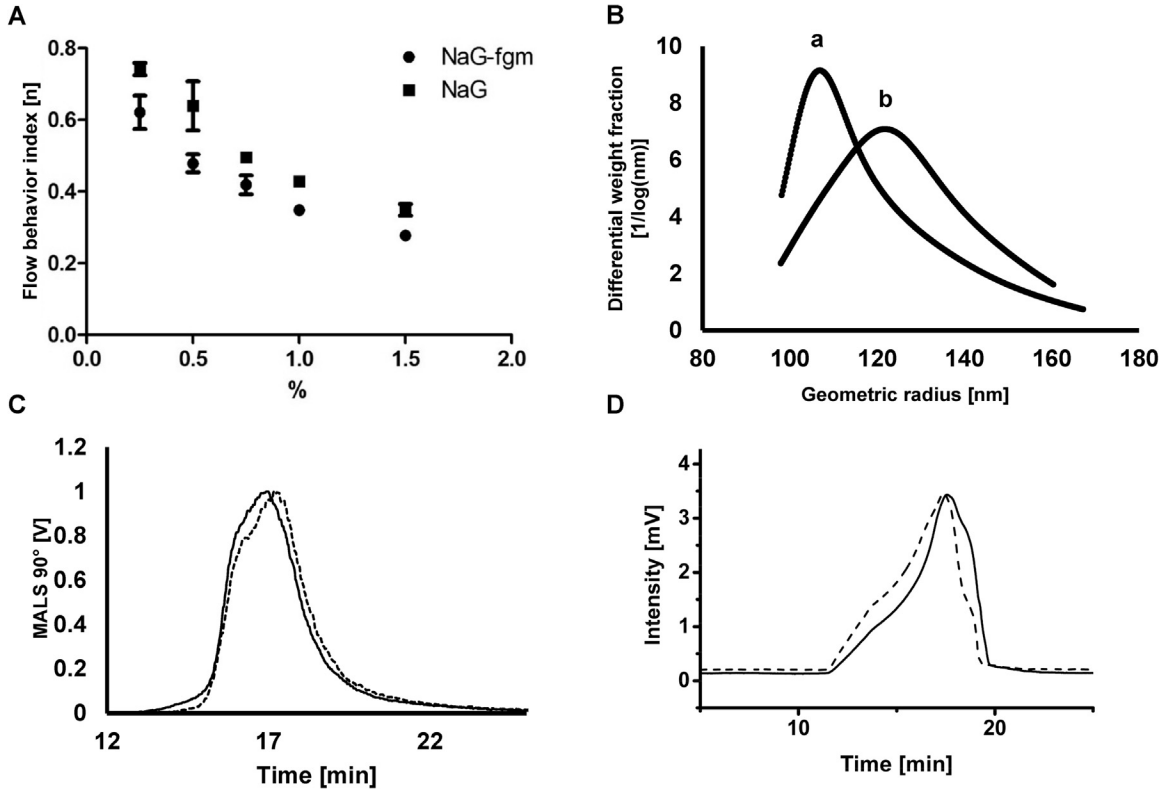


Fig. 3. (A) Flow behavior index (n) for different concentrations of *K. baliensis* NBRC 16680 HePS isolated from NaG- and NaG-fgm-media. The dissolved HePS amounts are represented as percentages (w/v). (B) rms radius distributions of *K. baliensis* NBRC 16680 formed HePS isolated from NaG medium (a) and NaG-fgm medium (b). (C) AF4 analysis comparing retention times between *K. baliens*is HePS derived from NaG medium (black solid) and NaG-fgm medium (black dashed). (D) HPSEC analysis of the *K. baliens*is HePS derived from NaG medium (black solid) and NaG-fgm medium (black dashed).

3.2. Determination of linkage types and assignment of structural components

To obtain further information about the monosaccharide composition of *K. baliensis* NBRC 16680 HePS, a combination of methanolysis and TFA hydrolysis was used. This procedure in combination with subsequent HPAEC-PAD analysis was described to be most suitable for the complete hydrolysis and analysis of uronic acid containing soluble polysaccharides [20]. The HPAEC-PAD chromatograms of both the NaG and NaG-fgm derived HePS demonstrated the presence of two peaks eluting after glucose, galactose, and mannose (Supplementary file 2). One of the peaks was identified as glucuronic acid, whereas the other could not be identified by using common uronic acid standards (galacturonic acid, mannuronic acid, and guluronic acid). However, the strong retention on the CarboPac PA-20 column suggests the presence of a second uronic acid or a uronic acid derivative. To roughly estimate the portions of the two uronic acids, an external calibration with glucuronic acid was used. By using this approach, it was possible to demonstrate that the sum of the two uronic acids is comparable to the amount of mannose.

For further information about the glycosidic linkages of the monosaccharide units, NaG-fgm HePS were analyzed by GC–MS after methylation analysis (Fig. 2A). In addition, the portions of the partially methylated alditol acetates were semiquantitatively estimated via GC-FID (Supplementary file 3). Besides small amounts of terminal units, both mannose and galactose are incorporated only via one main-type of linkage into the polysaccharide as demonstrated by the identification of the PMAAs derived from 1,2-linked mannose and 1,6-linked galactose. In contrast, higher portions of terminal units and various linkage types were observed for glucose. The detection of roughly comparable amounts of two PMAAs derived from *O*4-substituted glucopyranose units (1,4-Glc*p* and 1,3,4-Glc*p*) suggests the presence of a polysaccharide with a (1 → 4)-linked backbone with branches at position *O*3. In addition, 1,6-substituted glucose was present in a higher amount than each of the two putative backbone units. However, based on methylation analysis data only it cannot be concluded whether one or more polysaccharides are present or what kind of side chains are linked to the backbone. Methylation analysis of HePS derived from NaG yielded the same glycosidic linkages in comparable portions (Supplementary file 3). The only larger difference is a lower portion of the 1,6-substituted glucose units. However, because the portions of all other PMAAs remain relatively constant, these results indicate that the HePS from *K.baliensis* have a conserved structure.

The analysis of the PMAAs after carboxyl reduction allowed for the detection of the dideuterated PMAA derived from 1,4-Glc*p*, which demonstrates the presence of 1,4-substituted glucuronic acid. The detection of other uronic acid derived, dideuterated PMAAs was not possible. Therefore, it can be concluded that glucuronic acid is mainly involved in (1 → 4)-linkages. In addition, the absence of dideuterated PMAAs derived from the unidentified uronic acid indicates that this compound is a derivative of glucuronic acid.

To obtain information about the anomeric configuration of the monosaccharide units, NMR spectroscopy was applied. However, due to the high viscosity of the solutions and the resulting low T2 relaxation time, very low signal intensities were obtained if analyzed at room temperature. Therefore, spectra were recorded at 350 K, which allowed for the acquisition of a proton spectrum of the HePS derived from NaG-fgm (Fig. 2B). Besides the signals of the non-anomeric protons (3.30–4.10 ppm), several resonances between 4.40 ppm and 4.60 ppm were observed. These signals are most likely derived from 1,4-, 1,3,4-, and 1,6-substituted β-glucose units

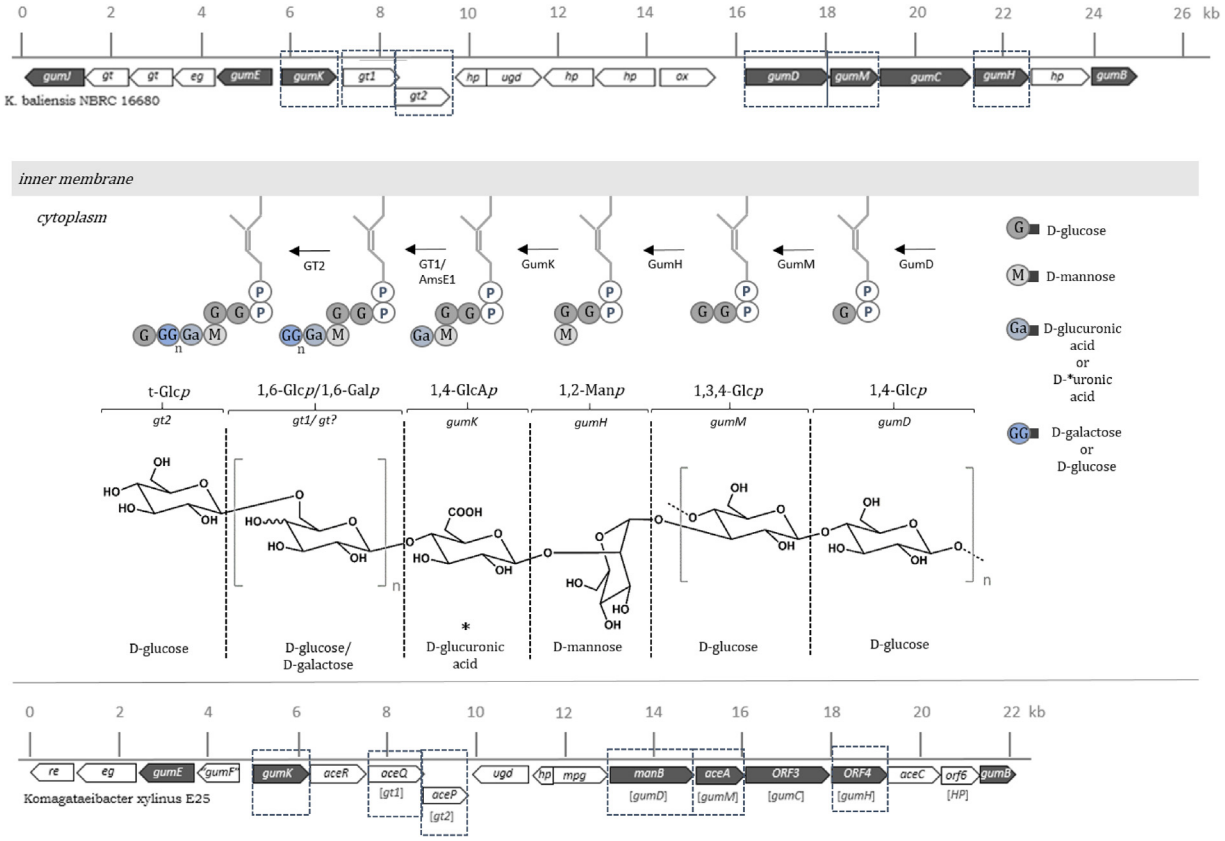


Fig. 4. Hypothetical structure of the *K. baliensis* HePS in correlation with the responsible biosynthesis pathway. (a) HePS biosynthesis encoding gene cluster of *K. baliensis* NBRC 16680. The cluster has an overall size of ~25 kb and involves 19 genes, including genes encoding for glycosyltransferases (*gt*), hypothetical proteins (*hp*), and eight *gum* like genes (*gumB, –C, –D, –E, –H, –J, –K, and –M*), which are marked in *grey*. Furthermore, the cluster contains a putative endoglucanase (*eg*), oxidoreductase (*ox*) and a UDP-glucose dehydrogenase (*ugd*) [9]. (b) outlines the proposed partial HePS biosynthesis pathway (deduced from xanthan and acetan biosynthesis; Vorhölter et al. [38], Griffin et al., [12]. (c) Scheme of the hypothetical acetan-like structure of one repeating unit of *K. baliensis* NBRC 16680 HePS, which can finally be polymerized (n) to the macromolecular, secreted HePS. The particular monosaccharides and their linkage types are depicted together with the corresponding genes, which encode the respective putative transferase reactions. The currently supposed positions of the unidentified uronic acid ("*") and of 1,6-linked glucose/galactose, both of which could not be clearly correlated with the biosynthesis cluster, are indicated. The relative high portions of 1,6-linked glucose/galactose (supplementary file 3) are assumed to result from a possibly multiple (indicated by "n") incorporation of 1,6-linked galactose instead and/or 1,6-linked glucose into the repeating unit. "~" indicates that this sugar may be a glucose or a galactose unit. Because of the high portions of β-configured glucose/galactose in the NMR spectrum, the β-configuration is depicted for the 1,6-linked glucose/galactose residues. However, considering the similarity of the *K. baliensis* HePS to acetan, it is also possible that an α-1,6-linked glucose is attached to the uronic acid residue. (D) *Gum* cluster of the acetan producer *Komagataeibacter xylinus* E25. The nomenclature for the acetan cluster is based on AM Griffin, Morris, V. J., Gasson, M. J. [12], while brackets under the particular genes mark the homologous *gum* genes. The dotted squares in the particular *gum*-clusters mark genes, which are homologous between the specific clusters. Initial ORF predictions and annotations were accomplished automatically using the program RAST, a SEED-based, prokaryotic genome annotation service [61]. Annotations were corrected via the NCBI Prokaryotic Genome Annotation Pipeline. EPS cluster annotations were performed via homology searches against GeneBank/EMBL using the function BLASTP as described previously [9]. t = terminal, Glc = glucose, Gal = galactose, Man = mannose, *p* = pyranose. Numbers indicate the substituted positions of a sugar unit.

and 1,4-substituted β-glucuronic acid units [7,28]. The ^{1}H chemical shift of the anomeric protons of 1,6-substituted β-galactose units should be roughly comparable with the chemical shift of the anomeric protons of 1,6-substituted β-glucose units. Therefore, the signals derived from this structural element might also be present. However, due to the overlapping signals, it was not possible to make unambiguous assignments. Only very small peaks were observed at the expected chemical shifts of the anomeric protons of 1,2-substituted α-mannose units (~5.30 ppm) and 1,6-substituted α-glucose units (~5.50 ppm), which might be a result of a lower abundance or a lower T2 relaxation time.

3.3. *Rheological properties and sizes of the* K. baliensis *NBRC 16680 derived HePS*

In order to characterize the flow behavior of HePS from *K. baliensis* NBRC 16680, flow curves were recorded using different HePS solutions, with a shear rate profile between 0.1 to 1,000 s^{-1}. The influence of different production media on the rheological properties of isolated HePS were studied via comparison of NaG versus NaG-fgm media derived HePS. The resulting shear rheological data were plotted as viscosity ([Pa·s]) versus shear rate (r, [1/s]) for different HePS concentrations (15 g/L, 10 g/L, 7.5 g/L, 5 g/L, 2.5 g/L), resulting in flow curve profiles (supplementary file 1). The obtained flow curve data were fitted to the power law model (Eq. (1)), enabling the calculation of flow behaviour indices (n) (Fig. 3A). The samples with the highest HePS concentration (15 g/L) exhibited flow behaviour indices (n) between 0.28 (NaG-fgm) and 0.35 (NaG) (Fig. 3A). The n values generally decreased with an increasing HePS concentration, showing the dependence of the HePS concentration on the viscosity of these shear-thinning solutions. Generally, n values were lower for NaG-fgm derived HePS if compared to unmodified NaG derived HePS at equal concentrations (Fig. 3A).

Both HePS types (NaG + NaG-fgm) were further characterized by AF4-MALS and HPSEC to get insights into the possible macromolecular differences of these differently produced HePS (Fig. 3B-D). NaG-fgm derived HePS contained a larger fraction of molecules, which exhibited higher rms radii in contrast to NaG derived HePS

(Fig. 3B). The existence of larger molecules (potentially of larger molecular weight) is also indicated by a slight shift of retention times among both samples (molecules of higher molecular weight elute later in AF4 separation, Fig. 3C). However, it was not possible to calculate molecular weights for both samples due to the relatively low amount of injected sample (10 μg), which limited detection of utilizable concentration signals (RI or UV detection). Injection of higher HePS amounts (>10 μg) resulted in complete aggregation and consequently inadequate separation of HePS molecules. The chromatograms obtained from the HPSEC analysis demonstrated that NaG-fgm derived HePS exhibit a lower retention time than the NaG derived HePS (Fig. 3D), which is most likely the result of a higher molecular weight. However, because both polysaccharides eluted at about the retention time of Dextran Blue (molecular weight: 2000 kDa), no detailed information about the molecular weight can be obtained.

4. Discussion

4.1. Influence of the provided carbon source on the yield and composition of K. baliensis produced HePS

K. baliensis NBRC16680 was cultivated in NaG media with different carbon sources (10 g/L mannitol (NaG), 10 g/L glucose (NaG-g), 10 g/L fructose and 10 g/L glucose (NaG-fg)). In addition, a combination of d-glucose (10 g/L), d-fructose (10 g/L) and magnesium (200 mg/L) was included in the experimental series, because magnesium showed stimulating effects on HePS production [29]. During growth of *K. baliensis* NBRC16680 the smallest fraction of HePS was isolated in NaG medium, followed by the glucose-supplemented NaG medium (Fig. 1C). Both media contain same amounts of sugar, but the respective sugars are metabolized in different ways, *inter alia* reflected in different pH-value profiles (Fig. 1B). Mannitol is oxidized by AAB to fructose by mannitol-dehydrogenase (A0U90_07780) [30], while glucose can be either directly activated to UDP-glucose and incorporated into the HePS biosynthesis, but also be oxidized to gluconic acid, which usually leads to decreasing pH values [13]. In the NaG-fgm media, the fastest cell growth of *K. baliensis* NBRC 16680 was observed (Fig. 1A). This is probably due to the doubled amount of (different) sugars in this medium (10 g d-fructose & 10 g d-glucose), compared to the NaG medium (10 g d-mannitol). Furthermore, there is a generally activating effect of magnesium, which can be seen by comparing the growth curves of NaG-fgm and NaG-fg medium. In NaG-fg medium, the cells grow more slowly than in the NaG-fgm medium, despite the same sugar concentrations. At the time point of HePS isolation (32 h), the bacteria within the four media were situated in the same growth phase and showed similar cell numbers (~$1*10^{10}$ CFU/mL). Thus, changing the carbon source does not result in a significantly increased biomass production in the stationary growth phase, whereas intercellular processes can be certainly influenced by variation of the carbon source as demonstrated by the largely different HePS yields (Fig. 1C). *K. baliensis* NBRC 16680 produced copious amounts of HePS in any media tested. However, lowest yields of HePS were produced in NaG media (1.73 ± 0.03 g/L), demonstrating that mannitol as the main carbon source is not an ideal choice for HePS production of *K. baliensis*. The effect of the carbon source becomes apparent when the HePS yields from NaG medium are compared to those of the NaG-g medium (2,64 ± 0.09 g/L), which contains the same amount of sugar/sugar alcohol. Other bacteria such as *Rhizobium tropici* or *Lactobacillus delbrueckii ssp. bulgaricus* also produce only low HePS amounts in medium with the sugar alcohol mannitol as the main carbon source [31,32]. A reason for the preferred utilization of other sugars might be the metabolic pathway of mannitol in AAB. Mannitol is first converted into fructose, by mannitol dehydrogenase, in order to be accessible for primary metabolism. The pentose phosphate route then probably further metabolizes a large proportion of the gained fructose, which is no longer available for HePS production. The respective increases in pH after ~10 h of fermentation (Fig. 1B) possibly resulted from overoxidation of acetate to CO_2 and H_2O and uptake and consumption of initially formed gluconic acid via the pentose phosphate way [30]. While *K. baliensis* NBRC 16680 seems to neutralize its growth environment independently of the provided carbon source, it could be recently shown that the formation of this acetan-like HePS is not necessary for acetic acid and ethanol tolerance in this strain [33]. Therefore, the differences in produced HePS amounts are more likely due to regulatory physiological processes in dependence of available sugar sources than to changes in the pH environment.

Besides the general improvement of the *K. baliensis* NBRC 16680 HePS yield, the structural and functional variability of the formed HePS was analyzed. Therefore, we analyzed the neutral sugar composition of *K. baliensis* NBRC 16680 HePS derived from media with different carbon sources (NaG, NaG-fgm, NaG-fg, NaG-g). It was demonstrated that the HePS of *K. baliensis* NBRC 16680 are composed of comparable amounts of glucose, galactose, and mannose, independent of the cultivation medium (Fig. 1D). Constant composition of the *K. baliensis* HePS is in agreement with the postulated progression of HePS synthesis [34]. HePS are formed via the sequential addition of nucleoside diphosphate sugars via glycosyltransferases (GT) to growing, repeating units of five to seven monomers. These repeating units are subsequently polymerized to form the final HePS [35] and therefore reflect the monomer composition of the entire HePS (Fig. 4). This kind of HePS synthesis was already shown for xanthan [36], and acetan, a HePS derived from *Ko. xylinus* [12,37]. In a previous study, we were able to show that the HePS cluster of *K. baliensis* strains NBRC 16680 has similarity to the *gum*-cluster of *X. campestris*, which is involved in the formation of xanthan [38,39], Furthermore, it shows high similarities with different HePS forming clusters of AAB, such as *Ga. diazotrophicus* [40] and *Ko. xylinus* [12]. Accordingly, it is assumed that the *K. baliensis* NBRC 16680 HePS is produced in a similar way, which would result in a constant structural composition of the HePS. Another study on acetan polysaccharides produced by *Ko. xylinus* demonstrated that the monomer composition of HePS is independent of the provided carbon source [8].

4.2. Structural characterization of the K. baliensis produced HePS in context to the connected HePS biosynthesis

To clarify the general structure of *K. baliensis* NBRC16680 HePS and to find potential similarities to acetan, structural analysis of the HePS was carried out by using HPAEC-PAD, methylation analysis, and NMR. Methylation analysis allowed for the detection of 1,4-/1,3,4-substituted glucose units, 1,2-substituted mannose units, and 1,4-substituted glucuronic acid units, which indicates the presence of an acetan-like core structure (Fig. 4B). This hypothesis is supported by the roughly equal portions of mannose and glucuronic acids, by the comparable amounts of 1,2-substituted mannose, 1,4-substituted glucose, and 1,3,4-substituted glucose units detected by methylation analysis, and by the presence of mainly β-linked glucose and/or glucuronic acid residues detected by NMR spectroscopy. The hypothesis that the monomer composition of the HePS is independent from the provided carbon source was supported by the comparable PMAA portions observed in the NaG and NaG-fgm derived HePS. However, some variations in the portions of the 1,6-linked glucose units were observed between the NaG and NaG-fgm derived HePS. Therefore, it cannot be excluded that small amounts of other polysaccharides are present or that minor deviations in the (side) chain length occur. It is also possible that the

small variations are a result of the different behaviors of the HePS during sample purification/treatment.

With the obtained data, it was not possible to elucidate the exact structure of the *K. baliensis* formed HePS. However, in our previous work, we were able to identify the HePS biosynthesis connected gene-cluster of *K. baliensis* NBRC 16680 [9]. This gene-cluster has similarities to the *gum*-cluster of *X. campestris*, responsible for xanthan gum formation, as well as to AAB gene clusters, which are involved in acetan formation of *Ko. xylinus* [38] and HePS formation of *Ga. diazotrophicus* [40]. By comparing the particular genes in the HePS cluster of *K. baliensis* NBRC 16680 with the corresponding genes in the HePS clusters of *X. campestris* and *Ko. xylinus* and by incorporating the knowledge of the detected linkage types of *K. baliensis* NBRC 16680 HePS, a hypothetical structure can be constructed (Fig. 4) [9]. In accordance with the β-1,4-linked glucose backbone of acetan and xanthan, the *K. baliensis* NBRC 16680 HePS contains 1,4-substituted glucopyranosyl residues [7,41] (Fig. 2A). The genes related to this linkage type are present in each HePS cluster (*K. baliensis* NBRC 16680, *X. campestris* and *Ko. xylinus).* The *gumD* gene codes for a protein that incorporates the initial glucose-1-phosphate residue and *gumM* codes for a protein that transfers a glucosyl residue from UDP-glucose to the previous glucose residue, resulting in a β-1,4-linked glucose-disaccharide [9,40,42] (Fig. 4). Furthermore, the second glucose residue is ramified with a mannose unit at position *O*3 in *X. campestris* and *Ko. xylinus*. The side chain is incorporated by an enzyme coded by *gumH* [10,42,43]. The *gumH* gene is also present in the HePS cluster of *K. baliensis* NBRC 16680 and 1,3,4-substituted glucose units/1,2-substituted mannose units were detected by methylation analysis, suggesting the presence of this structural element. The following step in the HePS synthesis is determined by the *gumK* gene, which is coding for a protein that adds a glucuronosyl residue from UDP-glucuronate to the glucose–glucose–mannose trisaccharide [44,45]. The *gumK* gene is also part of the HePS cluster of *K. baliens*is NBRC 16680 and consistently 1,4-substituted glucuronic acid was detected. In addition to a 1,4-substituted glucuronic acid unit, another putative uronic acid was found in *K. baliensis* NBRC 16680 HePS (Supplementary file 2). This could be caused by a modification of the glucuronic acid, which might occur before or after the incorporation of glucuronic acid into the repeating unit. In xanthan biosynthesis, the gene *gumI* is encoding for a glycosyltransferase that catalyzes the transfer of a mannosyl-residue from GDP-mannose to the growing xanthan tetrasaccharide. This gene is not present in the acetan and the *K. baliensis* NBRC 16680 HePS cluster (Fig. 4AC). Corresponding to this, different structures should be observable for xanthan and acetan and the *K. baliensis* NBRC 16680 HePS due to the varying biosynthesis. The mannose/glucuronic acid units are substituted by a terminal mannose-residue in the side chain of xanthan, whereas the side chains of acetan are further elongated by 1,6-substituted glucose units and finally terminated by a rhamnose unit [45]. In the *K. baliensis* HePS, high amounts of 1,6-linked glucose units were observed, too, whereas rhamnose was not detected by using monosaccharide and methylation analysis (Fig. 1,2). The gene that encodes for a rhamnosyl-transferase (*aceR*) in the acetan cluster [37] is absent in the *K. baliensis* NBRC 16680 HePS cluster, and consequently rhamnose residues are not incorporated into the HePS (Fig. 4). It is possible that the *K. baliensis* HePS is also substituted by one or more 1,6-substituted glucose residues, because relatively high amounts of 1,6-substituted and terminal glucose units were detected. *K. baliensis* NBRC 16680 shows the same GT genes (*aceP* and *aceQ*) as *Ko. xylinus* E25, which encode proteins that catalyze the incorporation of the two 1,6-linked glucose-1-phosphate to the glucose–glucose– mannose-glucuronic acid pentasaccharide [9]. Therefore, nearly all linkage types that are expected at the genetic level, based on the HePS cluster alignment, were identified via methylation analysis (Fig. 2). However, because xanthan and acetan from *X. campestris* and *Ko. xylinus* do not contain galactose [46], an exact assignment of the galactose residues in the *K. baliensis* (DSM 14400, NBRC 16680) HePS is currently not possible. Due to the presence of 1,6-linked galactose units (Fig. 2), we hypothesize that the galactose units are bound to or instead of the 1,6-linked glucose residues (Fig. 4C). BlastP alignments of the so far unassigned GT genes located upstream of the *gumJ* gene (Fig. 4A) [9] suggest that both GT (gt_{A0U90_06735}, gt_{A0U90_06740}) are associated with the incorporation of galactosyl-residues. Furthermore, similar glycosyltransferases are present in *Ga. diazotrophicus* (PAI5; GDIA_RS03825) [40], which also forms a HePS of glucose, galactose and mannose [15]. *Ga. diazotrophicus* PAI5 has a GT with homology to both of *K. baliensis* GTs (gt_{A0U90_06735}; Q 97%, identities 190/376 (51%) & gt_{A0U90_06740} Q 99%, identities 171/384 (45%)). The GT gt_{A0U90_06735} exhibits a predicted GT2_AmsE_like conserved domain (Bit Score: 119.34, E-value: 2.33e-32), which is involved in the amylovoran biosynthesis in the plant pathogenic bacteria *Erwinia amylovora*. AmsE is a glycosyltransferase, which is suggested to attach a galactose residue to the backbone of the chimeric EPS amylovoran [47]. Amylovoran is a complex anionic HePS composed of glucose, galactose and glucuronic acid, with 1,6-linked galactose residues [48]. The hypothesis that the 1,6-linked glucose/galactose units are incorporated into the HePS is further supported by the characterization of a HePS biosynthesis deficient mutant, which was not able to secrete any EPS [9]. However, from methylation analysis it cannot be concluded if glucose or galactose is bound to the uronic acid residues or if the 1,6-linked glucose/galactose residues alternate or appear blockwise. Nevertheless, the high portion of terminal glucose suggests that the side chains are terminated by glucose residues (Fig. 4C). The small portions of terminal mannose and terminal galactose are most likely derived from media components, co-precipitated oligo-/polysacccharides, or truncated side chains. It was previously demonstrated that acetan contains two acetyl-residues at specific positions [28] and that these residues are added by a protein encoded by the *gumF* gene. This gene is not present in the *gum*-like cluster of *K. baliensis* and the NMR spectra did not contain major non-sugar derived signals, suggesting that acetylation is not a structural feature of *K. baliensis* HePS. The combination of the results strongly suggests that the HePS produced by *K. baliensis* NBRC 16680 exhibits an acetan-core structure, which contains 1,6-linked glucose/galactose at currently non assignable position(s) and no rhamnosyl/acetyl residues in contrast to acetan.

4.3. *Rheological and macromolecular properties of HePS from* K. baliensis *NBRC 16680*

The HePS solutions of *K. baliensis* NBRC 16680 exhibited shear-thinning behavior with low flow behavior indices (n) obtained from the power law model (Fig. 3A). Solutions of *K. baliensis* HePS were highly viscous, even in low concentrations. By comparing HePS from NaG and NaG-fgm media it was demonstrated that HePS extracted from NaG media exhibited higher n-values than the HePS isolated from NaG-fgm media (Fig. 3A). *K. baliensis* HePS are possibly polyanionic as a result of the presence of glucuronic acid in the repeating units. The occurence of uronic acids enhances the anionic properties of the HePS, which possibly allows for the association of divalent cations such as calcium and magnesium and increases the interactions between the single strands [49]. However, the different n-values of the HePS produced in the two media (NaG, NaG-fgm) most likely do not result from a better entanglement of the dissolved HePS molecules in presence of magnesium in the NaG-fgm medium, because the cations were removed by dialyzing the HePS solutions during the purification process. Therefore, the different properties of the HePS produced in NaG-fgm medium could result from macromolecular effects, which are influenced by

magnesium during HePS biosynthesis. Magnesium often plays a crucial role in activation of enzymes that are involved in EPS biosynthesis, e. g. in the polymerization activity of glycosyltransferases [29,50]. Thus, NaG and NaG-fgm derived HePS from *K. baliensis* NBRC 16680 were further characterized via HPSEC and AF4-MALS analysis (Fig. 3BCD). The NaG-fgm derived HePS eluted later during AF4 separation and earlier during SEC separation, indicating a larger molecular weight of this HePS. Due to methodological limitations it was not possible to estimate the molecular weight. The rms radii of NaG-fgm HePS were indeed slightly higher (absolute determination from MALS signals independent of concentration signals). Based on the directed biosynthesis of HePS in general and the comparable monomer composition of all isolated HePS in this study (Fig. 1C), the repeating core EPS subunits are supposed to consist of a fixed monomer composition, whereas no clear assignment is possible for the detected 1,6-linked glucosyl/galactosyl residues as discussed above (Fig. 4C). However, magnesium in the medium may enhance the activation of enzymes, which are part of the HePS polymerization process and therefore oligomerization of the repeating core unit. Magnesium has been discussed in several studies as a potential promoter of EPS and biofilm production, especially related to the second messenger bis-(3′-5′)-cyclic-GMP [51–54]. The activity of the diguanylate cyclase is enhanced by magnesium, which leads to the formation of bis-(3′-5′)- cyclic dimeric GMP (c-di-GMP). The original discovery of the regulatory role of c-di-GMP in the synthesis of EPSs was established for cellulose synthesis in *Ko. xylinus*, where c-di-GMP acts as an allosteric, positive effector of cellulose synthase activity [50,55]. This stimulating effect can be probably traced back to an induced conformational change of the cellulose synthase subunit BcsA that allows UDP-glucose to access the catalytic site more effectively [56]. The situation is similar in case of alginate, produced by *Pseudomonas aeruginosa*, where the polymerization reaction of alginate requires c-di-GMP-binding to the inner-membrane protein Alg44; therefore increasing c-di-GMP levels enhanced the formation of alginate in mucoid strains [57]. Alg44 is a single-pass transmembrane protein that is involved in the polymerization and export of alginate across the inner membrane [58]. Consequently, magnesium influences the polymerization process of alginate and cellulose. Therefore, it may have similar effects on *K. baliensis* NBRC 16680 HePS synthesis, assuming an analogous function during the polymerization of the cellulose-like main chain/backbone of *K. baliensis* HePS. The responsible enzymes, which are possibly involved in the polymerization and export process of *K. baliensis* NBRC 16680 HePS are GumB, GumC and GumE [39]. GumC has three predicted trans membrane helices (TMH) and includes an N-terminal Wzz-domain (E-value: 1.23e-09, Bit Score: 55.82), and a C-terminal tyrosine autokinase domain (PEP-CTERM_TyrKin, E-value: 1.13e-14, Bit Score: 71.94). Both are characteristic for polysaccharide copolymerase 2 (PCP-2a) proteins, which regulate the chain length of high molecular weight EPS in *E. coli* [59,60]. Thus, magnesium in the medium may result in a longer main chain of the NaG-fgm medium derived *K. baliensis* NBRC 16680 HePS, resulting in larger molecules with a constant monomer composition similar to NaG medium derived HePS.

5. Conclusion

In this work, we described the structural characteristics of a novel HePS produced by *K. baliensis* NBRC 16680 based on methylation analysis, NMR spectroscopy, and on correlation of the obtained results to the previously identified HePS biosynthesis cluster. It was demonstrated that the HePS contains 1,4-substituted Glc*p* units and 1,3,4-substituted Glc*p* units, suggesting a (1 → 4)-linked acetan-like backbone with ramifications at position *O*3. Furthermore, we hypothesize – in correlation with the well described biosynthesis of the exopolysaccharide acetan –, that side chains are composed of 1,2-substituted mannose, 1,4-substituted glucuronic acid and 1,6-substituted glucose. In addition, 1,6-substituted galactose units as well as a putative uronic acid were detected, which might also be part of the side chains, but actually cannot be correlated with the responsible biosynthesis machinery. Rhamnose units and acetyl residues, which are structural elements of acetan, were not detected in the HePS of *K. baliensis* NBRC 16680. This observation is supported by the missing of homologous enzymes in the genome of *K. baliensis*, which incorporate rhamnose and acetyl residues into acetan. Instead of the terminal rhamnose in the side chains of acetan, a further attachment of 1,6-substituted glucose and 1,6-substituted galactose units is assumed. *K. baliensis* HePS derived from media with different carbon sources showed a very similar monomer composition, but varying rheological and macromolecular properties, which is possibly due to a different regulation of biosynthetic enzymes involved in HePS polymerization in dependence of initial carbon sources.

Acknowledgements

We want to thank Prof. Dr.-Ing. Heiko Briesen, Prof. Dr. Petra Först and Dipl.-Ing. Max Günther Müller (Lehrstuhl für Systemverfahrenstechnik, Technical University of Munich (TUM)) for giving us the opportunity to perform rheological measurements at their institute. Part of this work was funded by the German Federal Ministry for Economic Affairs and Energy via the German Federation of Industrial Research Associations (AiF) (FEI); project number AiF 18071N.

References

[1] I.S. Roberts, The biochemistry and genetics of capsular polysaccharide production in bacteria, Annu. Rev. Microbiol. 50 (1996) 285–315.
[2] A.S. Ferreira, J.H. Leita, I.N. Silva, P.F. Pinheiro, A.S. Sousa, C.G. Ramos, L.M. Moreira, Distribution of cepacian biosynthesis genes among environmental and clinical Burkholderia strains and role of cepacian exopolysaccharide in resistance to stress conditions, Appl. Environ. Microbiol. 76 (2) (2009) 441–450.
[3] T. Ophir, D.L. Gutnick, A role for exopolysaccharides in the protection of microorganisms from desiccation, Appl. Environ. Microbiol. 60 (2) (1994) 740–745.
[4] F. Jakob, Novel Fructans from Acetic Acid Bacteria, Technische Universität München, 2014.
[5] F. Jakob, D. Meißner, R.F. Vogel, Comparison of novel GH 68 levansucrases of levan-overproducing *Gluconobacter species*, Acetic Acid Bacteria 1 (1) (2012) 2.
[6] F. Jakob, A. Pfaff, R. Novoa-Carballal, H. Rübsam, T. Becker, R.F. Vogel, Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function, Carbohydr. Polym. 92 (2) (2013) 1234–1242.
[7] P. Jansson, J. Lindberg, K.M.S. Wimalasiri, M.A. Dankert, Structural studies of acetan, an exopolysaccharide elaborated by *Acetobacter xylinum*, Carbohydr. Res. 245 (1993) 303–310.
[8] H. Kornmann, P. Duboc, I. Marison, U.V. Stockar, Influence of nutritional factors on the nature, yield, and composition of exopolysaccharides produced by gluconacetobacter xylinus I-2281, Appl. Environ. Microbiol. 69 (2003) 6091–6098.
[9] J.U. Brandt, F. Jakob, J. Behr, A.J. Geissler, R.F. Vogel, Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*, Microb. Cell Fact. 15 (1) (2016) 170.
[10] P.L. Abdian, A.C. Lellouch, C. Gautier, L. Ielpi, R.A. Geremia, Identification of essential amino acids in the bacterial alpha-mannosyltransferase AceA, J. Biol. Chem. 275 (51) (2000) 40568–40575.
[11] J.K. Baird, P.A. Sandford, I.W. Cottrell, Industrial applications of some new microbial polysaccharides, Nat. Biotechnol. 1 (9) (1983) 778–783.
[12] A.M. Griffin, V.J. Morris, M.J. Gasson, Genetic analysis of the acetan biosynthetic pathway in *Acetobacter xylinum*, Int. J. Biol. Macromol. 16 (6) (1994) 287–289.
[13] H. Kornmann, P. Duboc, I. Marison, U. Stockar, Influence of nutritional factors on the nature, yield, and composition of exopolysaccharides produced by *Gluconacetobacter xylinus* I-2281, Appl. Environ. Microbiol. 69 (10) (2003) 6091–6098.
[14] G.J. Grobben, M.R. Smith, J. Sikkema, J.A.M. De Bont, Influence of fructose and glucose on the production of exopolysaccharides and the activities of enzymes involved in the sugar metabolism and the synthesis of sugar nucleotides in *Lactobacillus delbrueckii subsp. bulgaricus* NCFB 2772, Appl. Microbiol. Biotechnol. 46 (3) (1996) 279–284.

[15] R.V. Serrato, C.H. Meneses, M.S. Vidal, A.P. Santana-Filho, M. Iacomini, G.L. Sassaki, J.I. Baldani, Structural studies of an exopolysaccharide produced by *Gluconacetobacter diazotrophicus* Pal5, Carbohydr. Polym. 98 (1) (2013) 1153–1159.
[16] S. Moonmangmee, H. Toyama, O. Adachi, G. Teerakool, N. Lotong, K. Matsushita, Purication and characterization of a novel polysaccharide involved in the pellicle produced by a thermotolerant acetobacter strain, Biosci. Biotechnol. Biochem. 66 (4) (2001) 777–783.
[17] S. Moonmangmee, K. Kawabata, S. Tanaka, H. Toyama, O. Adachi, K. Matsushita, A novel polysaccharide involved in the pellicle formation of *Acetobacter aceti*, J. Biosci. Bioeng. 2 (2002) 192–200.
[18] S.V. Kaditzky, F. Rudi, Optimization of exopolysaccharide yields in sourdoughs fermented by lactobacilli, Eur. Food Res. Technol. 228 (2008) 291–299.
[19] K. Leontein, B. Lindberg, J. Lonngren, Assignment of absolute configuration of sugars by GLC of their acetylated glycosides formed from chiral alcohols, Carbohydr. Res. 62 (2) (1978) 359–362.
[20] G.A. De Ruiter, H.A. Schols, A.G.J. Voragen, F.M. Rombouts, Carbohydrate analysis of water-soluble uronic acid-containing polysaccharides with high-performance anion-exchange chromatography using methanolysis combined with TFA hydrolysis is superior to four other methods, Anal. Biochem. 207 (1) (1992) 176–185.
[21] F.M. Nunes, A. Reis, A.M.S. Silva, M.R.M. Domingues, M.A. Coimbra, Rhamnoarabinosyl and rhamnoarabinoarabinosyl side chains as structural features of coffee arabinogalactans, Phytochemistry 69 (7) (2008) 1573–1585.
[22] D. Wefers, M. Bunzel, Characterization of dietary fiber polysaccharides from dehulled, common buckwheat (*Fagopyrum esculentum*) seeds, Cereal Chem. 92 (6) (2015) 598–603.
[23] D.P. Sweet, R.H. Shapiro, P. Albersheim, Quantitative-analysis by various GLC response-factor theories for partially methylated and partially ethylated alditol acetates, Carbohydr. Res. 40 (2) (1975) 217–225.
[24] F.A. Pettolino, C. Walsh, G.B. Fincher, A. Bacic, Determining the polysaccharide composition of plant cell walls, Nat. Protoc. 7 (9) (2012) 1590–1607.
[25] H.E. Gottlieb, V. Kotlyar, A. Nudelman, NMR chemical shifts of common laboratory solvents as trace impurities, J. Org. Chem. 62 (21) (1997) 7512–7515.
[26] H.A. Barnes, J.F. Hutton, K. Walters, An Introduction to Rheology, Elsevier Applied Science, 1989.
[27] T. Ua-Arak, F. Jakob, R.F. Vogel, Characterization of growth and exopolysaccharide production of selected acetic acid bacteria in buckwheat sourdoughs, Int. J. Food Microbiol. 239 (2015) 103–112.
[28] I.J. Colquhoun, M. Defernez, V.J. Morris, NMR-studies of acetan and the related bacterial polysaccharide, CR1/4, produced by a mutant strain of *Acetobacter xylinum*, Carbohydr. Res. 269 (2) (1995) 319–331.
[29] M. Vargas-Garcıa, M.J. Lopez, M.A. Elorrieta, F. Suarez, J. Moreno, Influence of nutritional and environmental factors on polysaccharide production by *Azotobacter vinelandii* cultured on 4-hydroxybenzoic acid, J. Ind. Microbiol. Biotechnol. 27 (1) (2001) 5–10.
[30] U. Deppenmeier, A. Ehrenreich, Physiology of acetic acid bacteria in light of the genome sequence of Gluconobacter oxydans, J. Mol. Microbiol. Biotechnol. 16 (1–2) (2009) 69–80.
[31] A.K. Staudt, L.G. Wolfe, J.D. Shrout, Variations in exopolysaccharide production by Rhizobium tropici, Arch. Microbiol. 194 (3) (2012) 197–206.
[32] G.J. Grobben, S.J. Smith, M.R. Bont JAM de, Production of extracellular polysaccharides by *Lactobacillus delbrueckii ssp. bulgaricus* NCFB 2772 grown in a chemically defined medium, J. Appl. Microbiol. 79 (1) (1995) 103–107.
[33] J.U. Brandt, F.L. Born, F. Jakob, RF V: environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in Kozakia baliensis NBRC 16680, BMC Microbiol. (2017) (in press).
[34] J. Schmid, V. Sieber, Enzymatic transformations involved in the biosynthesis of microbial exo-polysaccharides based on the assembly of repeat units, Chembiochem 16 (8) (2015) 1141–1147.
[35] J. Schmid, S. Koenig, B. Rühmann, M. Rütering, Biosynthese und genomik mikrobieller polysaccharide, BIOspektrum 20 (3) (2014) 288–290.
[36] L. Ielpi, R.O. Couso, M.A. Dankert, Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*, J. Bacteriol. 175 (9) (1993) 2490–2500.
[37] T. Ishida, Y. Sugano, M. Shoda, Novel glycosyltransferase genes involved in the acetan biosynthesis of *Acetobacter xylinum*, Biochem. Biophys. Res. Commun. 295 (2) (2002) 230–235.
[38] F.J. Vorhölter, S. Schneiker, A. Goesmann, L. Krause, T. Bekel, O. Kaiser, B. Linke, T. Patschkowski, C. Rückert, J. Schmid, V.K. Sidhu, V. Sieber, A. Tauch, S.A. Watt, B. Weisshaar, A. Becker, K. Niehaus, A. Pühler, The genome of *Xanthomonas campestris pv: campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis, J. Biotechnol. 134 (1) (2008) 33–45.
[39] A. Becker, F. Katzen, A. Pühler, L. Ielpi, Xanthan gum biosynthesis and application: a biochemical/genetic perspective, Appl. Microbiol. Biotechnol. 50 (1998) 145–152.
[40] M. Bertalan, R. Albano, V. de Padua, L. Rouws, C. Rojas, A. Hemerly, K. Teixeira, S. Schwab, J. Araujo, A. Oliveira, et al., Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5, BMC Genomics 10 (2009) 450.
[41] P. Jansson, L. Kenne, B. Lindberg, Structure of the extracellular polysaccharide from *Xanthomonas campestris*, Carbohydr. Res. 45 (1) (1975) 275–282.
[42] N. Ielmini, F. Katzen, D.U. Ferreiro, C.G. Oddo, M. Vero, A. Becker, A. Pu, *Xanthomonas campestris pv. campestris* gum mutants: effects on xanthan biosynthesis and plant virulence, J. Bacteriol. 180 (7) (1998) 1607–1617.
[43] R.A. Geremia, M. Roux, D.U. Ferreiro, R. Dauphin-Dubois, A.C. Lellouch, L. Ielpi, Expression and biochemical characterisation of recombinant AceA, a bacterial alpha-mannosyltransferase, Mol. Gen. Genet. 261 (6) (1999) 933–940.
[44] M. Barreras, P.L. Abdian, L. Ielpi, Functional characterization of GumK, a membrane-associated beta-glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis, Glycobiology 14 (3) (2004) 233–241.
[45] E.A. Petroni, L. Ielpi, Isolation and nucleotide sequence of the GDP-mannose: cellobiosyl- gene from *Acetobacter xylinum*, J. Bacteriol. 178 (16) (1996) 4814–4821.
[46] R.O. Couso, L. Ielpi, M.A. Dankert, A xanthan-gum-like polysaccharide from *Acetobacter xylinum*, Microbiology 133 (1987) 2123–2135.
[47] C. Langlotza, M. Schollmeyera, D.L. Coplinb, M. Nimtzc, K. Geider, Biosynthesis of the repeating units of the exopolysaccharides amylovoran from *Erwinia amylovora* and stewartan from *Pantoea stewartii*, Physiol. Mol. Plant Pathol. 75 (4) (2011) 163–169.
[48] X. Wang, F. Yang, S.B. von Bodman, The genetic and structural basis of two distinct terminal side branch residues in stewartan and amylovoran exopolysaccharides and their potential role in host adaptation, Mol. Microbiol. 83 (1) (2012) 195–207.
[49] B. Vu, M. Chen, R.J. Crawford, E.P. Ivanova, Bacterial extracellular polysaccharides involved in biofilm formation, Molecules 14 (7) (2009) 2535–2554.
[50] P. Ross, M. Benziman, E. Vroom, A. Fidder, J.H. van Boom, The cyclic diguanylic acid regulatory in *Acetobacter xylinum* system of cellulose synthesis, Biochemistry 265 (31) (1990) 18933–18943.
[51] T.M. O'Shea, C.R. Deloney-Marino, S. Shibata, S. Aizawa, A.J. Wolfe, K.L. Visick, Magnesium promotes flagellation of *Vibrio fischeri*, J. Bacteriol. 187 (6) (2005) 2058–2065.
[52] T.M. O'Shea, A.H. Klein, K. Geszvain, A.J. Wolfe, K.L. Visick, Diguanylate cyclases control magnesium-dependent motility of *Vibrio fischeri*, J. Bacteriol. 188 (23) (2006) 8196–8205.
[53] M. Merighi, V.T. Lee, M. Hyodo, Y. Hayakawa, S. Lory, The second messenger bis- (3′ -5′) -cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*, Mol. Microbiol. 65 (4) (2007) 876–895.
[54] J.C. Whitney, P.L. Howell, Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria, Trends Microbiol. 21 (2) (2013) 63–72.
[55] W.H. Ross P, Y. Aloni, D. Michaeli, P. Weinberger-Ohana, R. Mayer, S. Braun, E. de Vroom, G.A. van der Marel, J.H. van Boom, M. Benziman, Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid, Nature 325 (1987).
[56] J.L.W. Morgan, J.T. McNamara, J. Zimmer, Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP, Nat. Struct. Mol. Biol. 21 (5) (2014) 489–496.
[57] M. Merighi, V.T. Lee, M. Hyodo, Y. Hayakawa, S. Lory, The second messenger bis- (3 '-5') -cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*, Mol. Microbiol. 65 (4) (2007) 876–895.
[58] L.L. Oglesby, S. Jain, D.E. Ohman, Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization, Microbiology 154 (6) (2008) 1605–1615.
[59] A.V. Franco, D. Liu, P.R. Reeves, The Wzz (Cld) protein in *Escherichia coli*: amino acid sequence variation determines o-antigen chain length specificity, J. Bacteriol. 180 (10) (1998) 2670–2675.
[60] L. Cuthbertson, I.L. Mainprize, J.H. Naismith, C. Whitfield, Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria, Microbiol. Mol. Biol. Rev.: MMBR 73 (1) (2009) 155–177.
[61] R. Overbeek, R. Olson, G.D. Pusch, G.J. Olsen, J.J. Davis, T. Disz, S. Edwards Ra Gerdes, B. Parrello, M. Shukla, et al., The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST), Nucleic Acids Res. (2014) 1–9.

PART III PAGES 65 - 76	**Environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in *Kozakia baliensis* NBRC 16680**

K. *baliensis* NBRC 16680 harbors two main HePS clusters, a *gum*-cluster coding for the previously described secreted HePS, and a *polABCD*E cluster coding for a HePS, forming a surface pellicle during static cultivation. Furthermore, *K*. *baliensis* NBRC 16680 comprises, compared to the DSM 14400 strain, a mobile element that inserts into the *gumD* gene in its activated state, leading to a morphology switch, from smooth wild-type (S) to a rough mutant colonies (R). To clarify, whether this is a random or directed switch, triggered by environmental factors, we checked the mutation frequency from smooth to rough colonies under different growth conditions. It could be presented, that the mutation frequency from smooth to rough strains was significantly increased upon growth on ethanol and acetic acid supplemented medium. However, both *K*. *baliensis* phenotypes (R and S) are still able to form a pellicle on the surface of the medium, whereas the morphologic switch got along with changes in the sugar contents of the produced pellicle HePS. To get closer insights into the composition and physiological function of the pellicle, a deletion of the *polE* gene of the *K*. *baliensis* NBRC 16680 mutant strain was conducted, as the *polE* gene was reported to be involved in pellicle formation in other AAB. According to literature, the *K*. *baliensis* R:Δ*polE* mutant strain was not able to form a regular pellicle anymore, but secreted an HePS into the medium, with a similar sugar monomer composition as the pellicle HePS from the R strain. However, solely the rough mutant strain showed in static and shaking growth experiments an markedly increased tolerance towards acetic acid and ethanol compared to the other NBRC 16680 strains (S, R:Δ*polE*). This increased tolerance is probably conditioned by a phenotypic transformation of *K*. *baliensis* NBRC 16680, from a secreted HePS in the wild type to a capsular HePS in the rough mutant strain. Thus, the results demonstrate that the alteration from a smooth to rough phenotype is presumably a genetically triggered event, which is related to an increased acetic acid and ethanol tolerance of the particular *Kozakia* strain. Moreover, the succeeded *polE* gene deletion in *K*. *baliensis* NBRC 16680 R, intended to clarify the involvement of the *polE* gene in the formation of a cell-associated, capsular polysaccharide, which seems to be essential for increased ethanol/acetic tolerance.

If *K*. *baliensis* DSM 14400 is included in this context, which shows a comparatively high intrinsic acetic acid tolerance from the beginning, it is likely that both strains have developed

different adaptive mechanisms to cope with changing natural habitats, like increasing acetic acid or ethanol concentrations. Hence, the genetic and morphological switch of *K*. *baliensis* NBRC 16680 could represent an adaptive evolutionary step during the bacterial development, which slowly leads to the drift of both strains.

Authors contributions: Julia U. Brandt was responsible for the independent planning and conducting the experiments. Julia U. Brandt evaluated the data and wrote the main text of the manuscript. Frank Jakob was the project leader responsible for supervision. Rudi Vogel conducted the conceptual setup and supervision of the study including infrastructural demands and final editing of the manuscript. Andi Geissler and Jürgen Behr were involved in the analysis of the whole genome sequencing data. Friederike L. Born was involved in the implementation of the *polE* knockout system.

Brandt et al. BMC Microbiology (2017) 17:172
DOI 10.1186/s12866-017-1070-y

BMC Microbiology

RESEARCH ARTICLE Open Access

Environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in *Kozakia baliensis* NBRC 16680

Julia U. Brandt, Friederike-Leonie Born, Frank Jakob* and Rudi F. Vogel

Abstract

Background: *Kozakia baliensis* NBRC 16680 secretes a *gum*-cluster derived heteropolysaccharide and forms a surface pellicle composed of polysaccharides during static cultivation. Furthermore, this strain exhibits two colony types on agar plates; smooth wild-type (S) and rough mutant colonies (R). This switch is caused by a spontaneous transposon insertion into the *gumD* gene of the *gum*-cluster, resulting in a heteropolysaccharide secretion deficient, rough phenotype. To elucidate, whether this is a directed switch triggered by environmental factors, we checked the number of R and S colonies under different growth conditions including ethanol and acetic acid supplementation. Furthermore, we investigated the tolerance of R and S strains against ethanol and acetic acid in shaking and static growth experiments. To get new insights into the composition and function of the pellicle polysaccharide, the *polE* gene of the R strain was additionally deleted, as it was reported to be involved in pellicle formation in other acetic acid bacteria.

Results: The number of R colonies was significantly increased upon growth on acetic acid and especially ethanol. The morphological change from *K. baliensis* NBRC 16680 S to R strain was accompanied by changes in the sugar contents of the produced pellicle EPS. The R:Δ*polE* mutant strain was not able to form a regular pellicle anymore, but secreted an EPS into the medium, which exhibited a similar sugar monomer composition as the pellicle polysaccharide isolated from the R strain. The R strain had a markedly increased tolerance towards acetic acid and ethanol compared to the other NBRC 16680 strains (S, R:Δ*polE*). A relatively high intrinsic acetic acid tolerance was also observable for *K. baliensis* DSM 14400^T, which might indicate diverse adaptation mechanisms of different *K. baliensis* strains in altering natural habitats.

Conclusion: The results suggest that the genetically triggered R phenotype formation is directly related to increased acetic acid and ethanol tolerance. The *polE* gene turned out to be involved in the formation of a cell-associated, capsular polysaccharide, which seems to be essential for increased ethanol/acetic tolerance in contrast to the secreted *gum*-cluster derived heteropolysaccharide. The genetic and morphological switch could represent an adaptive evolutionary step during the development of *K. baliensis* NBRC 16680 in course of changing environmental conditions.

Keywords: *Kozakia baliensis*, Heteropolysaccharides, Pellicle, Ethanol/acetic acid tolerance, Adaptive evolution

* Correspondence: frank.jakob@wzw.um.de
Technische Universität München, Lehrstuhl für Technische Mikrobiologie, Gregor-Mendel-Straße 4, 85354 Freising, Germany

Background

Gram-negative bacteria produce extracellular hetero- (HePS) or homopolysaccharides (HoPS), which are attached to the bacterial cell as capsular polysaccharide (CPS) or secreted into the environment as extracellular polysaccharide (EPS). Bacterial polysaccharides are important for the survival of bacteria, for instance in bacteria–host interaction, biofilm formation [1] and stress adaptation [2].

Acetic acid bacteria (AAB) are obligate aerobes and belong to the class of α-Proteobacteria. They are oxidative bacteria that strongly oxidize ethanol to acetic acid. AAB are well known for their ability to produce large amounts of EPSs, either HoPS, like dextrans, levans [3–5] and cellulose, or different kinds of HePS, such as acetan [6] and gluconacetan [7]. Furthermore, a variety of AAB has the ability to grow floating on the surface of a static culture by producing a pellicle enabling a high aeration state. The pellicle consists of an accumulation of cells, which are tightly associated with each other by capsular polysaccharides as connecting element. The pellicle CPS can be a HoPS of cellulose, which is produced by many *Komagataeibacter* species, like *Komagataeibacter (K.) xylinus* [8] (formerly *Gluconobacter (G.) xylinum* [9]), or a HePS, such as produced by many *Acetobacter* strains [10–12]. This HePS can be composed of different sugar monomers, like the HePS of *A. tropicalis* SKU1100. which consists of glucose, galactose, and rhamnose [13] or of *A. aceti* IFO3284 that contains only glucose and rhamnose [10].

The genes involved in the cellulose pellicle biosynthesis are arranged in an operon structure, like the acs operon [13] or the bcs operon [14], and widely studied. It is assumed that the genes involved in the synthesis of other pellicle HePS in AAB are assigned to a particular cluster, called *pol*-cluster. Deeraksa et al. (2005) could show, that the pellicle HePS produced by *A. tropicalis* SKU1100 could be traced back to a gene cluster, *polABCDE*, which is required for pellicle formation. In this operon, the *polABCD* genes showed high similarity to *rfbBACD* genes, which are involved in dTDP-l-rhamnose biosynthesis. The downstream located *polE* gene showed only low similarity to known glycosyltransferases, whereas a transposon-induced disruption of the *polE* gene resulted in a non pellicle forming strain, due to the absence of CPS production. Instead of the CPS, however, the Pel^- strain showed a smooth-surfaced colony and the HePS was now secreted into the medium, which had the same composition as the capsular pellicle polysaccharide [11].

*Acetobacte*r species are further known to exhibit high natural mutation frequencies [15, 16], also resulting in the formation of two or more different colony types. *A. pasteurianus* IFO3284 produces two altering types of colonies on agar medium that are inter-convertible by spontaneous mutation; a rough surface colony that can produce a pellicle (R strain) and smooth surface colony, which cannot produce a pellicle (S strain) [17]. Furthermore the R strains tolerate higher concentrations of acetic acid, whereas the pellicle formation is directly related to acetic acid resistance [18, 19]. It is further assumed that the pellicle CPS functions as a barrier-like biofilm against passive diffusion of acetic acid into the cells [18].

Microorganisms constantly face many difficult challenges, due to changing environmental conditions. The capacity to maintain functional homeostasis is essential for their survival. Recently, we have shown that the AAB *Kozakia (K.) baliensis* NBRC 16680 forms large amounts of a soluble unique HePS in the medium, as well as a pellicle during static cultivation [20]. Furthermore, *K. baliensis* NBRC 16680 forms a second type of colony form: a non-HePS producing rough-surfaced colony (R strain), caused by a transposon insertion into the *gumD* gene of the corresponding *gum*-like HePS cluster. The reason for this transposon insertion is unclear, whereby it was assumed that the transposon insertion represents a directed event, triggered by external factors.

Therefore, we investigated in this study, whether the morphology switch of *K. baliensis* NBRC 16680 is a random event, or triggered by environmental adaptations. In order to check if the morphological switch is connected to the formation of a CPS used in pellicle formation, we performed a *ΔpolE* deletion in *K. baliensis* NBRC 16680 R via a two step marker less gene deletion system. The different mutants were investigated regarding their growth and EPS production under different growth conditions including ethanol and acetic acid stress.

Methods

Bacterial strains, culture media, and culture conditions.

K. baliensis (NBRC 16680; National Institute of Technology and Evaluation (NITE) Biological Resource Center, Japan, DSM 14400; German Collection of Microorganisms and Cell Cultures(DSMZ)), as well as a mutant strain of *K. baliensis* NBRC 16680 R (*ΔgumD*) [20], with a rough phenotype, were used in this study. *K. baliensis* and its derivatives were grown at 30 °C in NaG media consisting of 20 g/L sodium gluconate, 3 g/L yeast extract, 2 g/L peptone, 3 g/L glycerol, 10 g/L mannitol and a pH adjusted to 6.0. *E. coli* strain TOP10 (Invitrogen, Karlsruhe, Germany), was grown at 37 °C and 180 rpm on a rotary shaker in LB medium consisting of 5 g yeast extract, 5 g NaCl and 10 g peptone. For selection of plasmids in *E. coli* TOP10. 50 μg/mL kanamycin was added to the LB medium. For selection of recombinant *K. baliensis* strains 50 μg/mL kanamycin or 60 μg/mL Fluorocytosin (FC) were used.

Convertibility of *K. bali*ensis NBRC 16680 from wild-type (S) to rough strains (R)

The number of R strains was determined during/after cultivation of *K. bali*ensis NBRC 16680 in standard NaG

medium, NaG medium supplemented with 3% ethanol (NaG-EtOH) or 0.4% acetic acid (NaG-AA), respectively. *K. baliensis* NBRC 16680 was first cultivated overnight in 10 ml of unmodified NaG medium at 30 °C (200 rpm). About 1×10^8 CFU/mL seed culture was afterwards transferred to 10 ml of standard or modified NaG medium, respectively. The flasks were incubated at 30 °C with rotary shaking at 200 rpm for 48 h and samples adducted at 0, 24 and 48 h. Cell numbers of wild-type and rough colonies (*ΔgumD*) were counted on non-modified NaG agar plates. Each growth experiment in liquid culture was performed thrice in separate assays, while each assay contained further three technical plating replicates.

For targeting the transposon insertion side, random colony PCRs of 23 rough colonies were carried out with Phire Hot start DNA polymerase (Thermo Fisher scientific; Waltham, USA). A primer set of a genomic primer (G4F_Fw) and a primer, targeting the mobile element (TE_Rv) were used; primers are listed in Additional file 1. PCR products were subsequently sequenced via Sanger sequencing by GATC Biotech (Konstanz, Germany).

Deletion of the *polE* gene with a two step marker less gene deletion system

For plasmids preparation, the GeneJET Plasmid Miniprep Kit (Thermo Fisher scientific, Waltham, USA) was used. Genomic DNA from *K. baliensis* NBRC 16680 R was extracted with the E.Z.N.A. Bacterial DNA Kit (Omega Bio-tek, Norcross, USA) and DNA purification was done with the E.Z.N.A. Cycle-Pure Kit (Omega Bio-tek, Norcross, USA). Restriction enzymes, DNA ligase, and alkaline phosphatase (FastAP) were obtained from Fermentas (Waltham, USA). PCRs were performed according to the Phusion High-Fidelity DNA Polymerase manuals from New England Biolabs (Frankfurt, Germany). For construction of the deletion vector, a fusion PCR technique was used to ligate the PCR products of flanking regions according to a long flanking homology (LFH) protocol [21, 22]. The length of the homology sequences were 20 bp. Primers are listed in Additional file 1. An enzyme-free cloning technique [23] was used for the further construction of the deletion vector, with the pKOS6b plasmid as basis, including a multiple cloning site (MCS), a kanamycine resistance gene (KM^R) and the *codA* and *codB* gene (Additional file 2A) [24]. Flanking regions of the *polE* gene, covering approximately 950 bp of the upstream and downstream region of the *polE* gene were amplified via PCR. The upstream region of 957 bp was amplified with primer P1_*polE*_KpnI_Fw and P2-*polE*_Rv, and a second primer set amplified the downstream region (968 bp) of the *polE* gene, containing the primers P3_*polE*_Fw and P4-*polE*_-XbaI_Rv (Additional file 1 & Additional file 2A). In the following step, a LFH PCR [21] was performed with P1_*polE*_KpnI_Fw and P4-*polE*_XbaI_Rv to merge the two previously amplified fragments (1898 bp). The fused fragment, as well as the pKOS6b vector, were both digested (KpnI, XbaI) and finally ligated. The resulting deletion vector, pKOS6bΔ*polE*, was verified via sanger sequencing (pK18MCS_Fw & pK18MCS_Rv), and further amplified in *E. coli* TOP 10. The transformation of pKOS6bΔ*polE* into *K. baliensis* NBRC 16680 R, was carried out by electroporation [25–28] with Gene Pulser Xcell™ Electroporation Systems from Bio Rad (München, Germany). Therefore, cells were inoculated to an OD_{600} of 0.3 in NaG medium and finally grown to an OD_{600} of 0.9. The culture was centrifuged at 5000 g, at 4 °C for 10 min, and washed three times in 1 mM HEPES buffer (pH 7). Cells were resuspended in 1 mM HEPES buffer, supplemented with ¼ volume of glycerin and shock frozen in 50 μL aliquots. The electroporation took place in cuvettes with 2 mm electrode distance from Bio Rad (München, Germany). The electroporation was carried out under constant conditions: 2.5 kV, 25 μF, and 400 Ω. Fresh enriched NaG medium (450 mM mannitol, 15 g/L yeast extract, 15 mM $CaCl_2$, 10 mM $MgSO_4$ and 6 mM glycerin) was added immediately after the pulse. The treated cells were incubated on a rotary shaker over 14 h and subsequently plated on NaG plates containing 50 μg/mL kanamycin for the first selection step with an incubation time of 48 h. During the first recombination step on NaG-Kan plates, a random chromosomal integration of the plasmid took place, which was checked by colony PCR using a specific primer set of a plasmid (pK18MCS_Fw or pK18MCS_Rv) and a genome (CL_*polE*_Fw or CL_*polE*_Rv) specific primer. Phire Hot start DNA polymerase (Thermo Fisher scientific; Waltham, USA) was used for colony PCR reactions, to screen for mutants or to confirm integration of the deletion vector into the genome. PCR products were sequenced via sanger sequencing by GATC Biotech (Konstanz, Germany). The positive clones were further grown on NaG-plates with 60 μg/ml FC, to drive the directed loss of the plasmid and the final selection of *K. baliensis* NBRC 16680 R Δ*polE* mutant colonies. After 3 days of incubation, the correct Δ*polE* mutant colonies could be identified via colony PCR, with a genome specific primer set (CL_*polE*_Fw & CL_*polE*_Rv) resulting in a 1950 bp fragment for the Δ*polE* mutant and a 3000 bp fragment for *K. baliensis* NBRC 16680 R (Additional file 2C).

Growth behavior of different *K. baliensis* strains in acetic acid and ethanol

*K. bali*ensis DSM 14400. NBRC 16680. the *ΔgumD* mutant [20] and the *ΔpolE* mutant strain (see 2.3) were grown on NaG agar plates, directly plated from the particular cryo stock, with either ethanol or acetic acid, in different concentrations. The ethanol supplemented plates contained 1% - 10% of ethanol (*v*/v, at intervals of 1%) and the acetic acid supplemented plates 0.1% - 1% of acetic acid (*v*/v, at

intervals of 0.1%). Each strain was streaked onto the plates from cryo-stocks and incubated for 3 days at 30 °C.

Furthermore, a static cultivation in NaG medium with 3% ethanol and 0.6% acetic acid was carried out. The *K. baliensis* NBRC 16680 strain, R mutant [20] and the *ΔpolE* mutant (see below) were grown as seed cultures in unmodified NaG media, overnight. Cultures were inoculated with an OD_{600} 0.3 and cultivated up to 0.9, respectively. For static cultures, 300 μl of the seed culture were inoculated into 3 ml NaG medium and cultivated statically at 30 °C. Cells were harvested by centrifugation at 6000 g and cell pellets were dried overnight at 120 °C. The dry weight was measured each day, over a time span of 7 days.

Analysis of HePS composition

Main cultures of *K. baliensis* NBRC 16680. *ΔgumD* mutant [20] and *ΔpolE* mutant (see below) were performed in 500 mL Erlenmeyer flasks with 50 mL of modified NaG media, inoculated with 500 μl of the pre-cultures and kept at 30 °C in a rotary shaker (200 rpm) for 32 h. Afterwards, cells were removed and the EPS containing supernatants were precipitated with cold ethanol (2:1, *v*/v) and kept overnight at 4 °C. This step was repeated three times, followed by a dialysis step (MWCO 14 kDa) of the recovered (centrifugation) and in ddH_2O re-dissolved HePS. Finally, the purified HePSs were lyophilized and quantified by weighing. To obtain large amounts of pellicle EPS, *K. baliensis* NBRC 16680 R [20] and the *ΔpolE* mutant (see 2.3) were cultured in unmodified NaG medium in cell culture flasks (Greiner Bio-One, Austria), to ensure a large surface for oxygen supply. Briefly, 10% of the seed culture was inoculated to 30 mL NaG medium and incubated statically at 30 °C for 14 days. The pellicle EPS was purified from the culture and cells were separated from EPS via ultra-sonification (10 min) and mechanical disruption, related to the method of IAI Ali, Y Akakabe, S Moonmangmee, A Deeraksa, M Matsutani, T Yakushi, M Yamada and K Matsushita [11]. The culture was centrifuged (10 min, 10.000 g) and the supernatant was saved in another flask. The cell pellet was washed 2 times with 10 mM HEPES buffer (pH 7), and suspended in the same buffer. The suspension was again ultra-sonicated for 10 min, followed by a centrifugation step for 10 min, 13.000 g. The resulting supernatant was combined, with the present supernatant from the first centrifugation, and EPS was precipitated with cold ethanol (2:1, *v*/v) and kept overnight at 4 °C. This step was repeated three times, followed by a dialysis step (MWCO 14 kDa) of the recovered (centrifugation) and in ddH_2O re-dissolved EPS.

The monosaccharide composition of the isolated *K. baliensis* NBRC 16680 R pellicle or secreted HePS (NBRC 16680. *ΔpolE*) was investigated via high performance liquid chromatography (HPLC). For HPLC analysis the purified polysaccharide samples were hydrolyzed with 10% of perchloric acid over 7 h at 100 °C, followed by a centrifugation step (4 °C, 10 min, 13,000 g) for removal of possible impurities, such as proteins. For the HPAEC analysis polysaccharide samples were hydrolyzed with 10% of perchloric acid over 2 h at 100 °C, as well followed by a centrifugation step (4 °C, 10 min, 13,000 g). The samples were further dissolved (1:10 or 1:100). The supernatant was analyzed using a Rezex RPM column (Phenomenex, Germany) coupled to a refractive index (RI) detector (Gynkotek, Germany) corresponding to the method of [29]. Sugar monomers were identified according to their retention time using suitable monosaccharide standards (D-glucose, D-galactose, D-mannose, D-rhamnose). The mobile phase was water, with a flow rate of 0.6 mL/min.

Results

Mutation of *K. baliensis* NBRC 16680 from S to R phenotype in dependence of different growth conditions

In *K. baliensis* NBRC 16680 spontaneous mutations occur, which cause a non-slimy phenotype, referred as rough (R) strain. In a previous work we have demonstrated, that this mutation can result from a transposon insertion in the *gumD* gene of the HePS forming cluster of *K. baliensis* NBRC 16680 [20]. The *gumD* gene encodes the first step of HePS formation, whereas a loss of the functional *gumD* gene leads to a total disruption of the HePS production and secretion in *K. baliensis* NBRC 16680. To clarify, if this mutation is a random event, or if it is a directed mutation possibly triggered by environmental factors, we performed an experimental series using different growth media. Different sugar combinations were tested (see 2.2) as well as stress inducing conditions like growth in ethanol (3%) or in acetic acid (0.6%) supplemented media. In addition, dilutions of the cryo-culture of *K. baliensis* NBRC 16680 were directly plated on NaG plates. A distinction was made between slimy glossy wild-type colonies (S) and rough dull mutant colonies (R), which were transparent, held against the light. The NaG plates with the directly plated dilutions of the cryo-cultures from *K. baliensis* NBRC 16680 showed only wild-type colonies. At 0 h, i. e. after inoculation with the overnight culture, solely wild-type colonies could be identified in standard NaG-medium, with cell numbers of about 1×10^8 CFU/mL (Fig. 1a). On the contrary, after inoculation of *K. baliensis* NBRC 16680 in NaG medium supplemented with ethanol or acetic acid, rough colonies could be detected at time point 0 in one of the three biological replicates, respectively (Fig. 1b*c*), which could have resulted from a too long contact to (and concomitant mutations in response to) acetic acid or ethanol. At 24 h the cell numbers of rough and wild-type strains were around $1{,}4 \times 10^9$ CFU/mL in the NaG medium and bacteria merged already into the starvation phase, with final cell counts of about 1×10^9 CFU/mL after 48 h (Fig. 1a). In

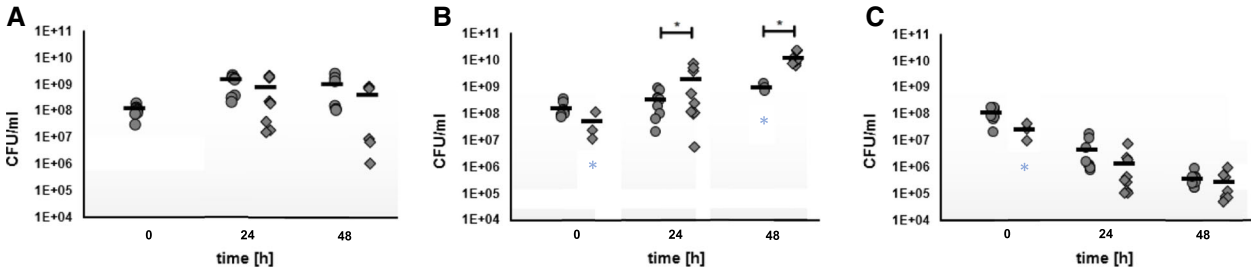


Fig. 1 Growth of *K. baliensis* NBRC 16680 in different media and influence of the provided carbon source on the morphology switch to a rough colony morphology (Δ*gumD*). **a** Growth of *K. baliensis* NBRC 16680 in NaG medium, $n = 9$ replica: wild-type (●), Δ*gumD* (♦), mean value (▬). **b** Growth of *K. baliensis* NBRC 16680 in NaG medium with ethanol (NaG-EtOH, 3%), $n = 9$ replica, * means that only $n = 3$ replica were above the detection limit of 10^4 CFU/mL: wild-type (●), Δ*gumD* (♦), mean value (▬).Asterisks centered over the error bars indicate the relative level of the *p*-value. In general, "*" means $p < 0.05$. (**c**) Growth of *K. baliensis* NBRC 16680 in NaG medium with acetic acid (NaG-AA, 0.4%), $n = 9$ replica, * means that only $n = 3$ replica were above the detection limit of 10^4 CFU/mL: wild-type (●), Δ*gumD* (♦), mean value (▬)

standard NaG medium the number of wild-type colonies exceled the number of mutant colonies or showed approximately equal numbers of wild-type and mutant colonies. During stress-inducing conditions (NaG-EtOH, NaG-AA), however, for NaG-EtOH an inverted picture emerged. At both survey marks (24 h, 48 h) significantly more mutant (R) than wild-type colonies were detectable, with around ten-power difference (Fig. 1b). After 48 h only in three of nine cases, wild-type colonies with a detection limit above of 10^4 CFU/mL could be detected (Fig. 1b*). In the NaG-AA medium, a continuous reduction of the cell numbers could be observed, in which the number of wild-type colonies always exceeded the number of mutated colonies (Fig. 1c).

In order to verify a possible integration of a mobile element in the *gumD* gene [20], or in front of the *gumD* gene, random colony PCR reactions were carried out, with a forward primer (G4F_Fw) targeting a location in front of the *gumD* gene (1.471.189–1.471.208 bp) and a reverse primer (TE-Rv) targeting the mobile element (Additional file 1) (Fig. 2b). Subsequently, the obtained PCR fragments were sequenced. For each mutated colony, a transposon insertion in the region of the *gumD* gene could be identified (Fig. 2c). These insertions, however, were not always located at the same site, but in a defined region of about 300 bp, around and in the *gumD* gene (Fig. 2c). Furthermore, it could be observed that in certain areas an integration of the mobile element occurred more often than in other areas, like up to nine times at 1471567 bp.

Influence of the *polE* gene on HePS formation and composition of *K. baliensis* NBRC 16680

In order to decode the role of the *polE* gene during pellicle formation of *K. baliensis* NBRC 16680 a *polE* deletion was carried out (see 2.3), with a markerless deletion system established by Kostner et al. [22, 24]. The pellicle forming ability of *K. baliensis* NBRC 16680 R Δ*polE* was further analyzed under static conditions in NaG-medium for 5 days. The growth behavior, as well as the HePS production, was compared with the wild-type and the rough mutant strain of *K. baliensis* NBRC 16680. All three *K. baliensis* strains showed pellicle formation after three days (Fig. 3a). For the wild-type strain of *K. baliensis* NBRC 16680 as well as for the rough mutant strain, a distinct pellicle formation at the edge of the test-tube was visible, which spreads over the entire boundary surface after three days. *K. baliensis* NBRC 16680 R Δ*polE* showed only a slight pellicle production (Fig. 3a). The physiology of the pellicle was different from the other two *K. baliensis* strains, instead of a surface spanning layer, only a loose conglomerate was present. The time for pellicle formation, also varied between the three strains. In comparison to the other *K. baliensis* strains, the pellicle of *K. baliensis* NBRC 16680 R Δ*polE* was only scarcely visible after three days of static incubation. During cultivation of the three *K. baliensis* strains (NBRC 16680, NBRC 16680 R and RΔ*polE*) under shaking conditions, no significant aberration in the growth behavior could be observed (Fig. 3b). In case of the HePS formation in shaking cultures with NaG medium, after 48 h a slight HePS production of *K. baliensis* NBRC 16680 R Δ*polE* could be demonstrated, with HePS amounts of 180 mg/L (Fig. 3c). The rough mutant *K. baliens*is strain, which differs only in the presence of an intact *polE* gene from *K. baliensis* NBRC 16680 R Δ*polE*, showed no EPS production in shaking cultures (Fig. 3c). The wild-type strain of *K. baliensis* NBRC 16680 with an intact *gum*- and *pol*-cluster, showed the highest EPS production under shaking conditions, with 1,73 g/L EPS.

K. baliensis NBRC 16680 HePS was furthermore isolated from shaking cultures and the monomer compositions were determined. It was possible to obtain pellicle EPS from the rough mutant strain (*K. baliensis* NBRC 16680 R), as well as EPS from shaking cultures from the *K. baliensis* NBRC 16680 wild-type strain and the *polE*

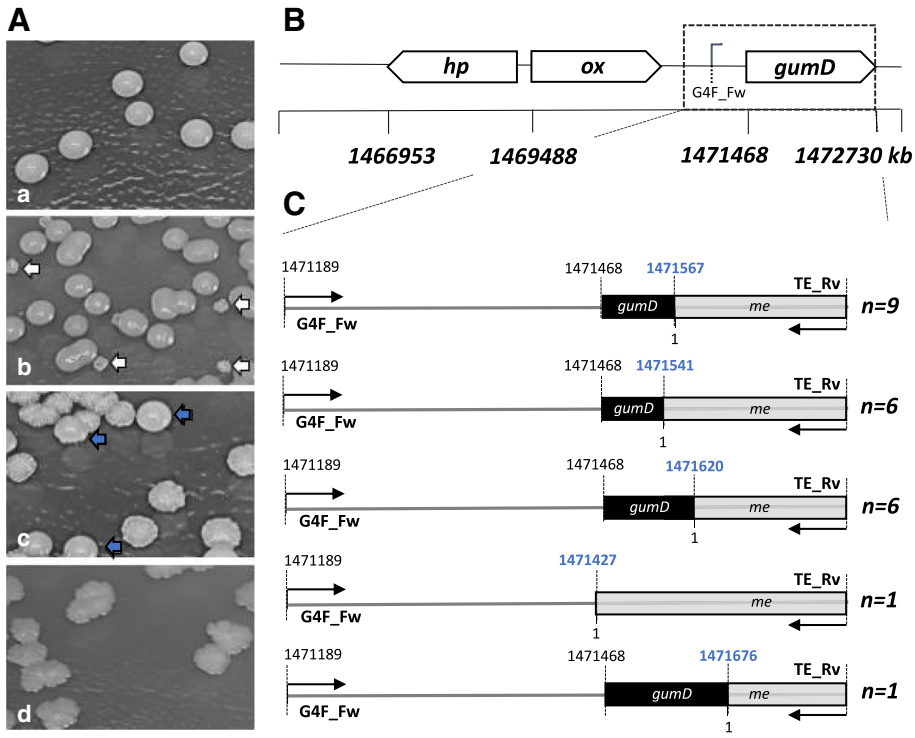


Fig. 2 Morphology and genetic switch of *K. baliensis* NBRC 16680 wild-type during cultivation in different media (**a**) In (*a*), growth of *K. baliensis* NBRC 16680 on NaG agar plates, plated at time point one after inoculation in NaG medium (0 h). Growth of *K. baliensis* NBRC 16680 on NaG agar plates, plated after 48 h of incubation in NaG medium (*b*), NaG-AA medium (*c*) and NaG-EtOH medium (*d*). Rough mutant colonies (Δ*gumD*) are indicated via a *white arrow*. Wild-type colonies are marked with a *grey arrow*. **b**) Shows a section of the *gum*-cluster of *K. baliensis* NBRC 16680 with the genomic location of the *gumD* gene (1471468–1,472,730 bp), oxidoreductase gene (*ox*, 1,469,488–1,470,627 bp) and a gene coding for a hypothetical protein (*hp*, 1,467,916–1,469,289 bp), based on JU Brandt, Jakob, F., Behr, J., Geissler, A.J., Vogel, R.F. [20]. Random colony PCRs of the respective rough colonies were carried out, targeting the transposon insertion side, using a genomic primer (G4F_Fw) and a primer, targeting the mobile element (TE_Rv). PCR products were subsequently sequenced. (**c**) Shows a schematic representation of the transposon insertion at the *gumD* locus in the rough colonys of *K. baliensis* NBRC 16680. The mobile element (*me*) is shown as grey bar with the corresponding insertion locus written in bold *blue*. The frequency of the found insertion site is marked on the left side, as *n* = x

deficient mutant, since for *K. baliensis* NBRC 16680 R, no EPS formation could be detected under shaking conditions. Because of the impaired pellicle formation of *K. baliensis* NBRC 16680 R:Δ*polE* (even during 14 days of cultivation in 30 mL cell culture flasks), no adequate amounts of EPS could be isolated for the monomer analysis. The monomer composition of wild-type *K. baliensis* NBRC 16680 HePS was composed of D-glucose, D-mannose and D-galactose (Fig. 3d). The statically cultivated rough mutant strain of *K. baliensis*, which is not able to produce HePS via the *gum*-cluster, showed a divergent monomer distribution, with generally higher amounts of D-mannose (46,2 ± 3,20%), and a consequently lower D-glucose (33,9 ± 5,03%) and D-galactose (19,9 ± 1,83%) level (Fig. 3d). The HePS derived from shaking cultures of the *polE* deficient *K. baliensis* NBRC 16680 R mutant showed a similar monomer distribution as the pellicle HePS of *K. baliensis* NBRC 16680 R, but with slightly variable percentages (Fig. 3d). Compared to *K. baliensis* NBRC 16680 R, a smaller proportion of D-glucose (21,8 ± 1,03%) was observable, while the proportion of D-mannose (51,4 ± 1,20%) was still higher.

Physiological effects of Δ*gumD* and Δ*gumD* + Δ*polE* mutations

Since the insertion of a mobile element into the *gumD* gene of the *gum*-cluster of *K. baliensis* NBRC 16680 appears to be a mechanism affected by external factors, the question arises, whether this is a physiological adaptation of the bacterium to changes in its external environment. Therefore, the growth behavior of the different *K. baliensis* strains, under different cultivation conditions, was investigated. *K. baliensis* NBRC 16680 and DSM 14400, *K. baliensis* NBRC 16680 R and the Δ*polE* mutant of the rough *K. baliensis* strain were plated on NaG-medium plates, which were supplemented with different acetic acid (0.1%, 0.2%, 0.3%, 0.4%, 0.5%, 0.6%, 0.7%, 0.8%, 0.9% and 1%) and ethanol (1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% and 10%) concentrations. Furthermore, the growth behavior of *K. baliensis* DSM 14400

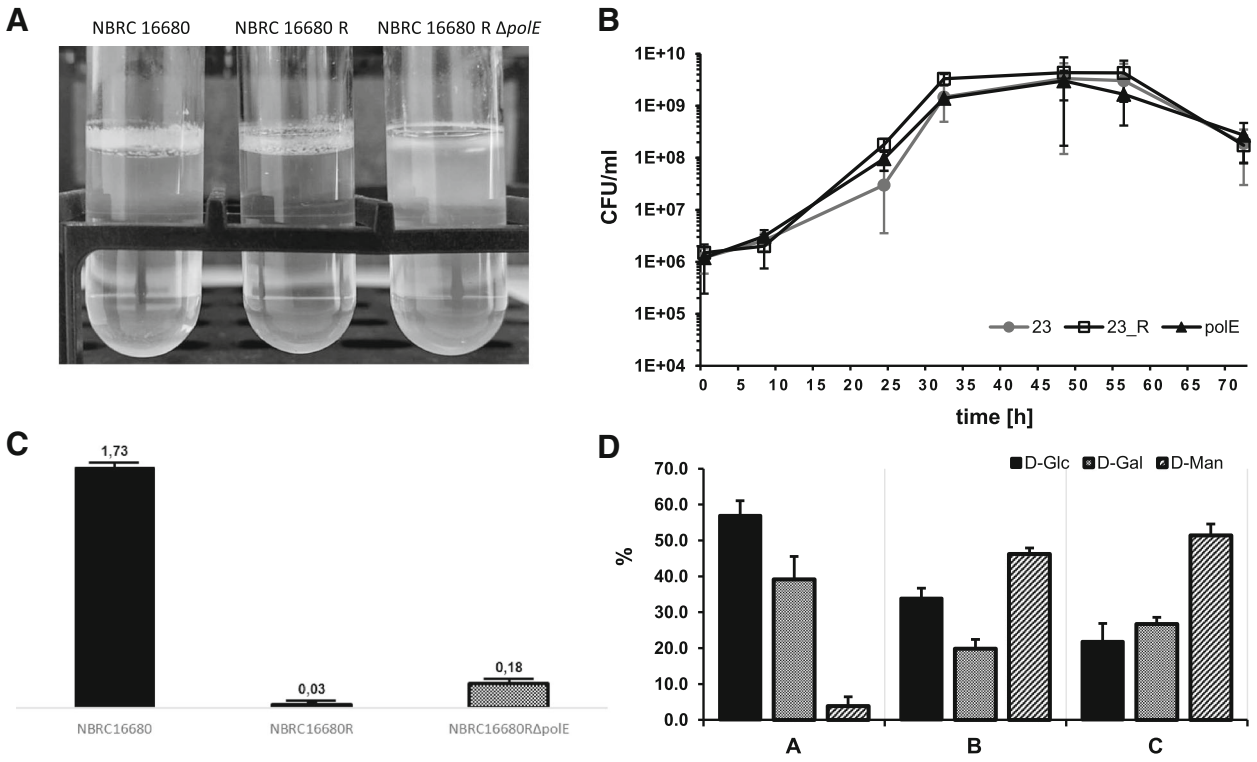


Fig. 3 Comparison of *K. baliensis* NBRC 16680 wild-type, the rough mutant strain and the *polE* deficient strain. **a** Phenotyps of *K. baliensis* NBRC 16680 during static cultivation of wild-type, rough mutant (*K. baliensis* NBRC 16680 R) and rough Δ*polE* mutant strain (*K. baliensis* NBRC 16680 R Δ*polE*). **b** Growth behavior of the different *K. baliensis* strains (NBRC 16680 (•), NBRC 16680 R (□), NBRC 16680 R Δ*polE*(▲)) in NaG medium. (**c**) Amount of precipitated HePS [g/L] from different *K. baliensis* strains (NBRC 16680. NBRC 16680 R, NBRC 16680 R Δ*polE*) during growth in NaG medium. (**d**) Neutral sugar composition of the isolated HePS for different *K. baliensis* strains during shaking or static cultivation. The proportions of the monomer occurring in the particular HePS are represented as percentage. (*a*) NBRC 16680 (shaking), (*b*) NBRC 16680 R (static), (*c*) NBRC 16680 R Δ*polE* (shaking)

was additionally tested, to investigate the variability in the growth behavior between two different wild-type strains of *K. baliensis*. *K. baliensis* DSM 14400 forms a HePS which is also composed of D-glucose, D-galactose and D-mannose, but with a deviating ratio compared to NBRC 16680 [20]. After a cultivation period of three days, distinct differences in the growth behavior of the different strains could be observed. Acetic acid had a significant influence on the growth of the different *K. baliensis* NBRC 16680 strains, while 0.7% acetic acid was the highest concentration, at which colony formation on agar plates was still possible for *Kozakia* strains NBRC 16680 R and DSM 14400 (Fig. 4a). The *K. baliensis* NBRC 16680 wild-type and the Δ*polE* mutant of the rough *K. baliensis* NBRC 16680 strain were not able to form colonies above 0.6% of acetic acid. In case of growth on NaG-EtOH plates, *K. baliensis* NBRC 16680 R and *K. baliensis* NBRC 16680 Δ*polE*, showed the highest EtOH tolerance, whereas *K. baliensis* NBRC 16680 R could even grow at 10% EtOH. Both wild-type strains of *K. baliensis* (NBRC 16680, DSM 14400) showed only poor growth behavior upon 7% EtOH and no growth at all, above 8% EtOH in the medium (Fig. 4b).

Furthermore, the effect of acetic acid and ethanol on the growth behavior of the different *K. baliensis* strains was monitored under static cultivation, by measuring the dry weight of the cells. It has to be noted, that the secreted, *gum*-cluster based HePS of the wildtype strain from *K. baliensis* NBRC 16680 was largely removed by centrifugation in advance. However, it can still slightly contribute to the dry weight as a small residue, which is still connected with the bacteria. In the NaG medium with 0.6% of acetic acid, the rough mutant strain displayed the fastest growth and reached the highest dry weight, after 7 days (1,77 ± 0.19 mg/ml). The wild-type strain of *K. baliensis* NBRC16680 showed slightly reduced final dry weight (1,64 ± 0.04 mg/ml), but an offset log phase resulting in a slower growth rate than the rough mutant strain (Fig. 5a). Similar results were obtained for the growth of the corresponding strains in NaG medium with 3% ethanol. *K. baliensis* NBRC16680 R exhibited the fastest growth, with a final dry weight of 1,17 ± 0.16 mg/ml. For *K. baliensis* NBRC16680 and the Δ*polE* mutant strain, a markedly reduced growth could be demonstrated (Fig. 5b).

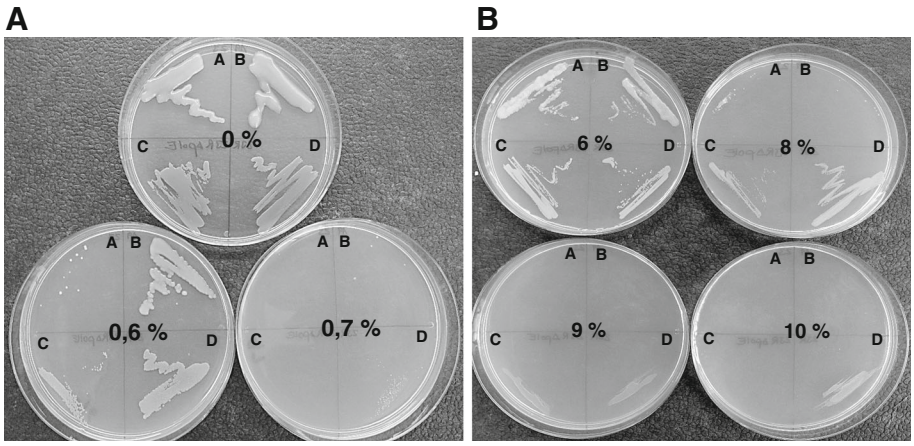


Fig. 4 Effect of acetic acid and ethanol on the growth behavior of different *K. baliensis* strains (**a**) Effect of acetic acid on growth of different *K. baliensis* strains (DSM 14400 (*a*), NBRC 16680 (*b*), NBRC 16680 R Δ*polE* (*c*), NBRC 16680 R (*d*)) at different acetic acid concentrations (0%, 0.6%, 0.7%). (**b**) Effect of ethanol on growth of different *K. baliensis* strains (DSM 14400 (*a*), NBRC 16680 (*b*), NBRC 16680 R Δ*polE* (*c*), NBRC 16680 R (*d*)) at different ethanol concentrations (6%, 8%,9%, 10%)

Discussion

K. baliensis NBRC 16680 produces and secretes large amounts of HePS via a *gum*-cluster encoded HePS biosynthesis. By insertion of a mobile element into the *gumD* locus of *K. baliensis* NBRC 16680 a mutant strain of *K. baliensis* is formed (*K. baliensis* NBRC 16680 R), which is unable to produce and secrete the *gum*-cluster derived HePS [20]. After cultivation of *K. baliensis* NBRC 16680 in standard or modified media (NaG, NaG-EtOH, NaG-AA), *K. baliensis* NBRC 16680 rough mutant strains were more frequently found in the presence of ethanol in the medium. For AAB an induced loss of various physiological properties has been observed, such as acetic acid resistance [30], ethanol oxidation, pellicle formation [31], and bacterial cellulose synthesis. The genetic mechanisms behind these instabilities is often unclear. Takemura et al. [32] reported that the loss of the ethanol oxidation ability in *A. pasteurianus* NC11380 occurs by an insertion of a mobile element into the alcohol dehydrogenase-cytochrome c gene, resulting in the loss of alcohol dehydrogenase activity. This has also been demonstrated for *A. pasteurianus* NCI 1452, where an insertion sequence element (IS) is associated with the inactivation of the alcohol dehydrogenase gene [33]. Gene inactivation provoked by transposable elements could also be demonstrated for EPS forming clusters of AAB, like the cellulose synthase operon. In *A. xylinum* ATCC 23769, an IS element caused insertions 0.5 kb upstream of the cellulose synthase gene, associated with spontaneous cellulose deficiency [34]. In both cases, it was not possible to sustain revertants of the cellulose synthesis and ethanol oxidation

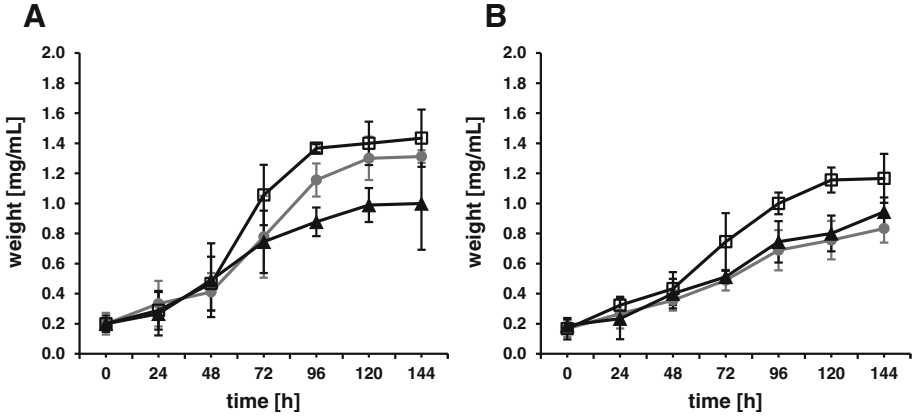


Fig. 5 Growth behavior of different *K. baliensis* strains during static cultivation in altered media (**a**) Growth behavior of the different *K. baliensis* strains (NBRC 16680 (•), NBRC 16680 R (□), NBRC 16680 R Δ*polE*(▲)) during static cultivation in NaG medium with acetic acid (0.6%). (**b**) Growth behavior of the different *K. baliensis* strains (NBRC 16680 (•), NBRC 16680 R (□), NBRC 16680 R Δ*polE*(▲)) during static cultivation in NaG medium with ethanol (3%) The growth was monitored by measuring the dry weight of cells

insufficient mutant strains, possibly as a consequence of remaining directed repeats (DR) of the IS element after relocation. Also in case of one previously investigated *K. baliensis* NBRC 16680 R mutant strain, the formation of directed repeats at the transposon insertion side could be observed [21]. Moreover, insertions of the mobile element in front of the *gumD* locus of the R strains as detected in the present study could result in a blocked transcription of *gumD* due to the presence of an energy-rich stem loop structure in the mobile element possibly causing *rho*-independent transcription termination [20]. It can actually not be ruled out that further mutations in the genome of *K. baliensis* NBRC 16680 simultaneously occurred. However, transposon insertion at the *gumD* locus of *K. baliensis* NBRC 16680 seems to be directly involved in R phenotype formation, since all of the 23 PCR checked R mutants exhibited the transposon insertion side at the *gumD* locus.

Morphotype variations, including the transition from a mucoid to a non-mucoid phenotype, are common events within the family of *Acetobacteraceae*, especially for *Acetobacter* species. This phenotypic change is often connected with the ability of the bacteria to form a pellicle on the medium surface [10, 11]. The pellicle is an assemblage of cells that permits them to float on the medium surface during static cultivation and ensures a high state of aeration. It was also shown that pellicle production could be associated with a change from a smooth phenotype to the rough, pellicle-forming strain. This change is accompanied by a transformation from secreted EPS to CPS, which could serve as a better barrier against ethanol [18]. Also the rough mutant of *K. balien*sis NBRC 16680 is still able to form a pellicle under static cultivation [20], suggesting that a second cluster, instead of the *gum*-cluster, is responsible for pellicle construction. For different *Acetobacter* strains, it has been shown, that the so-called *pol*-cluster is responsible for pellicle formation [35]. The *polABCDE* cluster shows a high level of homology to the *rfbBACD* genes of Gram-negative bacteria, which are involved in dTDP-rhamnose synthesis [36]. The *polE* gene has already undergone several assignments, since it has a relatively low homology level to glycosyltransferases, in general. A disruption of the *polE* gene in *A. tropicalis* SKU1100. however, leads to a defect in pellicle formation, thereby giving it a central role, either as rhamnosyl-transferase [35], or as galactosyl-transferase [11], which connects the CPS to the cell surface. Moreover, the mutant Pel^{-} cells secreted EPS into the culture medium. To investigate the relationship between *pol*-cluster and pellicle/CPS formation in *K. baliensis* NBRC 16680. *polE* knockouts were carried out in *K. baliensis* NBRC 16680 R, which does not form a *gum*-cluster dependent HePS [20]. Furthermore, we were interested to see, if a *polE* knockout has an effect on the ethanol tolerance of the respective *K. baliensis* strain. The Δ*polE* mutant of *K. baliensis* NBRC 16680 R showed the same rough colony morphology as the rough strain (*K. baliensis* NBRC 16680 R). In contrast to *A. tropicalis* SKU1100. *K. baliensis* NBRC 16680 R Δ*polE* was still able to form a pellicle on the surface, which was, however, only a loose conglomerate of cells and required considerably more time for formation. Furthermore, the Δ*polE* mutant was able to secrete small amounts of EPS into the medium. The incoherent pellicle could therefore be formed by the secreted EPS, thus leading to the formation of a weak EPS/ cell-layer on the surface. The secreted EPS showed a similar composition as the *K. baliensis* NBRC 16680 R capsular HePS. This supports the hypothesis of Ali et al. [11] that PolE is responsible for CPS formation, via addition of some residue(s) that connect the HePS with the cell surface, e.g. β-d-galactopyranosyl residues. A *polE* gene deletion or interruption results in a switch from a rough CPS producing, to an EPS producing phenotype, connected with growth behavior variations [35].

It was shown that a change from a smooth phenotype to the rough phenotype is associated with an increased tolerance against acetic acid [35]. By static cultivation and in combination with growth experiments on NaG plates with different ethanol and acetic acid concentrations of the different *K. baliensis* strains (NBRC 16680, NBRC 16680 R, NBRC 16680 R Δ*polE*), variations in the growth behaviors could be observed. Additionally to *K. baliensis* NBRC 16680 a second *K. baliensis* strain (DSM 14400) was tested, to investigate the variability in the growth behavior between two wild-type strains of *K. baliensis*. The rough mutant strain could form colonies on NaG plates under both ethanol and acetic acid stress in contrast to the wild-type NBRC 16680 even at high acetic acid (0.7%) and ethanol (10%) concentrations. During static growth of *K. baliensis* NBRC 16680 and the rough mutant, a similar pattern compared to the NaG plates appeared, with higher growth rates of *K. baliensis* NBRC 16680 R in NaG medium with acetic acid and ethanol. The Δ*polE* strain, however, showed a diminished growth on acetic acid (up to 0.6%) and ethanol (up to 9%) agar plates, as well as during steady state growth. It is assumed, that the knockout of the *polE* gene results in a change from CPS to EPS, whereas CPS serves as a better barrier against acetic acid [37]. This indicates that the CPS of *K. baliensis* NBRC 16680 R should be involved (inter alia) in the protection against acetic acid and ethanol. The direct relationship between pellicle formation and acetic acid resistance could be proven for *A. pasteurianus* (IFO3283, SKU1108, MSU10), where the rough strains had clearly higher acetic acid resistance abilities than the smooth phenotypes, respectively [19]. The pellicle functions in this case as a biofilm-like barrier and prevents the passive diffusion of acetic acid into the cells [38]. Furthermore, Perumpuli et al. [18] could show, that ethanol in the medium significantly induced pellicle formation. This is in agreement with our observations for *K. baliensis* NBRC 16680, which

mutated more frequently in NaG medium with ethanol, suggesting a directed mutation, which could at least partially be triggered under acetic acid fermenting conditions by ethanol or its oxidized product acetic acid.

K. baliensis DSM 14400 showed, similar to *K. baliensis* NBRC 16680 R, a high resistance against acetic acid, but was derogated by ethanol concentrations up to 8%. The high tolerance of *K. baliensis* DSM 14400 against acetic acid shows, that this strain has a different defense strategy in dealing with acetic acid, compared to *K. baliensis* NBRC 16680. Both strains were isolated from different environments (DSM 14400 isolated from palm brown sugar, NBRC 16680 isolated from ragi [39]). *K. baliensis* DSM 14400 might, therefore, have been commonly confronted with higher acetic acid concentrations, while they represent a new environmental factor for *K. baliensis* NBRC 16680. In the genome of *K. baliensis* DSM 14400. a cellulose synthase operon was identified on plasmid 3 (pKB14400_3) in contrast to *K. baliensis* NBRC 16680 including genes encoding the three cellulose synthase subunits A, B and C [20], suggesting that additionally formed cellulose could act as barrier against acetic acid. This is also the case for *Komagataeibacter (Ko.) xylinus* E25, isolated from vinegar, which generally deals with high acetic acid concentrations. Via comparison of diverse *pol* clusters from different acetic acid bacteria, a direct genetic connection between the *pol*- and *gum*-like clusters can be observed in *Ko. xylinus* (Fig. 6). *Ko. xylinus* E25 contains no *polE* gene in its genome, but is well known to produce a pellicle polysaccharide consisting of bacterial cellulose [13] and is also able to produce an extracellular HePS called acetan, composed of D-glucose, D-mannose, D-rhamnose, and D-glucuronic acid [6, 40]. In case of *Ko. xylinus* E25, the *pol*-cluster is flanking the acetan producing *gum*-like cluster, with the *polAB* gene upstream and the *polCD* gene downstream of the cluster (Fig. 6a). Acetan includes rhamnosyl residues, whose activated form is processed by *polABCD* [35], coding for the enzyme of the TDP-rhamnose synthesis, connecting the acetan synthesis with the enzymes of the *pol*-cluster. The *polE* gene is missing in the *Ko. xylinus* E25 genome (GCA_000550765.1), possibly resulting in no HePS mediated pellicle formation, while a tight cellulose pellicle is commonly formed on the surface of the medium [41], which again contributes to increased acetic acid resistance of the bacterium [42]. This shows that the *gum*-like HePS biosynthesis and the *pol*-cluster are possibly related with each other and are most likely regulated according to the environmental requirements of the respective bacterium. A lack of oxygen, or an increasing ethanol concentration, could be a signal, which leads to an up-regulation of the *polE* gene in *K. baliensis* NBRC 16680. This can take place in association with an ethanol or acetic acid triggered stress response, probably resulting in an induced deactivation of the *gum* clusters, via a transposon insertion.

In the genome of *Gluconacetobacter (Ga.) diazotrophicus* PAl5, no *polE* gene is present as well (Fig. 6b). In contrast to *Komagataeibacter xylinus*, *Ga. diazotrophicus* is an obligate endophytic bacterium that lives symbiotically (N_2-fixing) in the intercellular space of roots, stem and leaves of sugarcane plants [30]. It produces a HePS composed of D-glucose, D-galactose and D-mannose, whose production is most likely attributable to its *gum*-like cluster similar to *K. baliensis* [31]. The produced HePS plays a key role during plant colonization, as shown for the molecular communication between the colonized plants and *Ga. diazotrophicus* [32]. The genetic switch of *K. baliensis* NBRC 16680 wild-type to its rough mutant could indicate an environmentally driven conversion/adaptation of a formerly more plant associated AAB (wild-type) to an increasingly acetic acid producing AAB (rough mutant). Furthermore, *pol*-cluster mediated capsular pellicle polysaccharide biosynthesis and concomitant inactivation of excess HePS production and secretion seems to increase the tolerance against

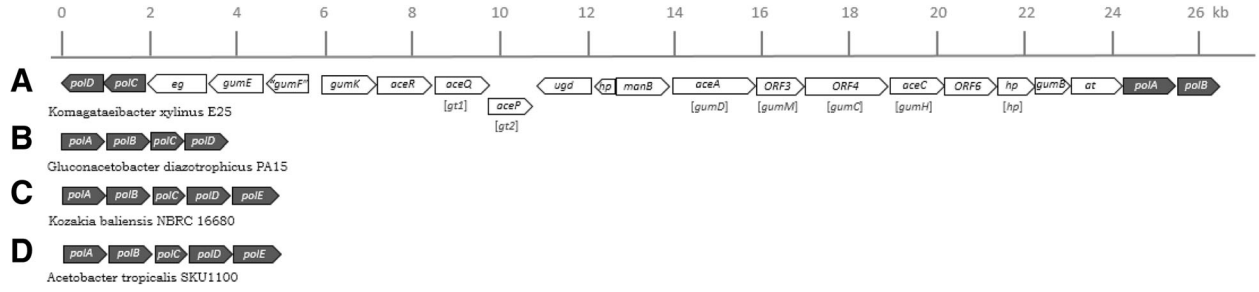


Fig. 6 Genetic organization of the *pol*-clusters from different acetic acid bacteria. The *pol*-clusters in combination with the *gum*-like cluster, responsible for the acetan synthesis, of *Ko. xylinus* E25 is depicted in (**a**). The so called acetan cluster has an overall size of ~26 kb and involves 22 genes, including glycosyltransferases (*gt*), mannose-phosphate-guanyltransferase (*mpg*), rhamnosyl transferase (*aceR*), hypothetical proteins (*hp*), and eight *gum* like genes marked as brackets under the particular genes (*gumB, –C, –D, –E, –F, –H, –K, and –M*). Furthermore, the cluster is flanked with the *polAB* gene upstream and the *polCD* gene downstream of the cluster. The nomenclature for the acetan cluster is based on AM Griffin, Morris, V. J., Gasson, M. J. [33]. In (**b**) is the *pol*-cluster of *Ga. diazotrophicus* Pal5, which consists of the *polABCD* genes. **c** and (**d**) show the related *pol*-clusters of *K. baliensis* NBRC 16680 [20] and *A. tropicalis* SKU1100 [35], including the *polABCDE* genes

acetic acid in certain AAB such as *K. baliensis* NBRC 16680, while specialized AAB additionally produce tight cellulose pellicles, which are typical for extremely acetic acid tolerant starter cultures used in vinegar production such as *Komagataeibacter xylinus*.

Conclusion

In summary the results obtained in this study show, that a switch from the *K. baliensis* NBRC 16680 wild-type to a rough mutant strain leads to a significantly increased acetic acid and ethanol resistance. This increased tolerance is probably accompanied by a morphologic switch, from secreted HePS in the wild type to capsular HePS in the rough mutant strain (Fig. 3d). The *polE* gene turned out to be involved for the formation of the resulting CPS. Although the exact role of PolE is still unknown, the results presented here show that *K. baliensis* NBRC 16680 R Δ*polE* displays a reduced tolerance against acetic acid and ethanol, most likely caused by a lack of cell-bound CPS and thus of pellicle formation. Since the morphological change is not reversible, this can also be understood as an adaptive evolutionary step of *K. baliensis* NBRC 16680 resulting in a shift of its ecological niche more towards an acetic acid-rich milieu.

Additional files

Additional file 1: Primers used for *polE* gene deletion in *K. baliensis* NBRC16680 RTable of primers used fort the *polE* gene deletion in *K. baliensis* NBRC 16680 R. (PPTX 341 kb)

Additional file 2: Deletion of the *polE* gene (A0U90_11950).The deletion mechanism is depicted in (A) displayed with the particular basic vector pKos6b. The deletion of the *polE* gene (A0U90_11950) is shown in (B). In (C) Agarose gel of colony PCR verifying the deletion of the *polE* gene in the genome; lane 1 showing *K. baliensis* NBRC 16680 R with the *polE* gene (3000 bp), lane 2 showing *K. baliensis* NBRC 16680 R with a *polE* deletion (Δ*polE*), with a PCR product of 1950 bp. (PPTX 418 kb)

Abbreviations

AA: Acetic acid; AAB: Acetic acid bacteria; *acs*: Cellulose synthesizing operon; *bcs*: Cellulose synthesizing operon; CPS: Capsular polysaccharide; DSM: Deutsche Sammlung von Mikroorganismen; EPS: Exopolysaccharide; EtOH: Ethanol; *G.*: *Gluconobacter*; *Ga.*: *Gluconacetobacter*; Gal: Galactose; Glc: Glucose; *gum*: Protein involved in HePS biosynthesis (derived from *X. campestris*); HEPES: 2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure; HePS: Heteropolysaccharide; HoPS: Homopolysaccharide; HPLC: High-performance liquid chromatography; *K*: *Kozakia*; KM: Kanamycin; *Ko*: *Komagataeibacter*; Man: Mannose; MWCO: Molecular weight cut-off; NaG: Modified sodium-gluconate medium; NBRC: Biological Research Centre NITE, Japan; OD: Optical density; *pol*: Genes involved in pellicle formation; R: Rough strain; *rfb*: Genes involved in dTDP-rhamnose synthesis; *X.*: *Xanthomonas*

Acknowledgements

We want to thank Prof. Dr. Wolfgang Liebl, Dr. Armin Ehrenreich and Dr. David Kostner (Department of Microbiology, Technical University of Munich) for giving us the opportunity to construct a Δ*polE* deletion mutant in *K. baliensis* NBRC 16680 R with their developed gene deletion system.

Funding

This work was supported by the German Research Foundation (DFG) and the Technische Universität München within the funding program Open Access Publishing. Part of this work was funded by the German Federal Ministry for Economic Affairs and Energy via the German Federation of Industrial Research Associations (AiF) (FEI); project number AiF 18071 N. The funder had no role in the design, analysis, interpretation or writing of the manuscript.

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its Additional files.

Authors' contributions

JUB, FJ and RFV were involved in planning the experimental setup and writing the manuscript. JUB performed the main experimental work and FLB was involved in the implementation of the *polE* knockout system. JUB evaluated the data and wrote the main text of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Received: 20 April 2017 Accepted: 10 July 2017
Published online: 10 August 2017

References

1. Roberts IS. The biochemistry and genetics of capsular polysaccharide production in bacteria. Annu Rev Microbiol. 1996;50:285–315.
2. Ferreira AS, Silva IN, Oliveira VH, Cunha R, Moreira LM. Insights into the role of extracellular polysaccharides in *Burkholderia* adaptation to different environments. Front Cell Infect Microbiol. 2011;16(1):1–9.
3. Jakob F. Novel fructans from acetic acid bacteria. Doctoral Dissertation, Freising, Germany: Technische Universität München; 2014.
4. Jakob F, Meißner D, Vogel RF. Comparison of novel GH 68 levansucrases of levan-overproducing *Gluconobacter species*. Acetic Acid Bacteria. 2012;1(1):2.
5. Jakob F, Pfaff A, Novoa-Carballal R, Rübsam H, Becker T, Vogel RF. Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function. Carbohydr Polym. 2013;92(2):1234–42.
6. Jansson P, Lindberg J, Wimalasiri KMS, Dankert MA. Structural studies of acetan, an exopolysaccharide elaborated by *Acetobacter xylinum*. Carbohydr Res. 1993;245:303–10.
7. Kornmann H, Duboc P, Marison I, Stockar U. Influence of nutritional factors on the nature , yield , and composition of exopolysaccharides produced by *Gluconacetobacter xylinus* I-2281. Appl Environ Microbiol. 2003;69(10):6091–8.
8. Brown RM, Willison JH, Richardson CL. Cellulose biosynthesis in *Acetobacter xylinum*: visualization of the site of synthesis and direct measurement of the in vivo process. PNAS. 1976;73(12):4565–9.
9. Yamada Y, Pattaraporn Y, Vu L, Thi H, Yuki M, Duangjai O, Somboon T, Yasuyoshi N. Description of *Komagataeibacter gen. nov.*, with proposals of new combinations (Acetobacteraceae). J Gen Appl Microbiol. 2012;58(5):397–404.
10. Moonmangmee S, Kawabata K, Tanaka S, Toyama H, Adachi O, Matsushita K. A novel polysaccharide involved in the pellicle formation of *Acetobacter aceti*. J Biosci Bioeng. 2002;93(2):192–200.
11. Ali IAI, Akakabe Y, Moonmangmee S, Deeraksa A, Matsutani M, Yakushi T, Yamada M, Matsushita K. Structural characterization of pellicle polysaccharides of *Acetobacter tropicalis* SKU1100 wild type and mutant strains. Carbohydr Polym. 2011;86(2):1000–6.

12. Moonmangmee S, Toyama H, Adachi O, Teerakool G, Lotong N, Matsushita K. Purification and characterization of a novel polysaccharide involved in the pellicle produced by a thermotolerant *Acetobacter* strain. Biosci Biotechnol Biochem. 2001;66(4):777–83.
13. Saxena IM, Kudlicka K, Okuda K, Brown RM. Characterization of genes in the cellulose-synthesizing operon (*acs* operon) of *Acetobacter xylinum*: implications for cellulose crystallization. J Bacteriol. 1994;176(18):5735–52.
14. Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, Benziman M, Gelfand DH, Meade JH. Emerick aW: genetic organization of the cellulose synthase operon in *Acetobacter xylinum*. Proc Natl Acad Sci U S A. 1990;87: 8130–4.
15. Azuma Y, Hosoyama A, Matsutani M, Furuya N, Horikawa H, Harada T, Hirakawa H, Kuhara S, Matsushita K, Fujita N, et al. Whole-genome analyses reveal genetic instability of *Acetobacter pasteurianus*. Nucleic Acids Res. 2009;37(17):5768–83.
16. Shimwell JLC, J. G. Mutant frequency in *Acetobacter*. Nature. 1964;201:1051–2.
17. Matsushita K, Ebisuya H, Ameyama M, Adachi O. Change of the terminal oxidase from cytochrome a1 in shaking cultures to cytochrome o in static cultures of *Acetobacter aceti*. J Bacteriol. 1992;174(1):122–9.
18. Perumpuli PABN, Watanabe T, Toyama H. Pellicle of thermotolerant *Acetobacter pasteurianus* strains: characterization of the polysaccharides and of the induction patterns. J Biosci Bioeng. 2014;118(2):134–8.
19. Kanchanarach W, Theeragool G, Inoue T, Yakushi T, Adachi O, Matsushita K. Acetic acid fermentation of *Acetobacter pasteurianus:* relationship between acetic acid resistance and pellicle polysaccharide formation. Biosci Biotechnol Biochem. 2014;74(8):1591–7.
20. Brandt JU, Jakob F, Behr J, Geissler AJ, Vogel RF. Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*. Microb Cell Factories. 2016;15(1):170.
21. Wach A. PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S. cerevisiae*. Yeast. 1996;12(3):259–65.
22. Peters B, Junker A, Brauer K, Mühlthaler B, Kostner D, Mientus M, Liebl W, Ehrenreich A. Deletion of pyruvate decarboxylase by a new method for efficient markerless gene deletions in *Gluconobacter oxydans*. Appl Microbiol Biotechnol. 2013;97(6):2521–30.
23. Tillett D, Neilan B. Enzyme-free cloning: a rapidmethod to clone PCR products independent of vector restriction enzyme sites. Oxford university press. 1999;27(19):e26.
24. Kostner D, Peters B, Mientus M, Liebl W, Ehrenreich A. Importance of codB for new codA-based markerless gene deletion in *Gluconobacter* strains. Appl Microbiol Biotechnol. 2013;97(18):8341–9.
25. Hall PE, Anderson M, Johnston DM, Cannons RE. Transformation of *Acetobacter xylinus* with plasmid DNA by electroporation. Plasmid. 1992;28(3):194–200.
26. Creaven M, Fitzgerald RJ, O'Gara F. Transformation of *Gluconobacter oxydans subsp. suboxydans* by electroporation. Can J Microbiol. 1994;40(6):491–4.
27. Krajewski V, Simić P, Mouncey NJ, Bringer S, Sahm H, Bott M. Metabolic engineering of *Gluconobacter oxydans* for improved growth rate and growth yield on glucose by elimination of gluconate formation. Appl Environ Microbiol. 2010;76(13):4369–76.
28. Mostafa HE, Heller KJ, Geis A. Cloning of *Escherichia coli lacZ* and *lacY* genes and their expression in *Gluconobacter oxydans* and *Acetobacter liquefaciens*. Appl Environ Microbiol. 2002;68(5):2619–23.
29. Kaditzky SV, Rudi F. Optimization of exopolysaccharide yields in sourdoughs fermented by lactobacilli. Eur Food Res Technol. 2008;228:291–9.
30. Cavalcante A. V., Dobereiner, J.: a new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. Plant Soil. 1988;108(1):23–31.
31. Bertalan M, Albano R, de Padua V, Rouws L, Rojas C, Hemerly A, Teixeira K, Schwab S, Araujo J, Oliveira A, et al. Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. BMC Genomics. 2009;10:450.
32. Meneses CHSG. Rouws, L. F. M., Simões-araújo, J. L., Vidal, M. S., Baldani, J. I.: exopolysaccharide production is required for biofilm formation and plant colonization by the nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus*. Mol Plant-Microbe Interact. 2011;24(12):1448–58.
33. Griffin AM. Morris, V. J., Gasson, M. J.: genetic analysis of the acetan biosynthetic pathway in *Acetobacter xylinum*. Int J Biol Macromol. 1994;16(6):287–9.
34. Coucheron DH. An *Acetobacter xylinum* insertion sequence element associated with inactivation of cellulose production. J Bacteriol. 1991; 173(18):5723–31.
35. Deeraksa A, Moonmangmee S, Toyama H, Yamada M, Adachi O, Matsushita K. Characterization and spontaneous mutation of a novel gene, *polE*, involved in pellicle formation in *Acetobacter tropicalis* SKU1100. Microbiology. 2005;151(Pt 12):4111–20.
36. Marolda CLVM. Genetic analysis of the dTDP-rhamnose biosynthesis region of the *Escherichia coli* VW187 (O7:K1) rfb gene cluster: identification of functional homologs of rfbB and rfbA in the rff cluster and correct location of the rffE gene. J Bacteriol. 1995;177(19):5539–46.
37. Deeraksa A, Moonmangmee S, Toyama H, Adachi O, Matsushita K. Conversion of capsular polysaccharide, involved in pellicle formation, to extracellular polysaccharide by galE deletion in *Acetobacter tropicalis*. Biosci Biotechnol Biochem. 2006;70(10):2536–9.
38. Kanchanarach W, Theeragool G, Inoue T, Yakushi T, Adachi O, Matsushita K. Acetic acid fermentation of *Acetobacter pasteurianus* : relationship between acetic acid resistance and pellicle polysaccharide formation. Biosci Biotechnol Biochem. 2010;74(8):1591–7.
39. Lisdiyanti P, Kawasaki H, Widyastuti Y, Saono S, Seki T, Yamada Y, Uchimura T. *Kozakia baliensis* gen. Nov., sp. nov., a novel acetic acid bacterium in the α-Proteobacteria. Int J Syst Evol Microbiol. 2002;52:813–8.
40. Ishida T, Sugano Y, Shoda M. Effects of acetan on production of bacterial cellulose by *Acetobacter xylinum*. Biosci Biotechnol Biochem. 2002;66(8):1677–81.
41. Brown AJ. On an acetic ferment which forms cellulose. J Chem Soc. 1886; 49:432–9.
42. Toda K, Asakura T, Fukaya M, Entani E, Kawamura Y. Cellulose production by acetic acid-resistant *Acetobacter xylinum*. J Ferment Bioeng. 1997;84(3):228–31.

PART IV PAGES 78 - 80	Multiple genome Sequences of heteropolysaccharide-forming acetic acid bacteria

Since hitherto little is known about HePS clusters in AAB, whole genome sequencing of potentially HePS-forming AAB should reveal more information. We report here the complete genome sequences of *A*. *aceti* TMW 2.1153, *A*. *persici* TMW 2.1084 and *N*. *chiangmaiensis* NBRC 101099, acetic acid bacteria partly isolated from water kefir. Our objective was to use the whole genome sequencing data to make a cross-strain representation of existing HePS clusters in AAB, to get better insights in their biosynthesis pathways and the possible physiological relevancy of the particular HePS. The genomes were assembled in one (*A*. *aceti* TMW 2.1153, *N*. *chiangmaiensis* NBRC 101099) to two contigs in case of *A*. *persici* TMW 2.1084, with an additional plasmid of 526,169 bp. HePS forming gene clusters could be detected, in which each strain carries a whole *pol* gene clusters, *polABCDE*, associated with pellicle formation. Furthermore, in *A*. *aceti* TMW 2.1153, *gum*-like genes, similar to genes in the *gum* cluster of *Xanthomonas (X.) campestris*, could be identified. Their accessibility will allow a better investigation of AAB derived HePS, in context of the AAB specific HePS formation and usability, depending on the available habitat.

Authors contributions: Julia U. Brandt was responsible for the independent planning and conducting the experiments. Julia U. Brandt evaluated the data and wrote the main text of the manuscript. Frank Jakob was the project leader responsible for supervision. Rudi Vogel conducted the conceptual setup and supervision of the study including infrastructural demands and final editing of the manuscript. Andi Geissler and Jürgen Behr were involved in the analysis of the whole genome sequenci

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Multiple Genome Sequences of Heteropolysaccharide-Forming Acetic Acid Bacteria

Julia U. Brandt,[a] Frank Jakob,[a] Andreas J. Geissler,[a] Jürgen Behr,[b] Rudi F. Vogel[a]

Technische Universität München, Lehrstuhl für Technische Mikrobiologie, Freising, Germany[a]; Technische Universität München, Bavarian Biomolecular Mass Spectrometry Center, Freising, Germany[b]

ABSTRACT We report here the complete genome sequences of the acetic acid bacteria (AAB) *Acetobacter aceti* TMW 2.1153, *A. persici* TMW 2.1084, and *Neoasaia chiangmaiensis* NBRC 101099, which secrete biotechnologically relevant heteropolysaccharides (HePSs) into their environments. Upon genome sequencing of these AAB strains, the corresponding HePS biosynthesis pathways were identified.

Received 16 February 2017 **Accepted** 2 March 2017 **Published** 20 April 2017

Citation Brandt JU, Jakob F, Geissler AJ, Behr J, Vogel RF. 2017. Multiple genome sequences of heteropolysaccharide-forming acetic acid bacteria. Genome Announc 5:e00185-17. https://doi.org/10.1128/genomeA.00185-17.

Address correspondence to Frank Jakob, frank.jakob@wzw.tum.de.

Acetic acid bacteria (AAB) are Gram-negative obligate aerobes belonging to the *Alphaproteobacteria* subdivision. They are well known to produce large amounts of exopolysaccharides (EPS)—either homopolysaccharides (HoPSs) such as levans (1–3) and cellulose (4, 5), or heteropolysaccharides (HePSs) such as acetan (6) and gluconacetan (7). HePSs, in particular, have unique properties, since their complex, mostly branched structures are responsible for drastic viscosity increases of aqueous solutions. The food industry is taking increasing advantage of the unique rheological properties of bacterial HePSs. We identified three potential HePS-forming AAB, *Acetobacter aceti* TMW 2.1153, *A. persici* TMW 2.1084, and *Neoasaia chiangmaiensis* NBRC 101099, on sucrose-deficient media. To gain insights into the HePS biosynthesis we sequenced the complete genomes of the identified HePS-producing AAB for further identification of specific HePS biosynthesis clusters.

A. aceti TMW 2.1153 and *A. persici* TMW 2.1084 were isolated from water kefir in our laboratory in Freising, Germany, and *N. chiangmaiensis* NBRC 101099 was isolated from a Thai red ginger flower at the National Institute of Technology and Evaluation Biological Resource Center, Japan (8). High-molecular-weight DNA was purified from modified sodium-gluconate (NaG) medium liquid cultures using the Genomic-tip 100/G kit (Qiagen, Hilden, Germany), as described previously (9). Single-molecule real-time (SMRT) sequencing (PacBio RS II) was carried out at GATC (Konstanz, Germany) (10). A single library was prepared for each of the three strains, and an insert size of 8 to 12 kb was selected for library creation, resulting in at least 200 Mb of raw data from 1 SMRT cell (1 × 120-min movies), applying P4-C2 chemistry. The generated sequences were assembled with SMRT Analysis version 2.2.0.p2 using the Hierarchical Genome Assembly Process version 3 (HGAP3) (11). Initial open reading frame predictions and annotations were accomplished automatically using the NCBI Prokaryotic Genome Annotation Pipeline and Rapid Annotations using Subsystems Technology (RAST), a SEED-based, prokaryotic genome annotation service (12–14).

The genomes were assembled to one circularized chromosome with overall chromosome sizes ranging from 3.23 Mb for *A. persici* TMW 2.1084 to 3.72 Mb for *A. aceti* TMW 2.1153 and G+C contents of 56.83 to 61.50%. *A. persici* TMW 2.1084 harbors an additional plasmid comprising 526,169 bp. The detailed characteristic data, sequencing statistics, genome information, and accession numbers are given in Table 1.

TABLE 1 Strain characteristics, sequencing statistics, genome information, and accession numbers

Strain	Source	BioSample no.[a]	Accession no.	Size (Mb)	Coverage (×)[b]	No. of contigs[c]	G+C content (%)	No. of PEGs[d]	No. of CDSs[e]
A. persici TMW 2.1084	Water kefir (Freising, Germany)	SAMN 04396916	CP014687 to CP014688	3.23	172	2	57.44	3,387	3,044
A. aceti TMW 2.1153	Water kefir (Freising, Germany)	SAMN 04396917	CP014692	3.72	96	1	56.83	3,664	3,200
N. chiangmaiensis NBRC 101099	*Alpinia purpurata* (Chiang-Mai, Thailand)	SAMN 04396918	CP014691	3.40	59	1	61.50	3,296	3,002

[a]All biosamples are part of the BioProject PRJNA311264. Accession numbers are listed for all contigs of each whole genome (as a range).
[b]Average coverage of HGAP assembly.
[c]Chromosome plus plasmids and partial plasmids (only the case for TMW 2.1084).
[d]PEGs, protein-encoding genes (based on RAST annotation).
[e]CDSs, coding sequences (coding based on NCBI Prokaryotic Genome Annotation Pipeline).

The genome of *A. aceti* TMW 2.1153 harbors three complete rRNA operons (5S, 16S, and 23S) and 52 tRNA genes. *A. persici* TMW 2.1084 encodes 62 tRNAs, and *N. chiangmaiensis* NBRC 101099 encodes 54 tRNAs; both harbor four complete rRNA operons. Among the identified genes, HePS-forming gene clusters could be detected, including *pol* genes of the *polABCDE* cluster associated with pellicle formation (15). Furthermore, in *A. aceti* TMW 2.1153, *gum*-like genes, similar to genes in the *gum* cluster of *Xanthomonas campestris* (16, 17), could be identified. Their accessibility will allow a better investigation of AAB-derived HePSs and the connected biosynthesis, based on specific HePS clusters.

Accession number(s). The three complete genomes have been deposited in DDBJ/EMBL/GenBank under the accession numbers given in Table 1.

ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG) and the Technical University of Munich (TUM) in the framework of the Open Access Publishing Program and the German Federal Ministry for Economic Affairs and Energy via the German Federation of Industrial Research Associations (AiF) (FEI), project number AiF 18071 N.

None of the funding sources had any influence on the study design; the collection, analysis, and interpretation of the data; the writing of the report; or the decision to submit the article for publication.

REFERENCES

1. Jakob F. 2014. Novel fructans from acetic acid bacteria. PhD dissertation. Technische Universität München, Freising, Germany.
2. Jakob F, Meißner D, Vogel RF. 2012. Comparison of novel GH 68 levansucrases of levan-overproducing Gluconobacter species. Acetic Acid Bacteria 1:2. https://doi.org/10.4081/aab.2012.e2.
3. Jakob F, Pfaff A, Novoa-Carballal R, Rübsam H, Becker T, Vogel RF. 2013. Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function. Carbohydr Polym 92:1234–1242. https://doi.org/10.1016/j.carbpol.2012.10.054.
4. Chawla PR, Bajaj IB, Survase SA, Singhal RS. 2009. Microbial cellulose: fermentative production and applications. Food Technol Biotechnol 47:107–124.
5. Raspor P, Goranovic D. 2008. Biotechnological applications of acetic acid bacteria. Crit Rev Biotechnol 28:101–124. https://doi.org/10.1080/07388550802046749.
6. Jansson PE, Lindberg J, Wimalasiri KMS, Dankert MA. 1993. Structural studies of acetan, an exopolysaccharide elaborated by *Acetobacter xylinum*. Carbohydr Res 245:303–310. https://doi.org/10.1016/0008-6215(93)80079-T.
7. Kornmann H, Duboc P, Marison I, von Stockar UV. 2003. Influence of nutritional factors on the nature, yield, and composition of exopolysaccharides produced by *Gluconacetobacter xylinus* I-2281. Appl Environ Microbiol 69:6091–6098. https://doi.org/10.1128/AEM.69.10.6091-6098.2003.
8. Yukphan P, Malimas T, Potacharoen W, Tanasupawat S, Tanticharoen M, Yamada Y. 2005. *Neoasaia chiangmaiensis* gen. nov., sp. nov., a novel osmotolerant acetic acid bacterium in the α-proteobacteria. J Gen Appl Microbiol 51:301–311. https://doi.org/10.2323/jgam.51.301.
9. Brandt JU, Jakob F, Behr J, Geissler AJ, Vogel RF. 2016. Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*. Microb Cell Fact 15:170. https://doi.org/10.1186/s12934-016-0572-x.
10. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D. 2009. Real-time DNA sequencing from single polymerase molecules. Science 323:133–138. https://doi.org/10.1126/science.1162986.
11. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods 10:563–569. https://doi.org/10.1038/nmeth.2474.
12. 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:D206–D214. https://doi.org/10.1093/nar/gkt1226.
13. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures

(SOPs) for (meta)genomic annotation. OMICS 12:137–141. https://doi.org/10.1089/omi.2008.0017.
14. 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. https://doi.org/10.1186/1471-2164-9-75.
15. Deeraksa A, Moonmangmee S, Toyama H, Yamada M, Adachi O, Matsushita K. 2005. Characterization and spontaneous mutation of a novel gene, polE, involved in pellicle formation in *Acetobacter tropicalis* SKU1100. Microbiology 151:4111–4120. https://doi.org/10.1099/mic.0.28350-0.
16. Becker A, Katzen F, Pühler A, Ielpi L. 1998. Xanthan gum biosynthesis and application: a biochemical/genetic perspective. Appl Microbiol Biotechnol 50:145–152. https://doi.org/10.1007/s002530051269.
17. Katzen F, Ferreiro DU, Oddo CG, Ielmini MV, Becker A, Pühler A, Ielpi L. 1998. *Xanthomonas campestris* pv. campestris gum mutants: effects on xanthan biosynthesis and plant virulence. J Bacteriol 180:1607–1617.

4. DISCUSSION

Kozakia (K.) baliensis is known to produce high amounts of high molecular weight levan with sucrose as the main carbon source (Jakob 2012c; Jakob 2013). Besides the well-studied levan, *K. baliensis* secretes a so far not characterized HePS into the surrounding medium. The present work highlights the potential of *K. baliensis* DSM 14400 and NBRC 16680 to produce a new, unique HePS and further provides a fundamental contribution to the understanding of the particular HePS biosynthesis as well as hinds for its physiological assignment. Furthermore, a global comparison including newly generated and existing genome sequencing data was carried out, resulting in a comprehensive overview of HePS clusters in AAB. In this section, the results of the four underlying publications will be discussed and auxiliary complemented by additional findings.

Throughout the cultivation of 27 AAB strains on sucrose-deficient NaG plates, we found that at least five strains produce a mucous substance from mannitol and glycerol as (main) carbon sources (supplementary file 1): *A. aceti* TMW 2.1153, *A. persici* TMW 2.1084, *N. chiangmaiensis* NBRC 101099, *K. baliensis* DSM 14400 and *K. baliensis* NBRC 16680.

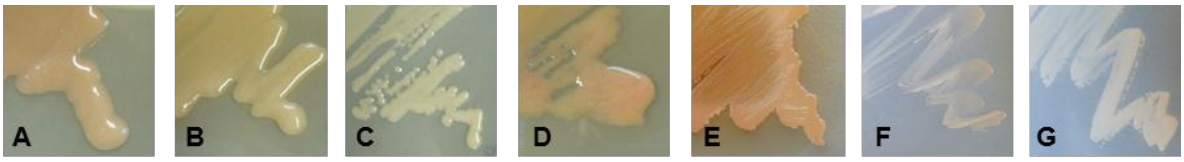


Figure 1 Colony morphology of different AAB strains on NaG plates without sucrose. (A) *K. baliensis* DSM 14400, (B) *K. baliensis* NBRC 16680, (C) *A. persici* TMW 2.1084, (D) *N. chiangmaiensis* NBRC 101099, (E) *Ga. johannae* DSM 13595, (F) *Asaia bogorensis* IFO 16594 and (G) *G. oxydans* TMW 2.877.

Since only both *K. baliensis* strains showed significant HePS formation in medium, they were selected for further experiments. Preliminary experiments from Jakob (2012b) with parenthetically analyzed EPS of *K. baliensis* DSM 14400, isolated from sucrose deficient NaG medium, revealed that it consists of several sugar monomers. Via subsequent HPLC analysis of both *K. baliensis* strains it could shown, that both HePS consist of glucose, galactose and mannose, respectively (Figure 1, PART I). However, both *K. baliensis* HePS differ in their monomer distribution, with significant higher amounts of glucose in case of the DSM 14400 strain. To examine the genomic background of the formed HePS and further, to be able to clarify the observed differences between *K. baliensis* DSM

Table 5 Replicon info of *K. baliensis* DSM 14400 and NBRC 16680

Type	Name	RefSeq	Size (bp)	GC%	CDS coding	rRNA	tRNA	Other RNA	Genes coding	Pseudogene
***K. baliensis* DSM 14400**										
Chromosom	-	NZ_CP014674.1	2.888.029	57.4	2,591	15	57	4	2.742	75
Plasmid	pKB14400_1	NZ_CP014675.1	253.508	58.8	230	-	-	-	268	38
Plasmid	pKB14400_2	NZ_CP014676.1	102.455	59.3	118	-	-	-	121	3
Plasmid	pKB14400_3	NZ_CP014677.1	176.991	53.4	132	-	-	-	156	24
Plasmid	pKB14400_4	NZ_CP014678.1	18.625	59.9	19	-	-	-	23	4
Plasmid	pKB14400_5	NZ_CP014679.1	35.463	56.4	30	-	-	-	40	10
Plasmid	pKB14400_6	NZ_CP014680.1	30.533	56.5	24	-	-	-	38	14
***K. baliensis* NBRC 16680**										
Chromosom		NZ_CP014681.1	2.807.246	57.5	2,498	15	57	4	2,632	58
Plasmid	pKB16680_1	NZ_CP014682.1	135.727	60.1	141	-	-	-	153	12
Plasmid	pKB16680_2	NZ_CP014683.1	88.099	59.4	99	-	-	-	104	5
Plasmid	pKB16680_3	NZ_CP014684.1	36.022	57.2	29	-	-	-	36	7
Plasmid	pKB16680_4	NZ_CP014685.1	14.635	53.4	9	-	-	-	14	5
Plasmid	pKB16680_5	NZ_CP014686.1	11.848	56.3	10	-	-	-	14	4

14400 and NBRC 16680, both *K*. *baliensis* genomes were sequenced and circularized chromosomes could be established (Figure 2, PART I). The genomes of both *K*. *baliensis* strains differ from each other; beside a slightly larger chromosome of the DSM 14400 strain (2,888,029 bp) compared to the NBRC 16680 strain (2,807,246 bp), *K*. *baliensis* DSM 14400 also contains a higher number of plasmids (Table 5). In order to be able to make more accurate statements about the *Kozakia* EPSs, initially HoPS forming clusters were searched, whereas both *K*. *baliensis* strains (DSM 14400, NBRC 16680) exhibited the expected levansucrase genes (Figure 5a, PART I). However, also clear differences between both strains were observable, primarily concerning plasmid 3 (pKB14400_3) of the DSM 14400 strain. This plasmid harbors a selection of additional HoPS forming clusters [a cellulose synthase operon and a levansucrase flanked by a type II secretion operon with high similarity to a *Ga*. *diazotrophicus* levansucrase gene (GDI_RS02220)], which do not occur in the NBRC 16680 strain (Figure 5, PART I). Both clusters are framed by mobile elements. While the levansucrase gene is interrupted by a transposon insertion and accordingly inactive (Jakob 2012a), the cellulose synthase operon includes all necessary elements for cellulose formation (Wong et al. 1990)(Figure 2A). The transcriptional activity of this cluster was checked via RT-qPCRs. Since it is known, that magnesium leads to an increased activation of the cellulose synthase operon (Ross P 1987), the expression level of the diguanylate cyclase gene (*dgc*) and the gene coding for the cellulose synthase subunit c (*bcsC*) in NaG and NaG-fgm medium was measured after 24 hours. The relative expression of both tested genes (*dgc*, *bcsC*) indicated that the genes are transcribed and not silent under the applied growth conditions. Further, the real-time PCR showed that both genes *dgc* and *bcsC* were expressed at an elevated level when the cells were grown in NaG-fgm medium (Figure 2C).

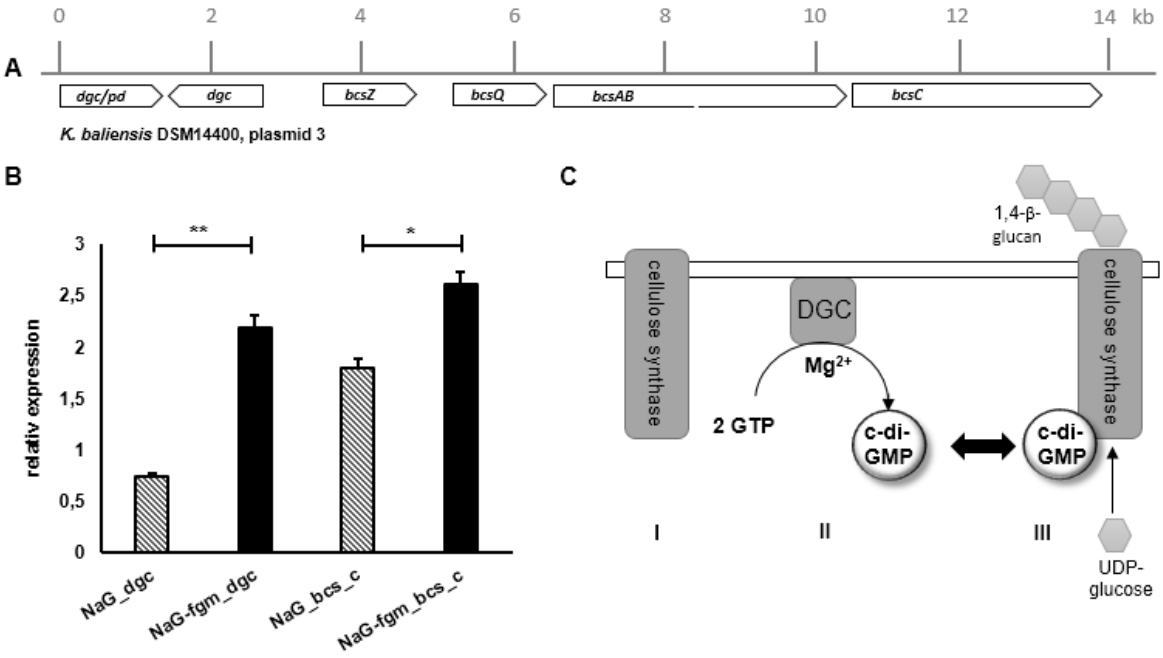


Figure 2 Growth and gene expression of the *K. baliensis* DSM 14400 *dgc* and *bcs-C* genes in NaG and NaG-fgm medium. In (A) the genetic organization of the of the *K. baliensis* DSM 14400 cellulose synthase operon is illustrated. The operon consists of six genes: *dgc/pd*: diguanylate cyclase/phosphodiesterase (GGDEF & EAL domains) with PAS/PAC sensor(s), *dgc*: diguanylate cyclase, *bcsZ*: Endoglucanase precursor (EC 3.2.1.4), *bcsQ*: NTPase, *bcsAB*: Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12), *bcsC*: Cellulose synthase operon protein C. (B) shows the relative expression of the *dgc*(striped bar) and *bcsC* gene (black bar) in NaG and NaG-fgm medium after 24 h. (C) Model of the Mg^{2+} dependent activation of the cellulose synthase. Inactivated cellulose synthase (I).The diguanylate cyclase (DGC) converts GTP into c-di-GMP (II) C-di-GMP acts as an allosteric, positive effector and activates the cellulose synthase (III). Schema adapted from Ross P (1987).

In several studies, the enzyme activating effect of magnesium has been discussed, especially related to the second messenger bis-(3´-5´)-cyclic-GMP (c-di-GMP). It was shown, that magnesium enhances the activity of the diguanylate cyclase that leads to the formation of c-di-GMP, which in turn acts as an allosteric, positive effector of the cellulose synthase activity (O'Shea et al. 2005; O'Shea 2006; Ross 1990; Ross P 1987; Whitney and Howell 2013). This stimulating effect can, most likely, be traced back to an induced conformational change of the cellulose synthase subunit BcsA that allows UDP-glucose to access the catalytic site more effectively (Morgan 2014). From the transcriptional activity of the cellulose operon of *K. baliensis* DSM 14400, it is assumed that this strain forms additional cellulose. The activity of this cluster was also checked on the EPS forming level via congored assays. However, no clear evaluation was possible due to conceivable interferences between the probably formed cellulose and the simultaneously produced HePS. The additionally formed cellulose of *K. baliensis* DSM 14400 would be an explanation of the mentioned, increased amount of glucose after the hydrolysis of the isolated *K. baliensis* DSM 14400 EPSs, preceding the HPLC analysis.

To figure out the exact genomic origin of the *K*. *baliensis* HePS (DSM 14400, NBRC 16680), before carrying out the complex and time-consuming structural analysis, both *K*. *baliensis* HePS were precisely analyzed at the genetic level. Via comparison of the genomes of both *K*. *baliensis* (DSM 14400, NBRC 16680) strains with related AAB and comprehension of known literature about HePS biosynthesis in AAB and the xanthan producer *X*. *campestris*, *gum*-like clusters could be identified.(Figure 3,PART I). These clusters showed high similarity to other HePS producing AAB, like *Ga*. *diazotrophicus* PA15 and *Ko*. *xylinus* E25 (Table 1, PART I). The *gum*-like HePS cluster of *K*. *baliensis* (DSM 14400, NBRC 16680) includes several *gum*-like genes, *gumB*, *gumC*, *gumD*, *gumE*, *gumH*, *gumJ*, *gumK* & *gumM,* which can be traced back to the well-characterized xanthan cluster (Becker et al. 1998; Pühler et al. 2008)). Further, by comparison of both *K*. *baliensis gum*-like clusters with the well characterized *gum*-cluster of *X*. *campestris*, genes could be identified and specific functions could be assigned to the particular genes. These genes (*gumD*, *gumH* and *gumM*) conceivably code for enzymes that are necessary for the incorporation of the HPLC analyzed neutral monomers glucose and mannose into the *K*. *baliensis* HePS (Pühler et al. 2008). In addition to the identified glycosyltransferase genes, the *K*. *baliensis gum*-like clusters harbor genes, which are necessary for the HePS polymerization and export (e.g. *gumB, gumC, gumE*). These genes arealso present in the acetan and *Ga*. *diazotrophicus* HePS cluster (Ishida 2002b). Moreover, a homology comparison between the glycosyltransferase genes *aceP* and *aceQ* of the acetan cluster of *Ko*. *xylinu*s E25 (Ishida 2002b), with the glycosyltransferase 1 and 2 genes of both *K*. *baliensis* strains, identified both as additional glucosyltransferases in the respective *gum*-like cluster (Table 1, PART I). Thus, it could be shown that by application of comparative genomics, possible functions of many genes in the cluster can already be assigned in advance. However, it has to be considered, that the HePS from *X*. *campestris* and *Ko*. *xylinus* do not contain galactose (Couso 1987; Ielpi 1993), hence, no function could be assigned via direct comparison (Table 2, PART I). *Ga*. *diazotrophicus* harbors the most related HePS biosynthesis cluster compared to *K*. *baliensis*. However, although the *Ga*. *diazotrophicus* HePS includes galactose, no galactosyltransferase is described in the cluster yet. Therefore it is assumed, that one or more of the homologous glycosyltransferases, solely present in the *K*. *baliensis* and *Ga*. *diazotrophicus gum*-like cluster, could function as specific galactosyltransferases. Since the individual AAB HePS differ from one another per se, also variances between the respective clusters could be determined. In contrast to the *K*. *baliensis gum*-like cluster, a *gumJ* gene, necessary for the translocation of the HePS across the membrane (Schmid 2015), is missing in the acetan cluster of *Ko*. *xylinus* E25

(Table 2, PART I). However, this might be functionally replaced by the *aceC* gene, which is only present in the acetan cluster. The AceC protein contains a MATE Wzx-like domain that could act as a flippase, supporting the membrane translocation of acetan. Furthermore, the acetan cluster is flanked by a mannose-1-phosphate guanylyltransferase gene (*manB)*, which is in both *K*. *balien*sis genomes located separately from the HePS cluster. The *Ko*. *xylinus* derived acetan consists of D -glucose, D -rhamnose, D -mannose and D-glucuronic acid, accordingly only in this AAB cluster a rhamnosyl- transferase gene (*aceR*) is involved (Griffin 1994). Furthermore, the acetan cluster, as well as the HePS cluster of *Ga*. *diazotrophicus* include the putative acetyltransferase gene *gumF*, coding for a protein that incorporates acetyl residues into the particular HePS, which is not present in the respective *K*. *bali*ensis cluster (Figure 3, PART I). This is in agreement with the preliminary HPLC analysis of both *K*. *baliensis* HePS, which contain no rhamnose, while the lack of acetyl-groups in both *Kozakia* HePS has to be further investigated in detail. The additional detection of a *gumK* gene in both *K*. *baliensis* clusters suggests the existence of an additional glucuronic acid (GlcA) residue in the HePS, which also could not be detected via HPLC sugar monomer analysis.

In order to get a better insight into the *K*. *baliensis* HePS structure and to find potential similarities to acetan, a structural analysis of the *K*. *baliensis* NBRC 16680 HePS was carried out by using HPAEC-PAD, methylation analysis, and NMR. Because the detected cellulose synthase operon that indicates a potential cellulose formation of *K*. *baliensis* DSM 14400, causing possible ambiguities, excludes this strain for the structural analysis. Here, it becomes clear how important, regarding time and costs, a prior review of the genomic data is.

Via the Methylation analysis of the *K*. *baliensis* NBRC 16680 HePS, 1,4- and 1,3,4-substituted glucose units, as well as 1,2-substituted mannose and 1,4-substituted glucuronic acid units could be detected (Figure2B, PART II). This indicates the presence of an acetan-like core structure, with a β-(1→4)-linked glucopyranosyl backbone, substituted with a tri- or pentasaccharide side chain at position *O*3 of every second glucose unit. This hypothesis is supported by the roughly equal ratio of mannose and glucuronic acids, and the simultaneously available, comparable amounts of 1,2-substituted mannose, 1,4-substituted glucose, and 1,3,4-substituted glucose units detected by methylation analysis. In contrast to acetan, a putative unidentified uronic acid and 1,6-substituted galactose units were found to be *K*. *baliensis* NBRC 16680 HePS constituents (Figure2A, PART II). This observation is, in case of the galactose, in consistency with the previous HPLC analyses. However, with the obtained data it was not possible to elucidate the exact structure of the *K*. *baliensis* formed HePS. However, by including

the genomic data and comparison of the particular genes in the HePS cluster of *K*. *baliensis* NBRC 16680 with the corresponding genes in the HePS clusters of *X*. *campestris* and *Ko*. *xylinus* and by incorporating the knowledge of the detected linkage types of *K*. *baliensis* NBRC 16680 HePS, a hypothetical structure could be constructed (Figure 4, PART II). In accordance with the β-1,4-linked glucose backbone of acetan and xanthan, the *K*. *baliensis* NBRC 16680 HePS contains 1,4-substituted glucopyranosyl residues (Jansson 1975; Jansson et al. 1993). The certain genes, *gumD* and *gumM*, were already identified in the *gum*-like cluster of *K*. *baliensis* NBRC 16680/DSM 14400, potentially coding for corresponding glycosyltransferases, henceforth associated with the incorporation of glucose units via a particular linkage type. The *gumD* gene codes for a protein that incorporates the initial glucose-1-phosphate residue and *gumM* codes for a protein that transfers a glucosyl residue from UDP-glucose to the previous glucose residue, resulting in a β-1,4-linked glucose-disaccharide (Bertalan et al. 2009; Brandt 2016; Ielmini 1998). Further, the presence of the *gumH* gene in the *gum*-like cluster of *K*. *baliensis* NBRC 16680 coincides with the identified 1,3,4-substituted glucose and 1,2-substituted mannose units, so that it could be assumed via the know function of GumH, that the second glucose residue is ramified with a mannose unit at position *O*3 (Abdian 2000; Geremia 1999; Ielmini 1998). The subsequent step in the HePS synthesis of *Ko*. *xylinus* and *X*. *campestris* is determined by the *gumK* gene, which codes for a protein that adds a glucuronosyl residue from UDP-glucuronate to the glucose–glucose–mannose trisaccharide (Barreras 2004; Petroni and Ielpi 1996). The *gumK* gene is also part of the *gum*-like cluster of *K*. *balien*sis NBRC 16680 and consistently 1,4-substituted glucuronic acid was detected during the methylation analysis. In addition to the 1,4-substituted glucuronic acid unit, a further putative uronic acid was detected via HPAEC-PAD chromatograms. This is so far unique and does not occur in the other described HePS. We assume that the so far unknown uronic acid is caused by a modification of the glucuronic acid, which might occur before or after the incorporation of glucuronic acid into the repeating unit. After the incorporation of glucuronic acid into the HePS, the biosynthesis deviates from the xanthan biosynthesis, based on the cluster sequence. In the xanthan biosynthesis, the following step is determined by the *gumI* gene, which codes for the ultimate glycosyltransferase that catalyzes the transfer of a mannosyl-residue from GDP-mannose to the xanthan tetrasaccharide. This gene is not present in the acetan and the *K*. *baliensis* NBRC 16680 HePS cluster (Figure 4, PART II), most likely resulting in different structures of xanthan, compared to acetan and the *K*. *baliensis* NBRC 16680 HePS. The side chains of acetan are in objective elongated by two 1,6-substituted glucose units, via AceP and

AceQ, and finally terminated by a rhamnose unit incorporated by AceR (Petroni and Ielpi 1996). It is conceivable that the growing *K*. *baliensis* tetrasaccharide (glucose–glucose–mannose-glucuronic acid) is also further substituted by one or more 1,6-substituted glucose residues, since the *gum*-like cluster harbors similar GT genes (*aceP* and *aceQ*) as *Ko*. *xylinus* E25, encoding proteins that catalyze the incorporation of the two 1,6-linked glucose-1-phosphates (Ishida 2002b). This is consistent with the detected relatively high amounts of 1,6-substituted and terminal glucose units. Equal as previously mentioned via the HPLC analysis, no rhamnose could be identified during the monosaccharide and methylation analysis, which is confirmed by the absence of the rhamnosyl-transferase gene (*aceR*) in the *K*. *baliensis* NBRC 16680 HePS cluster. Alike, there was no acetylation of side chains detected, like in all other HePS clusters (acetan, xanthan and the *Ga*. *diazotrophicus* cluster), that contain the acetyl-transferases gene *gumF*, which does not occur in the *K*. *baliensis* cluster (Figure 4, PART II). Accordingly, nearly all linkage types that are expected at the genetic level, based on the HePS cluster alignment, were identified via methylation analysis. However, it remains the problem of an exact assignment of the galactose residues in the *K*. *baliensis* (DSM14400, NBRC 16680) HePS, since xanthan and acetan from *X*. *campestris* and *Ko*. *xylinus* do not contain galactose (Couso 1987). Due to the occurrence of 1,6-linked galactose units, we hypothesize that the galactose units are bound to or replace the 1,6-linked glucose residues. To get further insights, we attempt to identify potential galactosyltransferases of the *K*. *baliensis* cluster via BlastP alignments with known galactosytransferase genes. Thus, to this point unassigned GT genes (*gtA0U90_06735, gt A0U90_06740*) located upstream of the *gumJ* gene of the *gum*-like cluster propose to be associated with the incorporation of galactosyl-residues. Further, an accordance of these two possible galactosyltransferase genes, with a corresponding gene in *Ga*. *diazotrophicus* (PAI5; GDIA_RS03825), also forming a HePS composed of glucose, galactose and mannose (Serrato et al. 2013), could be found (*gtA0U90_06735;* Q 97%, identities 190/376 (51%) & *gt A0U90_06740* Q 99 %, identities 171/384 (45%)). One of the *K*. *baliensis* GTs, GT *gtA0U90_0673*, comprises a predicted GT2_AmsE_like conserved domain (Bit Score: 119.34,E-value: 2.33e-32), primary known to be part of the glycosyltransferases AmsE, which is proposed to attach galactose residue to the backbone of the chimeric EPS amylovoran of the plant pathogenic bacteria *Erwinia amylovora* (Langlotza 2011). In accordance with the methylation analysis of the *K*. *baliensis* NBRC 16680 HePS, where 1,6-substituted galactose units were found, also amylovoran is a complex HePS composed of glucose, galactose and glucuronic acid, with 1,6-linked galactose residues (Wang 2012). From the methylation

analysis of the *K. baliensis* NBRC 16680 HePS, it cannot be concluded if glucose or galactose is bound to the uronic acid residues or if the 1,6-linked glucose/galactose residues alternate or appear blockwise. However, the high portion of terminal glucose suggests that the side chains are terminated by glucose residues. The combination of the structural and genomic data strongly suggests that the *K. baliensis* NBRC 16680 HePS exhibits an acetan core structure, which contains 1,6-linked glucose/galactose at currently non assignable position(s) and no rhamnosyl/acetyl residues.

Overall, it could be shown, that the applied mutual method efficiently supported the chemical analyses, resulting in the final clarification of the complex HePS structure, which leads to faster results, particular evident for HePS structure elucidation. We could demonstrate, that via forgoing whole genome sequencing of a strain of interest, it is possible to predict important structural components of a polysaccharide derived from a connected EPS clusters.

Nevertheless, no full statements about the resulting EPS can be made exclusively on the basis of genomic date, not taken into account transcriptional and translational factors. This was especially evident when the rheological data of the *K. baliensis* NBRC 16680 HePS were considered. In general, the HePS solutions of *K. baliensis* NBRC 16680 exhibited shear-thinning behavior with low flow behavior indices (n) obtained from the power law model (Figure 3a, PART II). Solutions of *K. baliensis* NBRC 16680 HePS were highly viscous, even in low concentrations. However, via a comparison of the HePS derived from NaG and NaG-fgm media it turned out that HePS extracted from NaG media exhibited higher n-values than those isolated from NaG-fgm media. As already mentioned in context with the cellulose synthase, magnesium often plays an essential role in activation of enzymes that are involved in EPS biosynthesis, e.g. in the polymerization activity of glycosyltransferases (Ross 1990; Vargas-Garcıa 2001). To explain the exact difference between both, the NaG and NaG-fgm derived HePS from *K. baliensis* NBRC 16680 were further characterized via HPSEC and AF4-MALS analysis. The NaG-fgm derived HePS eluted later during AF4 separation and earlier during HPSEC separation (Figure 3, PART II), indicating a larger molecular weight of this HePS. Additionally, slightly higher (absolute determination from MALS signals independent of concentration signals) rms radii of NaG-fgm HePS were observed (Figure 3b, PART II). In contrast, there was no influence of magnesium on the HePS composition detectable during growth in Nag and NaG-fgm medium. We additionally analyzed the neutral sugar composition of *K. baliensis* NBRC 16680 HePS derived from media with different carbon sources (NaG, NaG-fgm, NaG-fg, NaGg) and magnesium, where it could be demonstrated that the HePS of

K. *baliensis* NBRC 16680 are composed of comparable amounts of glucose, galactose, and mannose, independent of the cultivation medium (Figure 1, PART II). The consistent HePS composition is in conformity with the postulated progression of the directed HePS biosynthesis in general (Schmid 2015). HePS are formed via the sequential addition of nucleoside diphosphate sugars via glycosyltransferases (GT) to growing repeating units of five to seven monomers, as shown for the xantha (Ielpi 1993) and acetan biosynthesis (Griffin 1994; Ishida 2002b). These repeating units are subsequently polymerized to form the final HePS (Schmid 2014) and therefore reflecting the monomer composition of the entire HePS. A study on acetan polysaccharides produced by *Ko*. *xylinus* also demonstrated that the monomer composition of HePS is independent of the provided carbon source (Kornmann et al. 2003). Accordingly, it is assumed that the *K*. *baliensis* NBRC 16680 HePS is produced in a similar way, which would result in a constant structural composition of the HePS. Still, despite the same composition the NaG and NaG-fgm derived HePS show differences in their molecular weight. A possible explanation why the NaG and NaG-fgm derived HePS differ in their size, could be that magnesium in the medium enhances the activation of polymerization enzymes and therefore the oligomerization of the repeating core unit. The stimulated effect of magnesium, related to the second messenger bis-(3´-5´)-cyclic-GMP, has already been discussed in connection with the cellulose synthase of *K*. *baliensis* DSM 14400. A similar situation is found in case of alginate, produced by *Pseudomonas aeruginosa*, where the polymerization reaction of alginate requires c-di-GMP-binding to the inner-membrane protein Alg44; therefore increasing c-di-GMP levels enhanced the formation of alginate in mucoid strains (Merighi 2007). Alg44 is a single-pass transmembrane protein that is involved in the polymerization and export of alginate across the inner membrane (Oglesby 2008). Accordingly, magnesium influences in this case, the polymerization process of alginate. Therefore, it could have a similar effects on the *K*. *baliensis* NBRC 16680 HePS synthesis, assuming an analogous function during the polymerization of the backbone of the HePS. The responsible enzymes, which are possibly involved in the polymerization and export process of *K*. *baliensis* NBRC 16680 HePS are GumB, GumC and GumE (Becker et al. 1998). GumC has three predicted transmembrane helices (TMH), as well as an N-terminal Wzz-domain (E-value: 1.23e-09, Bit Score: 55.82), and a C-terminal tyrosine autokinase domain (PEPCTERM_TyrKin, E-value: 1.13e-14, Bit Score: 71.94). Both are characteristic for polysaccharide copolymerase2 (PCP-2a) proteins, which regulate the chain length of high molecular weight EPS in *E*. *coli* (Cuthbertson 2009; Franco 1998). Therefore, magnesium in the medium may possibly result in a longer main chain

of the NaG-fgm medium derived *K*. *baliensis* NBRC 16680 HePS, leading to an larger molecule with a constant monomer composition, similar to the composition of the NaG medium derived HePS.

In conclusion it can be said, that through the method of comparative genomics it is possible to predict particular structural components of the HePS, which considerably simplifies the following structural analysis. However, the holistic structure of the particular HePS can not be foreseen, due to further structural influencing transcriptional and translational factors. Furthermore, no direct link between the *gum*-like cluster and the *K*. *baliensis* HePS could be revealed up to this point.

The final proof, that the *gum*-like HePS cluster of *K*. *baliensis* is the genetic basis of the HePS, was verified via a mutation in an essential HePS-formation associated gene (*gumD*) of *K*. *baliensis* NBRC 16680. The mutant strain forms rough, dull colonies (R-strain), instead of smooth, slimy colonies (S-strain). Via PCR screening of the *gum*-like cluster of *K*. *baliensis* a transposon insertion in front of the *gumD* gene was found (Figure 7, PART I). This gene catalyzes the first step of HePS synthesis in *X*. *campestris*, by transferring a glucosyl-1-phosphate residue from UDP-glucose to an undecaprenyl phosphate residue. For *X*. *campestris* it was shown, that a mutation in this gene leads to a lack in xanthan synthesis, confirming the essential function of this protein during xanthan synthesis (Ielmini 1998). Therefore, a similar function can be assumed in the case of *K*. *baliensis* NBRC 16680. Mutations in other HePS forming clusters, like the *polABCDE* cluster, could be excluded by further PCR screenings.

The nucleotide sequence of the mobile element (1313 bp) is identically present at three further genomic regions of the *K*. *baliensis* NBRC 16680 chromosome (77,364–78,676 bp, 1,430,134–1,428,821 bp, 2,279,289– 2,280,601 bp), whereas one of them is closely located to the levansucrase gene (A0U90_10610; 2,276,746 – 2,275,586 bp). To shed light on, if the mutation is a random event, or if it is a directed mutation, possibly triggered by environmental factors, we performed an experimental series using different growth media. We suspected, that enriched sugar medium (NaG-s, NaG-sfg, NaG-fgm) will promote wild-type growth, whereas stress inducing medium (NaG-EtOH, NaG-AA) will enhance the transposon induced mutant colonies. It appeared that *K*. *baliensis* NBRC 16680 showed similar numbers of mutant and wild-type colonies in nearly all media, expect under acetic acid fermenting conditions in the presence of ethanol, where mutant strains were more frequently found, displaying high mutation frequencies (f: 0.93) (Figure 1, PART III). The transcriptional activity of mobile elements is

generally maintained at a low level, but a high frequency of transposition can occur in response to curtain environmental changes, as in case of nutrient deficiency, like carbon starvation (Lamrani et al. 1999) and temperature effects (Kretschmer and Cohen 1979). This stress-induced stimulation of mobile elements can serve as an adaptive response of environmental changes, to increase the genetic plasticity of a population by permitting new characteristic traits to evolve, and therefore to overcome life-threatening situations (Mont et al. 2000; Wessler 1996). Also in AAB an induced loss of various physiological properties could be observed, such as, acetic acid resistance (Ohmori 1981), ethanol oxidation, pellicle formation (Deeraksa et al. 2005), and bacterial cellulose synthesis. The genetic mechanisms behind these instabilities is often unclear. However, a connection with the insertion of mobile elements could be shown for some AAB. Beside the inactivation of an alcohol dehydrogenase gene in *A. pasteurianus* NC11380 (Takemura 1991) and NCI 1452 (Kondo 1997) by an insertion of a mobile element, a gene inactivation provoked by transposable elements could also be demonstrated for EPS forming clusters. In *A. xylinum* ATCC 23769, an IS element caused insertions 0.5 kb upstream of the cellulose synthase gene, associated with spontaneous cellulose deficiency (Coucheron 1991). In any cases, it was unsuccessful to sustain revertants, as a consequence of the remaining directed repeats (DR) of the IS element after relocation (Figure 7a, PART I). In case of *K. baliensis* NBRC 16680 R, likewise no revertants could be obtained after repeated cultivation on NaG agar plates. This leads to the conclusion, that the insertion of the mobile element leads to a non reversible, constant phenotypic change in *K. baliensis* NBRC 16680 by e. g. remaining direct repeats at the insertion site. The frequent occurrence of the mutant *K. baliensis* NBRC 16680 strain in the presence of ethanol, leads to the suggestion that the functional loss of the HePS production in a curtained part of the population, takes place to create diversity. This morphological switch between a slimy wild-type and a rough, maybe capsular phenotype would results in new competences of the strain, to cope with the increased ethanol concentrations. Morphotype variations, including the transition from a mucoid to a non-mucoid phenotype, are common events within the family of *Acetobacteraceae*, especially for *Acetobacter* species. This phenotypic change is accompanied by a transformation from secreted EPS to CPS, mostly in form of a pellicle floating on the medium surface (Ali et al. 2011b; Moonmangmee 2002a), which is able to serve as a barrier against ethanol (Perumpuli et al. 2014).

The accountable, so-called *pol*-cluster (*polABCDE*) is responsible for pellicle formation in different *Acetobacte*r strain (Deeraksa et al. 2005). The *polABCD* genes show a high level of homology to the *rfbBACD* cluster, which is involved in dTDP-rhamnose synthesis (Marolda

CL 1995). Whereas the *polE* gene has a relatively low homology level to known glycosyltransferases, which is why it has undergone several assignments. However, a disruption of the *polE* gene in *A. tropicalis* SKU1100 led to a defect in the pellicle formation, in this manner giving it a central role, either as rhamnosyl-transferase (Deeraksa et al. 2005), or as galactosyl-transferase (Ali et al. 2011b), connecting the CPS to the cell surface. Additionally, the obtained Pel$^-$ mutant cells secreted EPS into the culture medium. The rough mutant of *K. baliensis* NBRC 16680 doesn't produce a *gum*-cluster dependent HePS anymore, but is able to form a pellicle under static cultivation (Figure 3A, PART III), likewise to *A. tropic*alis, suggesting that the *pol*-cluster, instead of the *gum*-cluster, is responsible for the pellicle construction. To investigate the relationship between *pol*-cluster and pellicle/CPS formation in *K. baliensis* NBRC 16680, *polE* knockouts were carried out in *K. baliensis* NBRC 16680 R. Furthermore, we were interested to observe, if a *polE* knockout has an effect on the ethanol tolerance of the respective *K. baliensis* strain. It turned out, that the Δ*polE* mutant of *K. baliensis* NBRC 16680 R showed the same rough colony morphology as the rough strain (*K. baliensis* NBRC 16680 R). Surprisingly *K. baliensis* NBRC 16680 R Δ*polE* was still able to form a pellicle on the surface, in contrast to *A. tropicalis* SKU1100. However, the generated pellicle was only a loose conglomerate of cells and required considerably more time for formation, compared to the rough strain (Figure 3A, PART III). Similar to *A. tropicalis* SKU1100, the Δ*polE* mutant of *K. baliensis* NBRC 16680 R secreted small amounts of EPS into the medium (Figure 3C, PART III),. The weak pellicle could therefore be formed by the secreted EPS, thus leading to the formation of an incoherent EPS/ cell-layer on the surface. The secreted EPS showed a similar composition as the *K. baliensis* NBRC 16680 R capsular HePS (Figure 3D, PART III), supporting the hypothesis of Ali et al (2011b) that PolE is responsible for CPS formation, via addition of specific residues, which connect the HePS with the cell surface, e.g. β-d-galactopyranosyl residues. Therefore, a *polE* gene deletion or interruption would results in a switch from a rough CPS producing, to an EPS producing phenotype, which would be possibly connected with growth behavior variations. It was shown that a change from a smooth phenotype to the rough phenotype could be associated with an increased tolerance against acetic acid (Deeraksa et al. 2005). By growth experiments with different ethanol and acetic acid concentrations of the different *K. baliensis* strains (NBRC 16680, NBRC 16680 R, NBRC 16680 R Δ*polE*), deviations in the growth behaviors of the strains could be observed. Further, to investigate the variability in the growth behavior between the two wild-type strains of *K. baliensis*, also *K. baliensis* DSM 14400 was tested. It emerged, that the rough mutant strain

showed under ethanol and acetic acid stress a higher assertiveness compared to the wild-type strain (NBRC 16680) (Figure 4, PART III). During static growth, a similar pattern to the NaG plates appeared, with higher growth rates of *K*. *baliensis* NBRC 16680 R in NaG medium with acetic acid and ethanol (Figure 5, PART III), compared to the NBRC 16680 wild-type strain. However, the Δ*polE* strain, showed overall a diminished growth on acetic acid and ethanol. As already mentioned, we assume that the knockout of the *polE* gene results in a change from CPS to EPS. The resulting CPS could serves as a better barrier against acetic acid (Deeraksa et al. 2006), indicating that the CPS of *K*. *baliensis* NBRC 16680 R also could be involved (*inter alia*) in the protection against acetic acid and ethanol. A direct relationship between pellicle formation and acetic acid resistance was proven for *A*. *pasteurianus* (IFO3283, SKU1108, MSU10), where the rough strains sowed clearly higher acetic acid resistance than the smooth phenotypes, respectively (Kanchanarach 2014). The pellicle works in this case as a biofilm-like barrier and prevents the passive diffusion of acetic acid into the cells (Kanchanarach et al. 2010). Furthermore, Perumpuli et al (2014) showed, that ethanol in the medium significantly induced pellicle formation. This is in agreement with our observations for *K*. *baliensis* NBRC 16680, which mutated more frequently in NaG medium with ethanol (Fig.1), suggesting a directed mutation, which could be triggered under acetic acid fermenting conditions by ethanol or its oxidized product acetic acid. By implicating the behavior of the wild-type strains, it could be observed, that *K*. *baliensis* DSM 14400 showed also a high resistance against acetic acid (Figure 4A, PART III). However, this strain was derogated by ethanol concentrations up to 8%, compared to *K*. *baliensis* NBRC 16680 R, which was still able to grow at these concentrations. Since both *K*. *baliensis* wild-type strains demonstrate different behaviors during an equal stress conditions, it can be concluded, that both strains developed different defense strategy. Both strains were isolated from different environments (DSM 14400 isolated from palm brown sugar, NBRC 16680 isolated from ragi (Lisdiyanti 2002)). Therefore, it may be that *K*. *baliensis* DSM 14400 commonly challenges with high acetic acid concentrations, while they represent a new environmental factor for *K*. *baliensis* NBRC 16680. In contrast to *K*. *baliensis* NBRC 16680, in the genome of *K*. *baliensis* DSM 14400, a cellulose synthase operon was identified on plasmid 3 (pKB14400_3), with genes coding for the three cellulose synthase subunits A, B and C (Figure 5d, PART I). The additionally formed cellulose could act as barrier against acetic acid, thereby providing an increased protection for *K*. *baliensis* DSM 14400. The same mechanism is found for *Ko*. *xylinus* E25, isolated from vinegar, which generally deals with high acetic acid concentrations (Ishida 2002a). Further, also a direct genetic connection between the

pol- and *gum*-like clusters can be observed in *Ko*. *xylinus* (Figure 6, PART III). *Ko*. *xylinus* E25 contains no *polE* gene in its genome, but is well known to produce a pellicle polysaccharide consisting of bacterial cellulose (Saxena 1994) and is also able to produce an extracellular HePS called acetan, composed of D-glucose, D-mannose, D-rhamnose, and D-glucuronic acid (Ishida 2002a; Jansson et al. 1993). The included rhamnosyl residues, are processed by the *polABCD* genes (Deeraksa et al. 2005), coding for the enzymes of the TDP-rhamnose synthesis. In case of *Ko*. *xylinus* E25, the *pol*-cluster is flanking the acetan producing *gum*-like cluster, with the *polAB* gene upstream and the *polCD* gene downstream of the cluster (Figure 6A, PART III), connecting the acetan synthesis with the enzymes of the *pol*-cluster. The *polE* gene is missing in the *Ko*. *xylinus* E25 genome (GCA_000550765.1), possibly resulting in no HePS mediated pellicle formation, while a tight cellulose pellicle is commonly formed on the surface of the medium (Brown 1886), which contributes to increased acetic acid resistance of the bacterium (Toda 1997).

This displays that the *gum*- and the *pol*-cluster are possibly related with each other and could be at once regulated according to the environmental requirements of the respective bacterium. An increasing ethanol concentration or lack of oxygen, could be a signal, which leads to an up-regulation of the *polE* gene in *K*. *baliensis* NBRC 16680. This can take place in association with an ethanol or acetic acid triggered stress response, probably resulting in an induced deactivation of the *gum* clusters, via a transposon insertion. *Ga*. *diazotrophicus* PAl5 produces a HePS composed of D-glucose, D-galactose and D-mannose, whose production is most likely attributable to its *gum*-like cluster similar to *K*. *baliensis* (Bertalan et al. 2009). In the genome of *Ga*. *diazotrophicus* PAl5, no *polE* gene is present (Figure 6B, PART III). In contrast to *Ko*. *xylinus*, *Ga*. *diazotrophicus* is an obligate endophytic bacterium that lives symbiotically (N2-fixing) in the intercellular space of roots, stem and leaves of sugarcane plants (Cavalcante 1988), whereas the HePS plays a key role during plant colonization (Meneses 2011). In this context, the genetic switch of *K*. *baliensis* NBRC 16680 wild-type to its rough mutant could indicate an environmentally driven conversion/adaptation of a formerly more plant associated AAB (wild-type) to an increasingly acetic acid producing AAB (rough mutant). Furthermore, *pol*-cluster mediated capsular pellicle biosynthesis and concomitant inactivation of excess HePS production/secretion seems to increase the tolerance against acetic acid in certain AAB, such as *K*. *baliensis* NBRC 16680. On the other hand, specialized AAB additionally produce tight cellulose pellicles, which are typical for extremely acetic acid tolerant starter cultures used in vinegar production, such as *Komagataeibacter xylinus*. Thus, it could be a daring hypothesis

that the *polE* gene solely occurs in AAB strains, lacking a cellulose pellicle formation. So far, the *pol*-cluster was more or less exclusively observed for *Acetobacter* strains, with *A. tropicalis* SKU1100 as the most prominent example (Ali et al. 2011a; Deeraksa et al. 2005). To obtain a uniform overview of the *pol* genes in AAB, we compared *pol*-clusters of different AAB strains via whole genome sequencing of specific strains (*A. aceti* TMW 2.1153, *A. persici* TMW 2.1084 and *N. chiangmaiensis* NBRC 101099) and by us of already existing genome data. By comparing the individual clusters of different *Komagataeibacter* strains (*Ko. xylinus* E25, *Ko. europaeus* LMG 18890, LMG 18494, 5P3, NBRC 326, *Ko.oboediens* 174Bp2, *Ko. hansenii* ATCC 23769), it could be seen, that they do not include the *polE* gene in their *pol*-cluster (Andres-Barrao 2011; Iyer 2010; Kubiak 2014). Consistent with the *polE* gene absence, they form a net cellulose pellicle on the surface of a static culture, instead of a HePS pellicle (Andres-Barrao 2011; Iyer 2010; Saxena 1994; Yamada 2012). Further, it could be shown, that in case of *Ko. europaeus* NBRC 3261 and *Ko. xylinus* E25 the *pol*-cluster is fused with the *gum*-like cluster of the respective strain (Figure 3). This assembling of both clusters has not yet been shown. Further, *Ko. xyli*nus E25 secrets an HePS, acetan, in the external enviroment (Griffin 1994), whereas for *Ko. europaeus* NBRC 3261 a HePS formation has not yet been described. However, the strain displays a similar *gum*-like cluster compared to *Ko. xylinus* E25 (Figure 3), which suggests a comparable HePS production. It seems that the *pol*-cluster is no longer needed for the highly acetic acid resistant *Komagataeibacter* strains, since the cellulose pellicle appears to provide an adequate protection. This agrees with the observation described above, which showed that *K. baliensis* DSM 14400 already exhibits an intrinsic acetic acid resistance, also containing a cellulose synthase cluster, possibly forming cellulose. However, an exception to this is *Ko. xyli*nus E25, which forms an additional secreted HePS that contains rhamnose, therefore incorporating the *polABCD* genes, which are involved in the rhamnose synthesis, in its *gum*-like cluster. If you add *Acetobacter* to this comparison, it shows that they harbor complete *polABCDE* clusters (Figure 3) (Deeraksa et al. 2005; Kappock 2014) and mostly form HePS pellicles on the surface of a static culture (Perumpuli et al. 2014). In this context, we could identify complete *pol*-clusters in the genomes of *A. persici* TMW 2.1084 and *A. aceti* TMW 2.1153, which produces a HePS pellicle composed of glucose and rhamnose (Moonmangmee 2002a). This is also known for *A. tropicalis* SKU1100, which harbors a complete *polABCDE* cluster and produces a pellicle consisting of galactose, glucose and rhamose. Both genus (*Komagataeibacter* and *Acetobacter*) show a basic, up to high acetic acid resistance and mark main actors in the acetic acid fermentation. By the previous cohesion of

the *Acetobacter* and *Komagataeibacter* strains, which have been split up with increasing knowledge, the graphic illustrates a part of the difference between the different species. As in case of *Kozakia baliensis*, different kinds of resistance strategies seems to have developed during the evolutionary process and are among other things, represented by different EPS cluster formations.

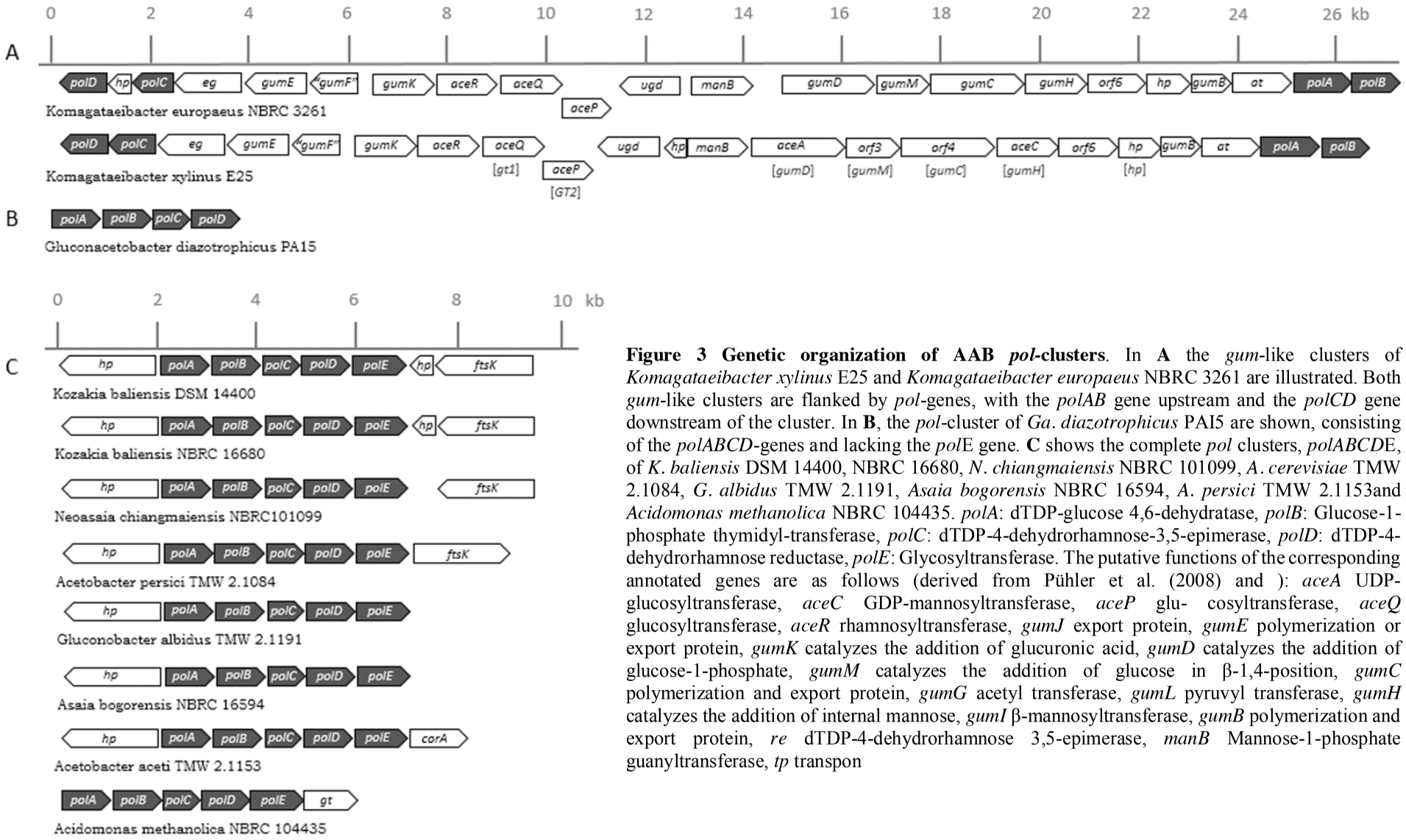


Figure 3 Genetic organization of AAB *pol*-clusters. In **A** the *gum*-like clusters of *Komagataeibacter xylinus* E25 and *Komagataeibacter europaeus* NBRC 3261 are illustrated. Both *gum*-like clusters are flanked by *pol*-genes, with the *polAB* gene upstream and the *polCD* gene downstream of the cluster. In **B**, the *pol*-cluster of *Ga. diazotrophicus* PAI5 are shown, consisting of the *polABCD*-genes and lacking the *pol*E gene. **C** shows the complete *pol* clusters, *polABCD*E, of *K. baliensis* DSM 14400, NBRC 16680, *N. chiangmaiensis* NBRC 101099, *A. cerevisiae* TMW 2.1084, *G. albidus* TMW 2.1191, *Asaia bogorensis* NBRC 16594, *A. persici* TMW 2.1153and *Acidomonas methanolica* NBRC 104435. *polA*: dTDP-glucose 4,6-dehydratase, *polB*: Glucose-1-phosphate thymidyl-transferase, *polC*: dTDP-4-dehydrorhamnose-3,5-epimerase, *polD*: dTDP-4-dehydrorhamnose reductase, *polE*: Glycosyltransferase. The putative functions of the corresponding annotated genes are as follows (derived from Pühler et al. (2008) and): *aceA* UDP-glucosyltransferase, *aceC* GDP-mannosyltransferase, *aceP* glu- cosyltransferase, *aceQ* glucosyltransferase, *aceR* rhamnosyltransferase, *gumJ* export protein, *gumE* polymerization or export protein, *gumK* catalyzes the addition of glucuronic acid, *gumD* catalyzes the addition of glucose-1-phosphate, *gumM* catalyzes the addition of glucose in β-1,4-position, *gumC* polymerization and export protein, *gumG* acetyl transferase, *gumL* pyruvyl transferase, *gumH* catalyzes the addition of internal mannose, *gumI* β-mannosyltransferase, *gumB* polymerization and export protein, *re* dTDP-4-dehydrorhamnose 3,5-epimerase, *manB* Mannose-1-phosphate guanyltransferase, *tp* transpon

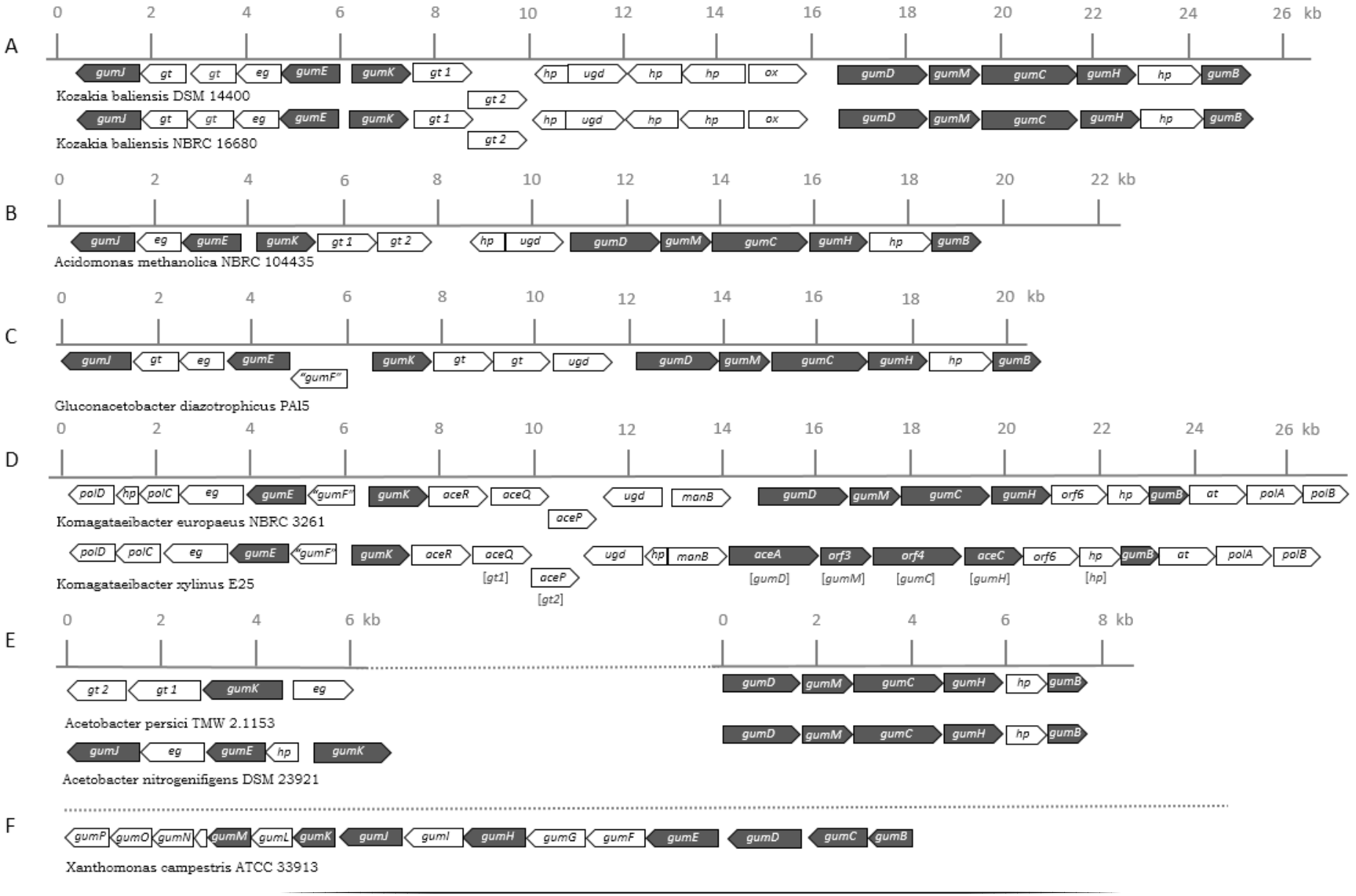

A
0 2 4 6 8 10 12 14 16 18 20 22 24 26 kb
gumJ gt gt eg gumE gumK gt 1 hp ugd hp hp ox gumD gumM gumC gumH hp gumB
Kozakia baliensis DSM 14400
gt 2
gumJ gt gt eg gumE gumK gt 1 hp ugd hp hp ox gumD gumM gumC gumH hp gumB
gt 2
Kozakia baliensis NBRC 16680
B
0 2 4 6 8 10 12 14 16 18 20 22 kb
gumJ eg gumE gumK gt 1 gt 2 hp ugd gumD gumM gumC gumH hp gumB
Acidomonas methanolica NBRC 104435
C
0 2 4 6 8 10 12 14 16 18 20 kb
gumJ gt eg gumE "gumF" gumK gt gt ugd gumD gumM gumC gumH hp gumB
Gluconacetobacter diazotrophicus PAl5
D
0 2 4 6 8 10 12 14 16 18 20 22 24 26 kb
polD hp polC eg gumE "gumF" gumK aceR aceQ ugd manB gumD gumM gumC gumH orf6 hp gumB at polA polB
aceP
Komagataeibacter europaeus NBRC 3261
polD polC eg gumE "gumF" gumK aceR aceQ ugd hp manB aceA orf3 orf4 aceC orf6 hp gumB at polA polB
[gt1] aceP [gumD] [gumM] [gumC] [gumH] [hp]
[gt2]
Komagataeibacter xylinus E25
E
0 2 4 6 kb
gt 2 gt 1 gumK eg
Acetobacter persici TMW 2.1153
gumJ eg gumE hp gumK
Acetobacter nitrogenifigens DSM 23921
0 2 4 6 8 kb
gumD gumM gumC gumH hp gumB
gumD gumM gumC gumH hp gumB
F
gumP gumO gumN gumM gumL gumK gumJ gumI gumH gumG gumF gumE gumD gumC gumB
Xanthomonas campestris ATCC 33913

Figure 4 Genetic organization of AAB *gum*-like clusters. The *gum*-like clusters of *K*. *baliensis* DSM 14400 and NBRC 16680 are depicted in **A**. The cluster has an overall size of ~25 kb and involves 19 genes, including glycosyltransferases (*gt*), hypothetical proteins (*hp*), and eight *gum* like genes (*gumB, -C, -D, -E, -H, -J, -K,* and *–M*), which are marked in grey. Furthermore, the cluster contains a putative endoglucanase (*eg*), oxidoreductase (*ox*) and a UDP-glucose dehydrogenase (*ugd*). **B** shows the *gum* like clusters of the AAB strains *Acidomonas methanolica* NBRC 104435. **C** and **D** show the related *gum* like clusters of the AAB strains *Ga*. *diazotrophicus* PAI5, *Komagataeibacter europaeus* NBRC 3261 and *Komagataeibacter xylin* E25. The *Ga*. *diazotrophicus* cluster exhibits, in comparison to both *K*. *baliensis* clusters, an additional *gumF* gene, that could putatively incorporate acetyl-residues at specific positions into the related HePS. The so called acetan cluster of *Ko*. *xylinus* harbors—besides an additional *gumF* gene—a rhamnosyl transferase, coded by *aceR*, as well as a mannose-phosphate-guanyltransferase (*mpg*). The nomenclature for the acetan cluster in **C** is based on Griffin (1994), while brackets under the particular genes mark the homologous *gum* genes. In **E** are the fragments of potentially *gum*-like clusters of *Acetobacter persici* TMW 2.1153 and *Acetobacter nitrogenifigens* DSM 23921. The particular fragments are located in different parts of the genome, respectively. In **F**, the *gum* cluster of *X*. *campestris*, which consists of *gumB, -C, -D, -E, -F, -G, -H, -I, -J, -K, -L, -M, -N* and *–P*, is depicted. The putative functions of the corresponding annotated genes are as follows (derived from Pühler et al. (2008) and): *aceA* UDP-glucosyltransferase, *aceC* GDP-mannosyltransferase, *aceP* glu- cosyltransferase, *aceQ* glucosyltransferase, *aceR* rhamnosyltransferase, *gumJ* export protein, *gumE* polymerization or export protein, *gumK* catalyzes the addition of glucuronic acid, *gumD* catalyzes the addition of glucose-1-phosphate, *gumM* catalyzes the addition of glucose in β-1,4-position, *gumC* polymerization and export protein, *gumG* acetyl transferase, *gumL* pyruvyl transferase, *gumH* catalyzes the addition of internal mannose, *gumI* β-mannosyltransferase, *gumB* polymerization and export protein, *re* dTDP-4-dehydrorhamnose 3,5-epimerase, *manB* Mannose-1-phosphate guanyltransferase, *tp* transpon

Ga. *diazotrophicus* PAl5 does not show a *polE* gene in his *pol*-cluster as well, whereby here, the cluster is still coherent and disconnected from the actual HePS cluster. The other observed AAB strains showed completed *pol* clusters, consisting of the *polABCDE* genes (Figure 3). As already described, a HePS pellicle formation could also be described for bot *K*. *baliensis* strains (DSM 14400, NBRC 16680), which is connected to the *polABCDE* cluster and based on the *polE* function.

An overview of the *gum*-like gene in AAB shows, that they occur mainly in *Gluconacetobacter*, *Komagataeibacter* and *Kozakia*, while they are distributed in the genome, or at least not present, in the most *Acetobacter* strains (Figure 4). *Acidomonas methanolica* NBRC 104435 shows the most similar cluster in comparison to *K*. *baliensis* (DSM 14400, NBRC 16680), however, no HePS formation has yet been described for this strain. *Ga*. *diazotrophicus* PAI5, *Komagataeibacter europaeus* NBRC 3261 and *Komagataeibacter xylin* E25 have already been described as HePS producers. The assignment of the *gum*-like genes in *Acetobacter* strains has not yet been explored, but for example, they could play a role during the pellicle formation. Although this cluster review does not yet provide complete answers to specific questions, it is still the first graphic, which provides an overview of the *gum*-like genes in AAB and thus makes it easier for the future to recognize cross-strain relationships.

LIST OF PUBLICATIONS DERIVED FROM THIS WORK

Peer-reviewed journals

Brandt, J. U., Jakob, F., Behr, J., Geissler, A.J., Vogel, R. F. (2016). **Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*.** Microbial Cell Factories, 15(1), 170.

Brandt, J. U., Jakob, F., Geissler, A.J., Behr, J., Vogel, R. F. (2017). **Multiple Genome Sequences of Heteropolysaccharadie-Forming Acetic Acid Bacteria.** Genome Announcements, 5(16).

Brandt, J. U., Born, F., Jakob, F., Vogel, R.F. (2017). **Environmentally triggered genomic plasticity and capsular polysaccharide formation are involved in increased ethanol and acetic acid tolerance in *Kozakia baliensis* NBRC 16680.** BMC Microbiology, 17(172).

Brandt, J. U., Welfers, D., Jakob, F., Bunzel, M., Vogel, R. F. (2017). **Characterization of an acetan-like heteropolysaccharide produced by *Kozakia baliensis* NBRC 16680.** International Journal of Biological Macromolecules, (106).

Oral presentations

Brandt, J. U., Jakob, F., Behr, J., Geissler, A.J., Vogel, R. F. (2015). Characterisation of a novel Heteropolysaccharide produced by *Kozakia baliensis*, 4th International Conference on Acetic Acid Bacteria, Taiyuan, China, Section VI, Presentation VI-4.

Oral presentations (co-author, speakers are underlined)

Welfers, D., Frauenhofer, M., **Brandt, J. U**., Xu, D., Jakob, F., Vogel, R. F., Bunzel, M. (2017). Strukturelle Eigenschaften verschiedener Lebensmittelrelevanter bakterieller Exopolysaccharide. Symposium der Gesellschaft Deutscher Lebensmitteltchnologen „Hydrokolloide VIII“, Gerlingen.

Poster presentation

6th European Conference on Prokaryotic and Fungal Genomics, ProkaGENOMICS, 29 September – 2 October 2015, Göttingen. Germany. Brandt, J. U., Jakob, F., Vogel, R.F., Exploitation of *Gluconobac*ter genomes towards the understanding of high levan production of *Gluconobacter species* TMW 2.1191.

Food Micro, 19 – 22 July 2016, Dublin, Ireland. Brandt, J. U., Jakob, F., Vogel, R.F., Characterisation of a novel food relevant heteropolysaccharide produced by *Kozakia baliensis*.

CURRICULUM VITAE

Personal Data

Name	Julia Brandt
Date of Birth:	10 June 1987
Place of Birth:	Koblenz

Personal and professional development

03/2014 –2018	**Doctorate (Dr. rer. nat.)** TU Munich, Center of Life and Food Science Weihenstephan, Lehrstuhl für technische Mikrobiologie Title of the thesis: *"Structure, biosynthesis and possible ecological role of a novel heteropolysaccharide from Kozakia baliensis"*
08/2013 - 02/2014	**Research Assistant** Bielefeld University
06/2013	**Diploma in Biology** Johannes Gutenberg Universität Mainz
10/2007 - 03/2013	**Studies in Biology** Johannes Gutenberg Universität Mainz Advanced subjects: Biochemistry, Molecular Genetics (1,0), Microbiology (1,0), Immunology (1,0) Title of the diploma thesis: *„Funktionelle Charakterisierung trans-interagierender Regionen in Abhängigkeit des methylierungsspezifischen Status der H19/IGF2-Imprintingkontrollregion" (Note 1,3)*
10/2006 - 07/2007	**Studies in Psychology** Université de Fribourg
03/2006	**Abitur,** Hilda-Gymnasium Advanced courses: Math, Biology and History

APPENDIX

Supplementary file 1 Screening of AAB strains for their ability to produce HePS from sucrose deficient medium HePS producing activity was observed macroscopically as follows: (+) weak, (++) adequate, (+++) strong and (-) no production of mucous slime on sucrose deficient NaG agar plates.

Strain	Origin	HePS forming potential
A. aceti	TMW 2.1153 (isolate from water kefir)	-
A. persici	TMW 2.1084 (isolate from water kefir)	++
A. pasteurianus	DSM 2324	-
A. pasteurianus	TMW 2.881	~
Asaia bogorensis	IFO 16594	~
Ga. azotocaptans	DSM 13594	-
Ga. frateurii	TMW 2.767	~
Ga. frateurii	DSM 7146	-
Ga. hansenii	DSM 5602	-
Ga. johannae	DSM 13595	-
Ga. liquefaciens	DSM 5603	-
Ga. sacchari	DSM 12717	-
G. albidus	TMW 2.1191 (isolate from water kefir)	+
G. cerinus	DSM 9533	-
G. cerinus	TMW 2.155	-
G. oxydans	DSM 46615	-
G. oxydans	DSM 50049	-
G. oxydans	DSM 3503	-
G. oxydans	NCIMB 8084	-
G. oxydans	DSM 2003	-
G. oxydans	TMW 2.1085	~
G. oxydans	TMW 2.877	-
G. species	DSM 3504	~
G. species	TMW 2.767 (isolate from water kefir)	+
K. baliensis	TMW 2.1087; DSM 14400	+++
K. baliensis	TMW 2.1340; NBRC 16680	+++
N. chiangmaiensis	TMW 2.1086; NBRC 101099	+++

REFERENCES

Abdian PL, Lellouch, A.C., Gautier, C., Ielpi, L., Geremia, R.A. (2000) Identification of essential amino acids in the bacterial alpha-mannosyltransferase AceA. Journal of Biological Chemistry 275(51):40568–40575

Alauzet C, Teyssier, C., Jumas-Bilak, E., Gouby, A., Chiron, R., Rabaud, C., Counil, F., Lozniewski, A., Marchandin, A. (2010) Gluconobacter as well as Asaia species, newly emerging opportunistic human pathogens among acetic acid bacteria. Journal of Clinical Microbiology 48:3935-3942

Ali IAI, Akakabe Y, Moonmangmee S, Deeraksa A, Matsutani M, Yakushi T, Yamada M, Matsushita K (2011a) Structural characterization of pellicle polysaccharides of Acetobacter tropicalis SKU1100 wild type and mutant strains. Carbohydrate Polymers 86(2):1000 - 1006

Ali IAI, Akakabe Y, Moonmangmee S, Deeraksa A, Matsutani M, Yakushi T, Yamada M, Matsushita K (2011b) Structural characterization of pellicle polysaccharides of *Acetobacter tropicalis* SKU1100 wild type and mutant strains. Carbohydrate Polymers 86(2):1000-1006 doi:10.1016/j.carbpol.2011.05.055

Altschul SF, Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) Basic local alignment search tool. Journal of Molecular Biology 215(3):403-410

Andres-Barrao C, Falquet, L., Calderon-Copete, S. P., Descombes, P., Perez, R. O., Barja, F. (2011) Genome sequences of the high-acetic acid-resistant bacteria *Gluconacetobacter europaeus* LMG 18890T and *G. europaeus* lmg 18494 (reference strains), *G. europaeus* 5P3, and *Gluconacetobacter oboediens* 174Bp2 (isolated from vinegar). Journal of Bacteriology 193(10):2670-2671

Angiuoli SV, Gussman, A., Klimke, W., Cochrane, G., Field, D., Garrity, G., Kodira, C.D., Kyrpides, N., Madupu, R., Markowitz, V., Tatusova, T., Thomson, N., White, O. (2008) Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. OMICS 12(2):137-141

Aziz RK, Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., Kubal, M., Meyer, F., Olsen, G.J., Olson, R., Osterman, A.L., Overbeek, R.A., McNeil, L.K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G.D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., Zagnitko, O. (2008) The RAST Server: rapid annotations using subsystems technology. BMC genomics 75(9):1-15

Baird JK, Sandford, P. A., Cottrell, I. W. (1983) Industrial application of some new microbial polysaccharides. Nature Biotechnolgy 1983(1):778 - 783

Barnes HA, Hutton, J.F., Walters, K. (1989) An Introduction to Rheology. Elsevier Applied Science

Barreras M, Abdian, P.L., Ielpi, L. (2004) Functional characterization of GumK, a membrane-associated beta-glucuronosyltransferase from *Xanthomonas campestris* required for xanthan polysaccharide synthesis. Glycobiology 14(3):233-241

Becker A, Katzen F, Pühler A, Ielpi L (1998) Xanthan gum biosynthesis and application : a biochemical/genetic perspective. Applied Microbiology and Biotechnology 50:145-152

Bertalan M, Albano R, de Padua V, Rouws L, Rojas C, Hemerly A, Teixeira K, Schwab S, Araujo J, Oliveira A, Franca L, Magalhaes V, Alqueres S, Cardoso A, Almeida W, Loureiro MM, Nogueira E, Cidade D, Oliveira D, Simao T, Macedo J, Valadao A, Dreschsel M, Freitas

F, Vidal M, Guedes H, Rodrigues E, Meneses C, Brioso P, Pozzer L, Figueiredo D, Montano H, Junior J, de Souza Filho G, Martin Quintana Flores V, Ferreira B, Branco A, Gonzalez P, Guillobel H, Lemos M, Seibel L, Macedo J, Alves-Ferreira M, Sachetto-Martins G, Coelho A, Santos E, Amaral G, Neves A, Pacheco AB, Carvalho D, Lery L, Bisch P, Rossle SC, Urmenyi T, Rael Pereira A, Silva R, Rondinelli E, von Kruger W, Martins O, Baldani JI, Ferreira PC (2009) Complete genome sequence of the sugarcane nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus* Pal5. BMC genomics 10:450 doi:10.1186/1471-2164-10-450

Brandt JU, Jakob, F., Behr, J., Geissler, A.J., Vogel, R.F. (2016) Dissection of exopolysaccharide biosynthesis in *Kozakia baliensis*. Microbial Cell Factories 15(1):170

Brown AJ (1886) On an acetic ferment which forms cellulose. Journal of the Chemical Society 49:432-439

Camacho C, Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L. (2009) BLAST+: architecture and applications. BMC Bioinformatics 421(10):1-9

Cavalcante A, V., Dobereiner, J. (1988) A new acid-tolerant nitrogen-fixing bacterium associated with sugarcane. Plant and Soil 108(1):23-31

Charney W, Montclair, N.J. (1987) Process for the production of dihydroxyacetone. USA, patent No 4076589

Chin C-S, Alexander DH, Marks P, Klammer Aa, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nature methods 10:563-571

Cleenwerck I, De Vos, P. (2008) Polyphasic taxonomy of acetic acid bacteria: an overview of the currently applied methodology. International journal of food microbiology 125(1):2-14

Coucheron DH (1991) An *Acetobacter xylinum* insertion sequence element associated with inactivation of cellulose production. Journal of Bacteriology 173(18):5723–5731.

Couso RO, Ielpi, L., & Dankert, M. A. (1987) A xanthan-gum-like polysaccharide from *Acetobacter xylinum.* Microbiology 133:2123-2135

Creaven M, Fitzgerald RJ, O'Gara F (1994) Transformation of *Gluconobacter oxydans subsp. suboxydans* by electroporation. Canadian Journal of Microbiology 40(6):491-494

Crotti C, Rizzi, A., Chouaia, B., Ricci, I., Favia, G., Alma, A., Sacchi, L., Bourtzis, K., Mandrioli, M., Cherif, A., Bandi, C., Daffonchio, D. (2010) Acetic acid bacteria, newly emerging symbionts of insects. Applied and Environmental Microbiology 76:6963–6970

Cuthbertson L, Mainprize, I.L., Naismith, J. H., Whitfield, C. (2009) Pivotal roles of the outer membrane polysaccharide export and polysaccharide copolymerase protein families in export of extracellular polysaccharides in gram-negative bacteria. Microbiology and molecular biology reviews : MMBR 73(1):155-177

De Ruiter GA, Schols HA, Voragen AGJ, Rombouts FM (1992) Carbohydrate analysis of water-soluble uronic acid-containing polysaccharides with high-performance anion-exchange chromatography using methanolysis combined with TFA hydrolysis is superior to four other methods. Anal Biochem 207(1):176-185 doi:10.1016/0003-2697(92)90520-h

Deeraksa A, Moonmangmee S, Toyama H, Adachi O, Matsushita K (2006) Conversion of capsular polysaccharide, involved in pellicle formation, to extracellular polysaccharide by galE deletion in *Acetobacter tropicalis*. Bioscience, Biotechnology, and Biochemistry 70(10):2536-9

Deeraksa A, Moonmangmee S, Toyama H, Yamada M, Adachi O, Matsushita K (2005) Characterization and spontaneous mutation of a novel gene, polE, involved in pellicle formation in Acetobacter tropicalis SKU1100. Microbiology 151(Pt 12):4111-20 doi:10.1099/mic.0.28350-0

Deppenmeier U, Ehrenreich A (2009) Physiology of acetic acid bacteria in light of the genome sequence of Gluconobacter oxydans. Journal of molecular microbiology and biotechnology 16(1-2):69-80 doi:10.1159/000142895

Deppenmeier U, Ehrenreich, A. (2009) Physiology of acetic acid bacteria in light of the genome sequence of *Gluconobacter oxydans.* Journal of molecular microbiology and biotechnology 16(1):69-80

Deppenmeier U, Hoffmeister, M., Prust, C. (2002) Biochemistry and biotechnological applications of Gluconobacter strains. Applied Microbiology and Biotechnology 60:233–242

Eid J, Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy, D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J., Turner, S. (2009) Real-time dna sequencing from single polymerase molecules. Science 5910(323):133- 138

Foster PL (2006) Methods for determining spontaneous mutation rates. Methods in Enzymology 409:195–213

Franco AV, Liu, D., Reeves, P. R. (1998) The Wzz (Cld) protein in *Escherichia coli:* Amino acid sequence variation determines o-antigen chain length specificity. Journal Of Bacteriology 180(10):2670-2675

Franke IH, Fegan, M., Hayward, C., Leonard, G., Stackebrandt, E., Sly, L.I. (1999) Description of *Gluconacetobacter sacchari sp. nov.*, a new species of acetic acid bacterium isolated from the leaf sheath of sugar cane and from the pink sugarcane mealy bug. International Journal of System Bacteriology 49:1681–1693

Geremia RA, Roux, M., Ferreiro, D. U., Dauphin-Dubois, R., Lellouch, A. C., Ielpi, L. (1999) Expression and biochemical characterisation of recombinant AceA, a bacterial alpha-mannosyltransferase. Molecular and General Genetics 261(6):933-940

Gottlieb HE, Kotlyar V, Nudelman A (1997) NMR chemical shifts of common laboratory solvents as trace impurities. J Org Chem 62(21):7512-7515 doi:10.1021/jo971176v

Griffin AM, Morris, V. J., Gasson, M. J. (1994) Genetic analysis of the acetan biosynthetic pathway in *Acetobacter xylinum.* International journal of biological macromolecules 16(6):287-289

Hall PE, Anderson M, Johnston DM, Cannons RE (1992) Transformation of *Acetobacter xylinus* with plasmid DNA by electroporation Plasmid 28(3):194-200.

Hanahan D (1983) Studies on transformation of Escherichia coli with plasmids. Journal of Molecular Biology 166(4):557-580

Hehre EJ (1951) The biological synthesis of dextran from dextrins. Journal of Biological Chemistry 192:161–174

Hekmat D, Bauer, R. & Fricke, J. (2003) Optimization of the microbial synthesis of dihydroxyacetone from glycerol with Gluconobacter oxydans. Bioprocess Biosyst Eng 26:109–116

Hernandez L, Arrieta J, Menendez C, Vazquez R, Coego A, Suarez V, Selman G, Petit-Glatron MF, Chambert R (1995) Isolation and enzymic properties of levansucrase secreted by Acetobacter diazotrophicus SRT4, a bacterium associated with sugar cane. The Biochemical journal 309:113-118

Ielmini N, Katzen, F., Ferreiro, D. U., Oddo, C. G., Vero, M., Becker, A., Pu, A. (1998) *Xanthomonas campestris pv . campestris* gum mutants : Effects on xanthan biosynthesis and plant virulence. Journal of Bacteriology 180(7):1607–1617

Ielpi L, Couso, R. O., Dankert, M. A. (1993) Sequential assembly and polymerization of the polyprenol-linked pentasaccharide repeating unit of the xanthan polysaccharide in *Xanthomonas campestris*. Journal of Bacteriology 175(9):2490-2500

Ishida T, Sugano, Y., Shoda, M. (2002a) Effects of acetan on production of bacterial cellulose by *Acetobacter xylinum.* Bioscience, biotechnology, and biochemistry 66(8):1677-81

Ishida T, Sugano, Y., Shoda, M. (2002b) Novel glycosyltransferase genes involved in the acetan biosynthesis of *Acetobacter xylinum*. Biochemical and Biophysical Research Communications 295 (2):230–235

Iyer PRea (2010) Genome sequence of a cellulose-producing bacterium, *Gluconacetobacter hansenii* ATCC 23769. Journal of Bacteriology 192:4256-4257

Jakob F (2012a) Mikrobiologie Novel fructans from acetic acid bacteria.

Jakob F (2012b) Novel fructans from acetic acid bacteria. Technische Universität München

Jakob F, Meißner, D., Vogel, R. F. (2012c) Comparison of novel GH 68 levansucrases of levan-overproducing *Gluconobacter species*. Acetic Acid Bacteria 1(1):2

Jakob F, Pfaff, A., Novoa-Carballal, R., Rübsam, H., Becker, T., Vogel, R. F. (2013) Structural analysis of fructans produced by acetic acid bacteria reveals a relation to hydrocolloid function. Carbohydrate Polymers 92(2):1234– 1242

Jansson P, Kenne, L., Lindberg, B. (1975) Structure of the extracellular polysaccharide from *Xanthomonas campestris*. Carbohydrate Research 45(1):275-282.

Jansson P, Lindberg J, Wimalasiri KMS, Dankert MA (1993) Structural studies of acetan, an exopolysaccharide elaborated by *Acetobacter xylinum*. Carbohydrate Research 245:303-310

Kaditzky SV, Rudi F. (2008) Optimization of exopolysaccharide yields in sourdoughs fermented by lactobacilli. European Food Research and Technology 228:291–299

Kanchanarach W, Theeragool G, Inoue T, Yakushi T, Adachi O, Matsushita K (2010) Acetic acid fermentation of *Acetobacter pasteurianus* : Relationship between acetic acid resistance and pellicle polysaccharide formation. Bioscience, Biotechnology, and Biochemistry 74(8):1591-1597

Kanchanarach W, Theeragool, G., Inoue, T., Yakushi, T., Adachi, O. & Matsushita K. (2014) Acetic acid fermentation of *Acetobacter pasteurianus:* Relationship between acetic acid resistance and pellicle polysaccharide formation. Bioscience, Biotechnology, and Biochemistry 74(8):1591-1597

Kappock TJea (2014) Draft Genome Sequence of Acetobacter aceti Strain 1023, a Vinigar Factory Isolate. Genome Announvements 3(2)

Kondo K, Horinouchi, S. (1997) Characterization of an insertion sequence , IS 12528 , from *Gluconobacter suboxydans*. Applied and Environmental Microbiology 63(3):1139-1142

Kornmann H, Duboc P, Marison I, Stockar U (2003) Influence of nutritional factors on the nature , yield , and composition of exopolysaccharides produced by *Gluconacetobacter xylinus* I-2281. Applied and Environmental Microbiology 69(10):6091-6098

Krajewski V, Simić, P., Mouncey, N. J., Bringer, S., Sahm, H., Bott, M. (2010) Metabolic engineering of *Gluconobacter oxydans* for improved growth rate and growth yield on glucose by elimination of gluconate formation. Applied and Environmental Microbiology 76(13):4369-4376

Kretschmer PJ, Cohen SN (1979) Effect of temperature on translocation frequency of the Tn3 element. Journal of Bacteriology 139(2):515-9.

Krumsiek J, Arnold, R., Rattei, T. (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics 23(8):1026-1028

Kubiak K, Kurzawa, M., Ludwicka, K., Krawczyk, M., Migdalski, A., Kacprzak, M. M., Loska, D., Krystynowicz, A., Bielecki, S. (2014) Complete genome sequence of *Gluconacetobacter xylinus* E25 strain -- valuable and effective producer of bacterial nanocellulose Journal of Biotechnology 176:9-18

Lamrani S, Ranquet C, Gama MJ, Nakai H, Shapiro Ja, Toussaint A, Maenhaut-Michel G (1999) Starvation-induced Mucts62-mediated coding sequence fusion: a role for ClpXP, Lon, RpoS and Crp. Molecular microbiology 32(2):327-43.

Langlotza C, Schollmeyera, M., Coplinb, D. L. ,Nimtzc, M. , Geider, K. (2011) Biosynthesis of the repeating units of the exopolysaccharides amylovoran from *Erwinia amylovora* and stewartan from *Pantoea stewartii*. Physiological and Molecular Plant Pathology 75(4):163-169

Leontein K, Lindberg B, Lonngren J (1978) Assignment of absolute configuration of sugars by GLC of their acetylated glycosides formed from chiral alcohols. Carbohydr Res 62(2):359-362 doi:10.1016/s0008-6215(00)80882-4

Lisdiyanti P (2002) Kozakia baliensis gen. nov., sp. nov., a novel acetic acid bacterium in the alpha-Proteobacteria. International Journal of Systematic and Evolutionary Microbiology 52(3):813-818 doi:10.1099/ijs.0.01982-0

Marolda CL VM (1995) Genetic analysis of the dTDP-rhamnose biosynthesis region of the *Escherichia coli* VW187 (O7:K1) rfb gene cluster: identification of functional homologs of rfbB and rfbA in the rff cluster and correct location of the rffE gene. J Bacteriol 177(19):5539-46.

Matsushita K, Fujii, Y., Ano, Y., Toyama, H., Shinjoh, M., Tomiyama, N., Miyazaki, T., Sugisawa, T., Hoshino, T., Adachi, O. (2003) 5-Keto- D-gluconate production is catalyzed by a quinoprotein glycerol dehydrogenase, major polyol dehydrogenase in Gluconobacter species. Applied Environmental Microbiology 69:1959–1966.

Matsushita K, Inoue, T., Adachi, O., Toyama, H. (2005) Acetobacter aceti possesses a proton motive force-dependent efflux system for acetic acid. Journal of Bacteriology 187(13):4346-4352

Matsushita K, Toyoma, H., Adachi, O. (1994) Respiratory chains and bioenergetics of acetic acid bacteria. Advances in Microbial Physiology 36:247–301

McCarthy A (2010) Third generation DNA sequencing: pacific biosciences' single molecule real time technology. Chemistry & Biology 17(7):675–676

Meneses CHSG, Rouws, L. F. M., Simões-araújo, J. L., Vidal, M. S., Baldani, J. I. (2011) Exopolysaccharide production is required for biofilm formation and plant colonization by the nitrogen-fixing endophyte *Gluconacetobacter diazotrophicus*. Molecular Plant-Microbe Interactions 24(12):1448-1458

Merighi M, Lee, V. T., Hyodo, M., Hayakawa, Y., Lory, S. (2007) The second messenger bis- (3 `-5`) -cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. Molecular Microbiology 65(4):876-895

Mont Â, Bazin C, Capy P, Gasperi G, Bie C (2000) Stress and transposable elements : co-evolution or useful parasites ? Heredity 82(2):101–106

Moonmangmee S, Kawabata, K., Tanaka, S., Toyama, H., Adachi, O., Matsushita, K. (2002a) A novel polysaccharide involved in the pellicle formation of *Acetobacter aceti.* Journal of Bioscience and Bioengineering 93(2):192–200

Moonmangmee S, Toyama, H., Adachi, O., Teerakool, G., Lotong, N. & Matsushita, K. (2002b) Purification and characterization of a novel polysaccharide involved in the pellicle produced by a thermotolerant *Acetobacter* strain. Bioscience, biotechnology, and biochemistry 66(4):777-783

Morgan JLW, McNamara, J.T. & Zimmer J. (2014) Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. Nature Structural & Molecular Biology 21(5):489-96

Mostafa HE, Heller KJ, Geis A (2002) Cloning of *Escherichia coli* lacZ and lacY Genes and Their Expression in *Gluconobacter oxydans* and *Acetobacter liquefaciens*. Applied and Environmental Microbiology 68(5):2619-2623

Mullins E, Kappock, T.J. (2013) Functional analysis of the acetic acid resistance (aar) gene cluster in Acetobacter aceti strain 1023. Acetic Acid Bacteria 2(1s):9-18

Naessens M, Cerdobbel, A., Soetaert, W., Vandamme E.J. (2005) Dextran dextrinase and dextran of Gluconobacter oxydans. Journal of Industrial Microbiology and Biotechnology 32:323-334

Nunes FM, Reis A, Silva AMS, Domingues MRM, Coimbra MA (2008) Rhamnoarabinosyl and rhamnoarabinoarabinosyl side chains as structural features of coffee arabinogalactans. Phytochemistry 69(7):1573-1585 doi:10.1016/j.phytochem.2008.01.021

Nwodo UU, Green E, Okoh AI (2012) Bacterial exopolysaccharides: functionality and prospects. International journal of molecular sciences 13:14002-14015

O'Shea TM, Deloney-Marino CR, Shibata S, Aizawa S, Wolfe AJ, Visick KL (2005) Magnesium promotes flagellation of *Vibrio fischeri*. Journal of Bacteriology 187(6):2058-65 doi:10.1128/JB.187.6.2058-2065.2005

O'Shea TM, Klein, A. H., Geszvain, K., Wolfe, A. J., Visick, K. L. (2006) Diguanylate cyclases control magnesium-dependent motility of *Vibrio fischeri.* Journal of Bacteriology 188(23):8196-205

Oglesby LL, Jain, S., Ohman, D.E. (2008) Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization. Microbiology 154(6):1605-1615

Ohmori S, Uozumi, T., Beppu, T. (1981) Loss of acetic acid resistance and ethanol oxidizing ability in an *Acetobacter* strain. Agricultural and Biological Chemistry 46(2):381-389

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 research:1-9
Pedraza RO (2008) Recent advances in nitrogen-fixing acetic acid bacteria. International journal of food microbiology 125:25-35
Perumpuli PABN, Watanabe T, Toyama H (2014) Pellicle of thermotolerant *Acetobacter pasteurianus* strains: characterization of the polysaccharides and of the induction patterns. Journal of bioscience and bioengineering 118(2):134-138
Peters B, Junker, A., Brauer, K., Mühlthaler, B., Kostner, D., Mientus, M., Liebl, W., Ehrenreich, A. (2013) Deletion of pyruvate decarboxylase by a new method for efficient markerless gene deletions in *Gluconobacter oxydans.* Applied microbiology and biotechnology 97(6):2521-30
Petroni EA, Ielpi L (1996) Isolation and nucleotide sequence of the GDP-mannose : Cellobiosyl- gene from *Acetobacter xylinum*. Journal of Bacteriology 178(16):4814-4821
Pettolino FA, Walsh C, Fincher GB, Bacic A (2012) Determining the polysaccharide composition of plant cell walls. Nature Protocols 7(9):1590-1607 doi:10.1038/nprot.2012.081
Prust C, Hoffmeister M, Liesegang H, Wiezer A, Fricke WF, Ehrenreich A, Gottschalk G, Deppenmeier U (2005) Complete genome sequence of the acetic acid bacterium Gluconobacter oxydans. Nature biotechnology 23(2):195-200 doi:10.1038/nbt1062
Pühler A, Vorhölter FJ, Schneiker S, Goesmann A, Krause L, Bekel T, Kaiser O, Linke B, Patschkowski T, Rückert C, Schmid J, Sidhu VK, Sieber V, Tauch A, Watt SA, Weisshaar B, Becker A, Niehaus K (2008) The genome of Xanthomonas campestris pv. campestris B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. Journal of biotechnology 134 33–45
Raspor P, Goranovic, D. (2008) Biotechnological applications of acetic acid bacteria. Critical Reviews in Biotechnology 28(2):101–124
Rehm BHA (2010) Bacterial polymers: biosynthesis, modifications and applications. Nature reviews Microbiology 8:578-592
Roh SW, Nam, Y.D., Chang, H.W., Kim, K.H., Kim, M.S., Ryu, J.H., Kim, S.H., Lee, W.J., Bae, J.W. (2008) Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in Drosophila melanogaster. Applied and Environmental Microbiology 74:6171-6177
Römling UaG, M. Y. (2005) Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends in microbiology 23(9)
Rosalam S, England R (2006) Review of xanthan gum production from unmodified starches by Xanthomonas comprestris sp. Enzyme and Microbial Technology 39:197–207
Ross P, Benziman, M., Vroom, E. , Fidder, A., van Boom, J. H. (1990) The cyclic diguanylic acid regulatory in *Acetobacter xylinum* system of cellulose synthesis. Biochemistry 265(31):18933-18943
Ross P WH, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, de Vroom E, van der Marel GA, van Boom JH, Benziman M. (1987) Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. Nature 325
Ruiz A, Poblet, M., Mas, A., Guillamon, J.M. (2000) Identification of acetic acid bacteria by RFLP of PCR-amplified 16s rDNA and 16s-23s DNA intergenic spacer. International Journal of Evolutionary Microbiology 50:1981–1987

Sambrook J, Fritsch, E. F. & Maniatis, T. (1989) Molecular Cloning: a Laboratory Manual, 2nd edn. Cold Spring Harbor NY: Cold Spring Harbor Laboratory

Saxena IM, Kudlicka, K., Okuda, K., Brown, R. M. (1994) Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. Journal of bacteriology 176(18):5735-52

Schmid J, Koenig S, Rühmann B, Rütering M (2014) Biosynthese und Genomik mikrobieller Polysaccharide. BIOspektrum 20(3):288–290

Schmid J, Koenig, S., Rühmann, B., Rütering, M. (2014) Biosynthese und genomik mikrobieller polysaccharide. BioSpektrum 20(3):288-290

Schmid J, Sieber, V. (2015) Enzymatic transformations involved in the biosynthesis of microbial exo-polysaccharides based on the assembly of repeat units. ChemBioChem 16(8):1141–1147

Serrato RV, Meneses CH, Vidal MS, Santana-Filho AP, Iacomini M, Sassaki GL, Baldani JI (2013) Structural studies of an exopolysaccharide produced by *Gluconacetobacter diazotrophicus* Pal5. Carbohydrate Polymers 98(1):1153-9 doi:10.1016/j.carbpol.2013.07.025

Sweet DP, Shapiro RH, Albersheim P (1975) Quantitative-analysis by various GLC response-factor theories for partially methylated and partially ethylated alditol acetates. Carbohydr Res 40(2):217-225 doi:10.1016/s0008-6215(00)82604-x

Takemura H, Horinouchi, S., Beppu, T. (1991) Novel insertion-sequence IS1380 from *Acetobacter pasteurianus* is involved in loss of ethanol-oxidizing ability. Journal of Bacteriology 173(22):7070-7076

Tao F, Wang X, Ma C, Yang C, Tang H, Gai Z, Xu P (2012) Genome sequence of Xanthomonas campestris JX, an industrially productive strain for Xanthan gum. Journal of bacteriology 17:4755–4756

Tieking M, Kaditzky, S., Valcheva, R., Korakli, M., Vogel, R.F., Gänzle, M.G. (2005) Extracellular homopolysaccharides and oligosaccharides from intestinal lactobacilli. Journal of Applied Microbiology 99:692-702

Tillett D, Neilan, B. (1999) Enzyme-free cloning: a rapidmethod to clone PCR products independent of vector restriction enzyme sites. Oxford university press 27(19):e26

Toda K, Asakura, T., Fukaya,M., Entani,E.,Kawamura, Y. (1997) Cellulose production by acetic acid-resistant *Acetobacter xylinum*. Journal of Fermentation and Bioengineering 84(3):228-231

Ua-Arak T, Jakob, F., Vogel, R. F. (2015) Characterization of growth and exopolysaccharide production of selected acetic acid bacteria in buckwheat sourdoughs. International journal of food microbiology 239:103-112

Vargas-Garcıa M, Lopez, MJ , MA Elorrieta, F Suarez and J Moreno (2001) Influence of nutritional and environmental factors on polysaccharide production by *Azotobacter vinelandii* cultured on 4-hydroxybenzoic acid. Journal of industrial microbiology & biotechnology 27(1):5-10

Vorhölter FJ, Schneiker, S., Goesmann, A, Krause, L., Bekel, T., Kaiser, O., Linke, B., Patschkowski, T., Rückert, C., Schmid, J., Sidhu, V.K., Sieber, V., Tauch, A., Watt, S.A., Weisshaar, B., Becker, A., Niehaus, K., Pühler, A. (2008) The genome of *Xanthomonas campestris pv. campestris* B100 and its use for the reconstruction of metabolic pathways involved in xanthan biosynthesis. Journal of Biotechnology 134(1):33-45

Wach A (1996) PCR-synthesis of marker cassettes with long flanking homology regions for gene disruptions in *S . cerevisiae*. Yeast 12(3):259-265

Wang X, Yang, F. and von Bodman, S. B. (2012) The genetic and structural basis of two distinct terminal side branch residues in stewartan and amylovoran exopolysaccharides and their potential role in host adaptation. Molecular Microbiology 83(1):195-207

Wefers D, Bunzel M (2015) Characterization of dietary fiber polysaccharides from dehulled, common buckwheat (*Fagopyrum esculentum*) seeds. Cereal Chem 92(6):598-603

Wessler SR (1996) Plant retrotransposons: Turned on by stress. Current Biology 8(6):959-61

Whitney JC, Howell PL (2013) Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. Trends in Microbiology 21(2):63-72

Wong HC, Fear AL, Calhoon RD, Eichinger GH, Mayer R, Amikam D, Benziman M, Gelfand DH, Meade JH, Emerick aW (1990) Genetic organization of the cellulose synthase operon in Acetobacter xylinum. Proceedings of the National Academy of Sciences of the United States of America 87:8130-8134

Yamada Y, Hoshino, K, Ishikawa, T (1997) The phylogeny of acetic acid bacteria based on the partial sequences of 16S ribosomal RNA: the elevation of the subgenus *Gluconacetobacter* to the generic level. Bioscience, Biotechnolog and Biochemistry 61:1244–1251

Yamada Y, Pattaraporn, Y., Vu, L., Thi, H., Yuki, M., Duangjai, O., Somboon, T., Yasuyoshi, N. (2012) Description of *Komagataeibacter gen. nov.*, with proposals of new combinations (Acetobacteraceae). The Journal of general and applied microbiology 58(5):397-404

Yamada YY, P. (2008) Genera and species in acetic acid bacteria. International journal of food microbiology 125:15-24