Mali Mærk

Looking for the Big Picture: Genome-Based Approaches to Improve Alginate Production in *Azotobacter vinelandii*

Thesis for the degree of Philosophiae Doctor

Trondheim, January 2014

Norwegian University of Science and Technology
Faculty of Natural Sciences and Technology
Department of Biotechnology

NTNU – Trondheim
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Natural Sciences and Technology
Department of Biotechnology

© Mali Mærk

ISBN 978-82-471-4939-3 (printed ver.)
ISBN 978-82-471-4940-9 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2014:18

Printed by NTNU-trykk

Things you thought impossible before breakfast just might be unavoidable by lunchtime.

- Rebecca Solnit

ACKNOWLEDGEMENTS

Although it might feel that way when you're stuck alone in the lab or office at ungodly hours, doing a PhD is generally not a solo run. Mine was no exception, and I am very grateful to a lot of people for their advice, assistance and good company during my years as a PhD student.

First of all, a big thanks to my supervisors Svein Valla and Helga Ertesvåg at the Department of Biotechnology, NTNU. Your considerable experience, insightful questions and invaluable advice have been of great importance for my research and thesis writing. Thank you for encouraging me, and for the interesting discussions about both research and other topics. I also want to thank my collaborators at SINTEF Materials and Chemistry, at Department of Cancer Research and Molecular Medicine (NTNU) and in the *Azotobacter vinelandii* genome project team. Without their expertise and efforts major parts of the work presented in this thesis would not have been possible. Thanks to all my colleagues at Department of Biotechnology for providing a friendly and helpful work environment, with an extra thanks to the department's excellent technicians for keeping the labs running smoothly.

I've been lucky enough to find some truly great friends among the people I have worked and studied with. Anna, Marte, Ane Kjersti, Marianne, Claudia, Kerstin, Hanne, Torunn, Stina, Susan, Marie and Vu: You rock. Thanks for all the coffee breaks, the shared frustrations and the pep-talks, and for making my life more fun both on and off campus. A big thanks to my outside-of-biotech friends for being the amazing people you are and reminding me that there are other interesting things in this world besides bacteria. I'll add a special thanks to Tommy for being my statistics (and other) helpline, to Egil for the office SPKs, and to Marte and Anna for keeping me sane through the final months of thesis writing.

Completing this PhD project took longer than expected, so for the last few years I have been working at the Centre for Renewable Energy (SFFE), and thereafter at the Norwegian Defence Research Establishment (FFI), while finishing my PhD in my spare

time. I am therefore very grateful to Gabriella Tranell (SFFE) and Janet M. Blatny (FFI) for their encouragement and understanding, and for allowing me the flexibility I sometimes needed to see this through. Thank you!

Finally, I want to thank my awesome family for always believing in me and having my back, especially mamma, pappa, Marius, Beate, Madelen and Bastian. You mean the world to me, and I'll be visiting Torshaug more often now that this is done. ☺ And Per, thank you for all the love and support, for bearing with me, and for the "meals on wheels". I look forward to spending more time together.

Trondheim, December 2013

Mali Mærk

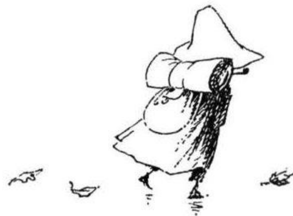


Illustration created by Tove Jansson and published in her book "Sent i november".
Reproduced with permission of the copyright owner. ©Moomin Characters™

ABSTRACT

Alginates are commercially important polysaccharides with a wide range of industrial and technological applications. Polymer chain length and monomer distribution greatly affect the material properties, which makes different alginate types ideal for different areas of use. All commercial alginate manufacture is currently based on extraction from brown algae, but the polymers are also produced by bacteria in the genera *Pseudomonas* and *Azotobacter*. Bacterial bioproduction is technically possible, but is not yet economically competitive with algal alginates. *A. vinelandii* is an attractive candidate for development of bacterial bioproduction strains due to its potential for producing homogenous alginates with tailored monomer compositions, and thus high market value. Successful development of strains and cultivation conditions for bioproduction is however dependent on extensive knowledge of the factors affecting the biosynthetic process in question.

Along with the availability of complete genome sequences, a multitude of new opportunities has arised with regard to investigations of gene functions, metabolic and regulatory relationships, environmental adaptations etc. Part of the work presented in this thesis is a contribution to the manual curation and annotation of the first published *A. vinelandii* genome, which provided the basis for our subsequent genome-based analyses of factors affecting alginate production in this organism. The genome annotation and analysis carried out as part of this thesis was mainly concerned with carbohydrate metabolism genes, and has provided a foundation for further investigations of sugar uptake and utilization in this organism. A notable outcome of the genome analysis was the discovery of numerous highly similar and apparently conserved intra-genome homologs among *A. vinelandii* core carbohydrate metabolism genes. Investigations of 943 bacterial and archaeal genomes confirmed that the number of such homologs is indeed unusually large in *A. vinelandii*. We propose that the retention of multiple gene copies confers adaptive benefits via gene dosage and/or increased regulatory flexibility.

Genes, and thus cellular processes, affecting alginate production in *A. vinelandii* were investigated by construction and screening of a transposon insertion library comprising 4000 mutant strains. Abolished or diminished alginate production was confirmed for ~70 transposon insertion mutants and the disrupted genes were identified by sequencing. The disrupted genes included structural and regulatory genes involved in alginate biosynthesis, as well as genes involved in iron uptake, peptidoglycan recycling, motility and synthesis of several cofactors and central metabolites. Based on these results the effect of various medium supplements on alginate production in wild type *A. vinelandii* was investigated, and addition of thiamine, succinate or a mixture of lysine, methionine and diaminopimelate was shown to result in significantly increased alginate levels. The screening results also revealed two possible new regulators of alginate biosynthesis; the fructose phosphotransferase system protein FruA and an IclR family transcriptional regulator. Two mutants were confirmed to have gained an increase in alginate production. For one of these the disrupted gene was identified as *mucA*, encoding the main negative regulator of alginate biosynthesis.

Global effects of inactivating MucA were investigated by phenotypic characterization and transcriptome analyses of fermentor-grown *A. vinelandii* wild type and *mucA*$^-$ strains. The *mucA* mutant has a lowered growth rate, elevated alginate production and diminished respiration rate compared to the wild type strain. Both medium composition and MucA inactivation had profound effects on carbon source utilization. The transcriptome analyses revealed new roles for the key regulators MucA/AlgU with regard to control of alginate composition, cell mass production, respiration and possibly nitrogen fixation. The redirection of carbon utilization in the *mucA* mutant was also reflected in transcriptional changes in genes involved in gluconeogenesis/glycolysis and energy production.

The results presented in this thesis will have importance for further work towards the long-term goal of establishing bacterial systems for commercial bioproduction of alginates.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS iii
ABSTRACT v
LIST OF PAPERS ix
1 INTRODUCTION 1
1.1 Alginate 1
1.1.1 Occurrence and functions in nature 1
1.1.2 Composition, structure and material properties 1
1.1.3 Industrial and technological applications 3
1.1.4 Bioproduction 4
1.2 Bacterial biosynthesis of alginates 5
1.2.1 Precursors and biosynthetic machinery 5
1.2.2 Regulation 8
1.3 *Azotobacter vinelandii* – biology and taxonomy 11
1.4 Genome-wide analyses 12
1.4.1 Comparative genomics 12
1.4.2 Functional genomics and high-throughput screening 13
2 AIMS OF THE STUDY 14
3 SUMMARY OF RESULTS AND DISCUSSION 15
3.1 *A. vinelandii* genome annotation and analysis 15
3.1.1 Summary of genome annotation and analysis 15
3.1.2 Reconstruction of central carbohydrate metabolism 17
3.1.3 *A. vinelandii* and *Pseudomonas* spp. glucose 6-phosphate dehydrogenase genes 19
3.1.4 Occurrence of highly similar carbohydrate metabolism synologs 21
3.2 Genome-wide studies of factors affecting *A. vinelandii* alginate biosynthesis . 24
3.2.1 Identification of genes affecting alginate biosynthesis 24
3.2.2 Medium supplements increasing alginate production 27
3.2.3 Phenotypic characterization of wild type and *mucA*⁻ strains 28
3.2.4 Comparative transcriptome analyses of wild type and *mucA*⁻ strains 29
4 CONCLUDING REMARKS 33
REFERENCES 35

LIST OF PAPERS

Paper I

Setubal JC, dos Santos P, Goldman BS, Ertesvåg H, Espín G, Rubio LM, Valla S, Almeida NF, Balasubramanian D, Cromes L, Curatti L, Du Z, Godsy E, Goodner B, Hellner-Burris K, Hernandez JA, Houmiel K, Imperial J, Kennedy C, Larson TJ, Latreille P, Ligon LS, Lu J, Mærk M, Miller NM, Norton S, O'Carroll IP, Paulsen I, Raulfs EC, Roemer R, Rosser J, Segura D, Slater S, Stricklin SL, Studholme DJ, Sun J, Viana CJ, Wallin E, Wang B, Wheeler C, Zhu H, Dean DR, Dixon R, Wood D (2009) Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. *J Bacteriol* **191(14)**:4534-45

Paper II

Mærk M, Johansen J, Ertesvåg H, Drabløs F, Valla S. Safety in numbers: multiple occurrences of highly similar homologs among *Azotobacter vinelandii* carbohydrate metabolism proteins probably confer adaptive benefits. *Submitted*

Paper III

Mærk M, Jakobsen ØM, Sletta H, Klinkenberg G , Tøndervik A, Ellingsen TE, Valla S, Ertesvåg H. Identification of genes and metabolic processes important for alginate biosynthesis in *Azotobacter vinelandii* by screening of a transposon insertion mutant library. *Manuscript*

Paper IV

Sletta H, Mærk M, Jakobsen ØM, Moreno S, Tøndervik A, Espín G, Ellingsen TE, Valla S, Ertesvåg H. Inactivation of the anti-sigma factor MucA causes major perturbations in carbon metabolism in *Azotobacter vinelandii*. *Manuscript*

Contributions to other publications not included in this thesis:

Steigedal M, Sletta H, Moreno S, Mærk M, Christensen BE, Bjerkan T, Ellingsen TE, Espín G, Ertesvåg H, Valla S (2008) The *Azotobacter vinelandii* AlgE mannuronan C-5-epimerase family is essential for the in vivo control of alginate monomer composition and for functional cyst formation. *Environ Microbiol* **10(7)**:1760-70

1 INTRODUCTION

1.1 Alginate

1.1.1 Occurrence and functions in nature

The polysaccharides known as alginates (first named algin or algic acid) were first discovered in marine brown algae (1), where they play a prominent structural role. Alginates can comprise up to 40% of the dry matter of brown algae and are located in the intercellular matrix as a gel containing divalent cations (2). Alginates are also synthesized and secreted by several bacteria in the genera *Azotobacter* and *Pseudomonas* (3-11). The soil bacterium *A. vinelandii* is able to form dormant, dessication-resistant cysts, in which alginate is an important component of the cyst coat (12). *A. vinelandii* also produces alginate when growing vegetatively, during which the extracellular alginate capsule is assumed to have a role in protecting the oxygen-sensitive nitrogenases used by this bacterium to fix atmospheric nitrogen (13). Unlike *A. vinelandii*, the *Pseudomonas* species investigated so far do not produce alginate constitutively. The role of alginate appears to vary for different pseudomonads, where the polysaccharide has been implicated in biofilm formation, virulence, antibiotic resistance and maintaining cellular hydration (14).

1.1.2 Composition, structure and material properties

Alginates comprise a family of unbranched (1-4)-linked copolymers consisting of the monosaccharides β-D-mannuronic acid (M) and its C5-epimer α-L-guluronic acid (G). The monomers in an alginate chain are arranged in stretches of alternating sequence (MG/GM-blocks) and/or stretches consisting of only one monomer (M-blocks and G-blocks) (15). Alginate chemical structure is shown in Figure 1. In bacterial alginates some of the M monomers may be *O*-acetylated, which inhibits epimerization and alginate lyase-catalyzed degradation (16, 17) and thus influences monomer distribution and chain length. Block length and distribution as well as the overall length of alginate chains vary widely, even for alginates isolated from the same material. G-blocks have not been observed in bacterial alginates from the species *Pseudomonas*,

even though *P. syringae* has been shown to contain a gene for an epimerase that can introduce G-blocks *in vitro* (18). Alginates consisting of pure polymannuronan can be isolated from *Pseudomonas* strains in which the mannuronan C5-epimerase-encoding *algG* gene has been inactivated (19, 20).

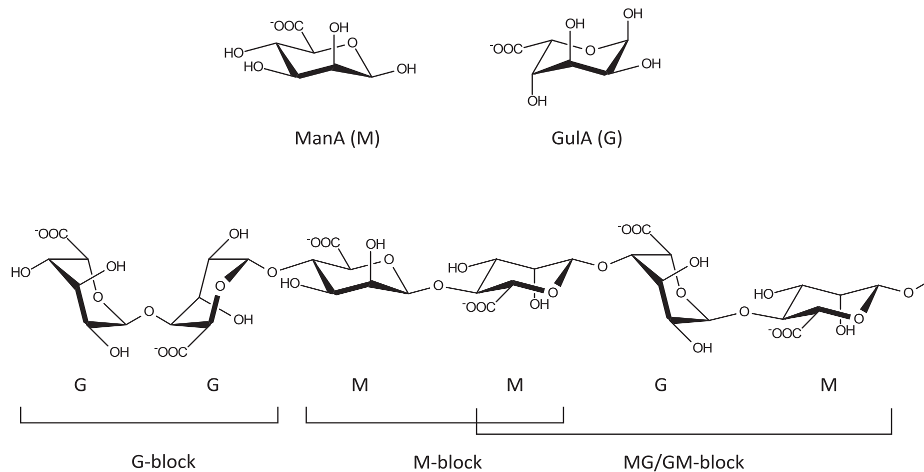


Figure 1: Chemical structure of alginate. The alginate monomers β-D-mannuronic acid (ManA; M) and α-L-guluronic acid (GulA; G), as well as an alginate chain illustrating linkage conformations and block composition. Alginates may contain G-blocks, M-blocks and/or MG/GM-blocks of varying lengths.

Important properties of alginates are their ability to form viscous solutions and gels. These properties are largely determined by the molecular mass, monomer composition and sequential structure of the alginate chains (21). The intrinsic viscosity of alginates is mainly determined by chain length, but is also affected by the pH and ionic strength of the solution (22) since alginates are polyanions at pH values above 5 (2). Alginates containing G-blocks form gels in the presence of divalent cations, such as Ca^{2+} (23, 24) (Figure 2). The strength and flexibility of ion gels depend on the fraction and lengths of G blocks in the alginate: generally G-rich alginates will form strong, brittle gels, while M-rich alginates will give more flexible gels with lower gel strength (26). Gel strength and the gel's ability to bind water is also influenced by the degree of acetylation (27). In addition to ion gels, alginates may also form acid gels when the pH is brought below the pK_a value of the uronic acid monomers. Both G-blocks and M-blocks form acid gels, but the former provides the strongest gels (28).

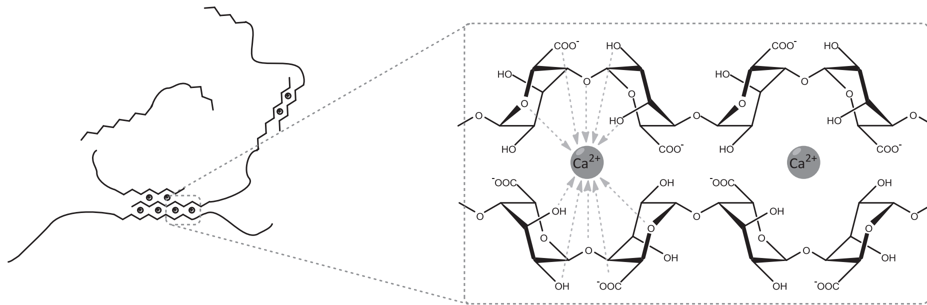


Figure 2: The "egg-box" model (25)**.** This model describes the interactions between alginate G-blocks and divalent cations (here exemplified by Ca2+), which result in ionic gel formation.

1.1.3 Industrial and technological applications

The material properties of alginates make them suitable for a wide range of applications. Examples of industrial applications include using alginate as a thickener in dye pastes for textile printing, as a plasticizer, stabilizer and binding agent in welding rods, as a co-flocculant in water purification and to bind water in manufacturing of ceramics. In food technology alginates are used as thickeners, stabilizers and emulsifiers in a variety of products, including ice cream, sauces, dressings and beverages. Furthermore, the gelling properties of alginates are utilized in the production of jams, desserts and restructured foods (29). Well-established pharmaceutical and medical applications of alginates include dental impression materials, immediate- and sustained-release tablets and capsules, wound dressings and formulations for preventing gastric reflux (29, 30). Additionally, the mild conditions for gelling make alginates well suited for encapsulation and immobilization of cells and enzymes; a technique with numerous applications including bioproduction of ethanol, antibiotics and other compounds, and mass production of artificial seeds (31). Alginate is among the most promising materials for encapsulating living cells and tissues for transplantation, where such capsules may protect the transplanted cells from the immune system of the host (32, 33), and is also receiving interest for its potential use in tissue engineering and regenerative medicine (34-36), artificial insemination and in vitro fertilization (37), encapsulation of probiotic bacteria to

increase viability both before and after ingestion (38), polymer-controlled drug-delivery (30), and purification of proteins (39).

1.1.4 Bioproduction

As of today, all commercially produced alginates are isolated from algal sources. However, harvesting alginates from algae yields complex mixtures of alginates with varying composition and molecular mass, depending on species, plant part and growth conditions (2, 40, 41), while the new and emerging medical, pharmaceutical and biotechnological applications for the polymer require high-quality alginates with specific material properties, and mainly alginates with high G contents. This poses an additional challenge, as G-rich algal alginates can mainly be harvested from the stipes of old *Laminaria hyperborea* plants (41), which is a limited resource. With an increasing demand for a stable supply of homogenous alginates with specific compositions, native algal alginates are thus falling short. One possible solution is to increase the G-content in alginates harvested from algae by *in vitro* epimerization using bacterial enzymes (42). Another attractive option is bacterial bioproduction (43), which could make it possible to establish reliable and consistent production of alginates with more well-defined qualities. Furthermore, strain engineering opens up new opportunities for *in vivo* production of tailored alginates with specific properties (44, 45).

A. vinelandii is an attractive candidate for development of bacterial alginate bioproduction strains due to its innate ability to synthesize alginates containing G-blocks (46, 47), a trait which has not been found in any *Pseudomonas* species studied so far. The G-blocks are introduced into the *A. vinelandii* alginate chains after polymerization and export, by a family of extracellular AlgE epimerases which are exported via a type I secretion system (48). Heterologous expression and export of these epimerases in *Pseudomonas* sp. for *in vivo* alginate bioproduction is currently not regarded as feasible, as one would need to introduce a functional secretion system in addition to the enzymes themselves. Intracellular epimerization is not an option, as

the alginates are not accessible to the AlgE epimerases until after export (see Chapter 1.2.1).

Research on bacterial bioproduction is ongoing and bacterial alginate production by fermentation is technically possible (49-54), but is not yet economically competitive with algal alginates. This is likely to change as new applications demanding high purity alginates with specific characteristics emerge and enter the market. High-quality alginates have the potential to reach higher market values, thus justifying the higher production costs (45). As an example, ultrapure alginates with defined molecular weights and M/G ratios are currently priced at about 100 USD per gram (NovaMatrix web catalogue prices, September 2013), while the bulk alginates used in food and cosmetic industry can be obtained for as little as 5 USD per kilogram. Additionally, the continued development and optimization of bacterial bioproduction processes can be expected to lower the production costs after commercial production has been established, thus contributing to further competitive advantages.

1.2 Bacterial biosynthesis of alginates

1.2.1 Precursors and biosynthetic machinery

Bacterial biosynthesis of alginates has mainly been studied in *A. vinelandii* and *P. aeruginosa*. Most of the structural genes involved alginate biosynthesis are localized in a gene cluster with the same organization in both these organisms (55-58), as well as in *P. syringae* and *P. fluorescens* (59, 60): *algD*-*alg8*-*alg44*-*algK*-*algJ*(*algE*)-*algG*-*algX*-*algL*-*algI*-*algV*(*algJ*)-*algF*-*algA*. The gene names in brackets indicate alternative nomenclature used in *Pseudomonas* spp. Homologous genes with the same or a similar organization are also found in *P. putida*, *P. entomophila*, *P. mendocina*, *P. denitrificans*, *P. protegens*, *P. resinovorans*, *P. poae*, *P. brassicacearum* and *P. fulva*, but not *P. stutzeri* (61).

The alginate monomers are derived from fructose 6-phosphate, so there is a close connection between alginate biosynthesis and the central carbohydrate metabolism. In pseudomonads glucose is primarily degraded to glyceraldehyde 3-phosphate and

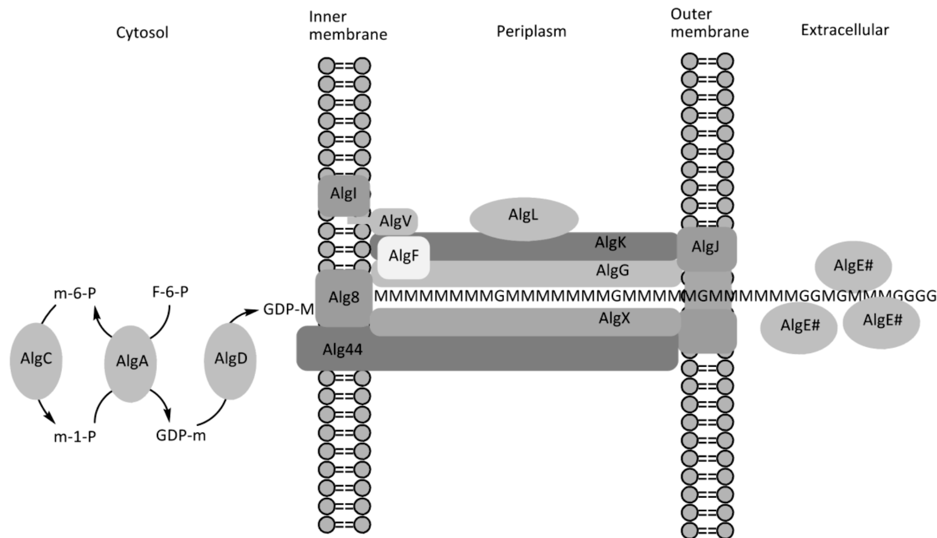


Figure 3: Bacterial alginate biosynthesis. Working model for alginate precursor biosynthesis, polymerization, modification and export, here exemplified by *Azotobacter vinelandii*. See text for details and references. F-6-P: fructose 6-phosphate, m-6-P: mannose 6-phosphate, m-1-P: mannose 1-phosphate, GDP-m: GDP-mannose, GDP-M: GDP-mannuronic acid, M: mannuronic acid monomer, G: guluronic acid monomer. AlgE# symbolizes AlgE1-7. Adapted from (77, 78).

pyruvate via the Entner-Doudoroff (ED) pathway (62). Prior to alginate biosynthesis glucose is converted to fructose 6-phosphate primarily via the ED and gluconeogenetic pathways in both *A. vinelandii* and *P. aeruginosa* (63-67), even though these organisms also have at least one phosphoglucose isomerase which catalyzes the interconversion of glucose 6-phosphate and fructose 6-phosphate. The limited alginate synthesis from directly isomerized glucose 6-phosphate could be due to inhibition of phosphoglucose isomerase by the ED pathway intermediate 6-phosphogluconate (65). Fructose is however channeled to alginate biosynthesis primarily as an intact hexose in both *A. vinelandii* and *P. aeruginosa* (64, 65, 68).

The enzymes necessary for conversion of fructose 6-phosphate to the activated alginate precursor GDP-mannuronic acid (69, 70) are encoded by the first and last genes in the *alg* cluster, *algD* (71, 72) and *algA* (73, 74), plus *algC* which is located outside of the cluster (75, 76). The proteins encoded by the remaining genes in the *alg* cluster are involved in polymerization, export and modification of the nascent alginate

chain. These steps are much more poorly understood, but are believed to be coordinated via the formation of a protein complex involving both inner and outer membrane proteins, as well as periplasmic proteins.

The current model for bacterial alginate biosynthesis from fructose 6-phosphate is shown in Figure 3. Alg8 and Alg44 are membrane proteins required for the polymerization process and transport across the inner membrane (54, 79-83). In the periplasm some of the M monomers are acetylated by a protein complex consisting of AlgI, AlgV (AlgJ in *Pseudomonas* sp.) and AlgF (58, 84-86) and some of the non-acetylated M monomers are epimerized to G monomers by AlgG (56, 87). The alginate-degrading lyase AlgL is also located in the periplasm (88, 89). The biological function of AlgL is not clear, but could be prevention of toxic accumulations of alginate in the periplasm (90, 91). In addition to their enzymatic activities, AlgG and AlgL have been proposed to form a structural complex with the periplasmic scaffold proteins AlgK and AlgX and the outer membrane protein AlgJ (AlgE in *Pseudomonas* spp.) (92). This complex is required for transport of the polymer through the periplasm and across the outer membrane, while at the same time protecting the nascent alginate chain from the lyase activity of AlgL (20, 57, 91, 93-98). It has been suggested that Alg44 might also be a subunit of this complex (92), or function as a physical link between Alg8 in the inner membrane and AlgJ/AlgE in the outer membrane (80, 83). Alginate is most likely secreted from the cell through the pore-forming AlgJ/AlgE protein (93, 98, 99).

In *A. vinelandii* non-acetylated M monomers can also be epimerized after export from the cell, as this organism encodes at least seven secreted mannuronan C5-epimerases (AlgE1-7) in addition to the periplasmic epimerase AlgG (46, 47, 100). With the exception of *algE5*, the genes encoding all these extracellular epimerases are located in the same gene cluster (separate from the *alg* cluster) (101). *P. syringae* also has an additional mannuronan C5-epimerase gene (*psmE*), and the functionality of the corresponding protein has been demonstrated *in vitro* (18). The biological role of PsmE is however not known, and the majority of the sequenced *Pseudomonas* genomes do not contain a *psmE* homolog (61).

1.2.2 Regulation

Alginate biosynthesis is governed by a complex hierarchical network of regulatory controls encompassing global as well as alginate-specific factors, and which includes transcriptional, post-transcriptional and post-translational regulation. The regulation of alginate biosynthesis has mainly been studied in *P. aeruginosa*, and is thus best understood in this organism. Known alginate regulatory genes in *P. aeruginosa* and *A. vinelandii* are listed in Table 1.

In *P. aeruginosa* the 12 genes in the *alg* cluster are organized as one operon which is transcribed from the promoter P*algD* (102), while the *A. vinelandii alg* cluster contains at least four, possibly overlapping, transcriptional units (57, 58, 72, 74, 79). Recently two putative internal promoters were also identified in the *P. aeruginosa alg* operon (103). Transcription from *P. aeruginosa* P*algD* requires four major transcription factors; the alternative sigma factor σ^{22}, encoded by *algU* (*algT*) (104-107), and the transcriptional regulators AlgR (108, 109), AmrZ (110, 111) and AlgB (106, 112, 113). AlgU also induces expression of these three regulators (106, 110, 114) as well as its own expression (114, 115). AlgR additionally activates expression of *algC*, which probably coordinates expression of this structural gene with expression of the *alg* cluster (116). In addition, AlgP and AlgQ are required for *algD* expression in *P. aeruginosa* (117-119).

In *A. vinelandii algD* is transcribed from three promoters, only one of which requires σ^{22} (72, 120-122), and the transcription of at least two of the other operons in the *A. vinelandii alg* cluster is also AlgU independent (74, 79). AlgU autoregulation has however been observed in this organism (123), similarly to *P. aeruginosa*. The functions of AmrZ, AlgB, AlgP and AlgQ have not yet been studied in *A. vinelandii*. AlgR inactivation has been shown to reduce alginate production, but does not affect *algD* transcription (124). Similarly to *algD*, *A. vinelandii algC* also has two promoters, only one of which is dependent upon AlgU. The elevated AlgC expression provided by the AlgU dependent promoter appears to be required for alginate biosynthesis (76).

Table 1: Regulatory genes involved in alginate biosynthesis in *Pseudomonas aeruginosa* and/or *Azotobacter vinelandii*.

Gene name	Gene product and known/proposed function(s) in alginate biosynthesis	References[1]	
		P. aeruginosa	***A. vinelandii***
Positive regulators			
algU/algT	alternative sigma factor (σ^{22}/σ^{E}) involved in transcription of several alginate biosynthetic genes	(106, 107, 110, 114)	(72, 76, 120, 123)
algR/algR1	P*algD* and P*algC* transcriptional activator (*P. aeruginosa*) / putative transcriptional regulator (*A. vinelandii*)	(108, 109, 116)	(124)
algQ/algR2	anti-σ^{70} factor, regulator of *algD* transcription	(117, 125, 126)	
algB	P*algD* transcriptional activator	(113, 127)	
amrZ/algZ	P*algD* transcriptional activator	(110, 111)	
algP/algR3	putative P*algD* transcriptional activator	(118, 119)	
algW	protease involved in MucA degradation	(128, 129)	
mucE	activator of MucA cleavage by AlgW	(128, 129)	n/a
mucP/yaeL	protease involved in MucA degradation	(128, 130, 131)	
clpX	subunit of ClpXP protease involved in MucA degradation	(132)	
clpP	subunit of ClpXP protease involved in MucA degradation	(132)	
clpP2	protease involved in MucA degradation	(132)	
himA	integration host factor (IHF) subunit; regulator of *algD* and *algB* transcription	(133, 134)	
himD	IHF subunit; regulator of *algD* and *algB* transcription	(133, 135)	
algH	putative transcriptional regulator	(136)	
cysB	P*algD* transcriptional activator	(137)	
rpoN	alternative sigma factor (σ^{54}) involved in *algD* transcription and regulation of MucA degradation	(138, 139)	
prc	protease putatively involved in MucA degradation	(140)	
mucR	c-di-GMP-synthesizing enzyme required for alginate biosynthesis	(141)	
gacS/lemA	sensor kinase that regulates *algD* expression via the GacA/Rsm system		(122, 142)
gacA	GacS cognate response regulator that activates *rsmZ1* and *rsmZ2* transcription		(122, 142)
rsmZ1	small regulatory RNA that antagonizes RsmA activity		(142)
rsmZ2	small regulatory RNA that antagonizes RsmA activity		(142)
rpoS	sigma factor (σ^{S}) involved in *algR* transcription (*P. aeruginosa*) / transcription from the p1-*algD* promoter (*A. vinelandii*)	(143, 144)	(145)
Negative regulators			
mucA/algS	anti-sigma factor that represses σ^{22} activity by sequestration	(146, 147)	(120, 123)
mucB/algN	periplasmic protein that inhibits MucA cleavage by AlgW	(129, 146)	(120)
mucC/algM	putative regulatory protein that may act synergistically with MucA or MucB	(148)	(123)
mucD/algY	regulator of AlgW/MucP-mediated MucA degradation	(131, 149)	(120)
rpoD	sigma factor σ^{70} involved in competitive σ^{22} regulation	(126)	
kinB	putative regulator of MucA degradation	(139)	n/a
ampR	suppressor of P*algU* activity	(150)	n/a
rsmA	mRNA-binding protein that destabilizes *algD* mRNA		(142)

[1]n/a: No close homologs identified in this organism.

In both *P. aeruginosa* and *A. vinelandii algU* is in an operon with *mucABCD*, whose protein products negatively regulate AlgU and thus alginate biosynthesis (120, 123, 148, 149, 151, 152). MucA, an inner-membrane anti-sigma factor that represses σ^{22} by sequestration, is the principal negative regulator of AlgU sigma factor activity, and the sequestration activity of MucA is dependent on the periplasmic MucB protein (146, 147, 153). While the function of MucC is still unclear, MucD appears to act by controlling the signals that activate regulated intramembrane proteolysis of MucA in *P. aeruginosa* (131), where proteolytic regulation has recently been identified as a mechanism for degradation of MucA and thus activation of σ^{22} (154). The process involves the protease AlgW which cleaves MucA in a MucE-dependent manner (128, 129, 155), and the protease MucP which probably catalyzes the complete degradation of AlgW-cleaved MucA (128, 130) as well as AlgW-independent MucA degradation (131). The proteases ClpXP and ClpP2 have also been implicated in proteolysis of MucA (132, 154). MucE, AlgW, MucP, ClpXP and ClpP2 are thus positive regulators of alginate production. *A. vinelandii* does not encode a MucE homolog (156), and the potential roles of the AlgW, MucP, ClpX and ClpP homologs in alginate production have not been investigated in this organism.

Other positive regulators of alginate biosynthesis in *P. aeruginosa* include the integration host factor (IHF) (133-135), the putative transcriptional regulator AlgH (136), the transcription factor CysB (137), the alternative sigma factor RpoN (138, 139), the protease Prc (140, 155) and the c-di-GMP-synthesizing enzyme MucR (141). *A. vinelandii* contains homologs of all these proteins, but none of them have so far been investigated for their possible roles in alginate biosynthesis. KinB, the cognate histidine kinase of AlgB, and the global transcriptional regulator AmpR have been shown to be negative regulators of alginate biosynthesis in *P. aeruginosa* (139, 150), but neither of these have any close homologs in *A. vinelandii*.

The two-component histidine kinase GacS and its global response regulator GacA positively regulate alginate production in *A. vinelandii* (122, 145). This occurs via an Rsm (regulator of secondary metabolism) system where GacA most likely relieves

RsmA repression of *algD* expression by inducing transcription of the small RNAs *rsmZ1* and *rsmZ2* (142). The role of GacS/GacA in *Pseudomonas* spp. is less clear. GacS/GacA has been shown to be required for alginate biosynthesis in *P. viridiflava* (157) and *P. syringae* (158), while a *P. aeruginosa gacA*$^-$ mutant was found to produce normal levels of alginate (159).

Several of the regulators mentioned in this chapter also have regulatory roles in other metabolic and signalling pathways besides those directly related to alginate biosynthesis, underlining that several cellular processes can be expected to influence alginate production levels. Regulatory connections have already been indicated between alginate biosynthesis and for example cell wall stress and recycling, biosynthesis of other polymers, biofilm formation, encystment, antibiotics resistance, quorum sensing, motility and iron uptake (121, 122, 150, 155, 160-164).

1.3 *Azotobacter vinelandii* – biology and taxonomy

A. vinelandii is a widely distributed Gram negative, motile and metabolically versatile free-living soil bacterium. It is an obligate aerobe and has the ability to fix athmospheric nitrogen even under fully aerobic conditions (165). A very high respiration rate is an important mechanism for protecting the oxygen-sensitive nitrogenases from being deactivated by O_2. This is referred to as respiratory protection (166, 167). Additional factors contributing to the protection of nitrogenases from oxygen damage include the FeSII protein (168), autoprotection (169), cellular ATP levels, and maintenance of a low redox state (170). Furthermore, the extracellular alginate capsule can act as a physical barrier against O_2 transfer into the cell (13).

The cell morphology of *A. vinelandii* can vary from rods to more coccoid shapes (165), and under certain conditions *A. vinelandii* cells may differentiate into a resting stage called cysts (12). Encystment renders the cells more resistant to desiccation and radiation compared to vegetative cells (171). Alginate is an important component of the cyst coat, but is also produced during vegetative growth (12). The monomer composition, and thus the physicochemical properties, of the alginates produced by *A.*

vinelandii is mainly controlled by the seven extracellular epimerases AlgE1-7. This presumably allows the organism to adapt to varying environmental conditions by producing alginates suitable for fulfilling different functions (101, 172, 173). The ability to generate alginates of varying composition makes *A. vinelandii* a very interesting candidate for strain engineering with the aim of commercial bioproduction of tailored alginates.

An interesting feature of *A. vinelandii* is that the bacterium can harbour a very high number of chromosomes when cultivated in rich media. In early exponential phase each cell contains only one or a few chromosome copies, but the number increases as the cultures grow and appears to reach >80 chromosome copies per cell in stationary phase (174). This has to be taken into account when generating *A. vinelandii* mutants, as cells may retain both wild-type and mutated chromosomes.

Taxonomically *Azotobacter* is classified as a genus in the family Pseudomonadacae, along with, among others, *Pseudomonas*. *A. vinelandii* is closely related to the *Pseudomonas* species, and genome-based studies have suggested that it might actually belong to the *Pseudomonas* genus (175, 176).

1.4 Genome-wide analyses

1.4.1 Comparative genomics

Genomics deals with the sequencing, assembly, annotation, and structural and functional analysis of genomes. Over the years since the first bacterial genome was sequenced in 1995 (177), bacterial genome sequencing has skyrocketed as development of new technologies keep making it ever easier and cheaper to sequence whole genomes (178, 179). As of today the Genomes OnLine Database contains 6569 completed (finished genomes and permanent drafts) and more than 16 000 ongoing prokaryotic genome projects (180). Despite this progress, functional annotation of bacterial genomes remains a major challenge, especially with regard to direct experimental evidence. There are however many applications of genomic data which do not require access to a gold-standard finished genome.

As the number of sequenced genomes has increased, comparative genomics has emerged as an interesting field in biological research. The ability to make broad comparisons over large numbers of prokaryotic genomes has provided general insights about genome sizes, organization, structure, plasticity, gene content and repetitive elements (181, 182). Furthermore, information about sequence similarity is a major resource for identifying genes and other functional regions, as well as for predicting the actual functions of such regions. Identifying conserved and divergent regions also provide clues about which properties are prevalent in a group of strains or species and which features are unique to a given organism (183, 184). Finally, comparative genomics is a powerful tool for studying taxonomy, evolutionary relationships and genome evolution (182, 185).

1.4.2 Functional genomics and high-throughput screening

The availability of complete genome sequences has opened up a plethora of new opportunities when it comes to investigating gene functions, metabolic and regulatory networks, environmental adaptations etc. While earlier genetic studies would generally be limited to one or a few genes at a time, functional genomics aims to elucidate the relationship between the whole genome of an organism and its observed phenotype, typically by utilizing microarrays and other high-throughput methods (186-188). The use of robotic colony-picking and liquid handling systems coupled with various detector units, as well as computerized data collection and processing, allows researchers to rapidly screen large strain libraries or compound collections to identify active compounds, specific antibodies, genes affecting particular cellular processes etc. (189, 190). Results from such screens are widely used as starting points for drug design and to provide new insights into the biological roles of genes, including their functional organization and interactions in metabolic pathways and cellular processes. Genome-wide functional studies have been fundamental to recent advances in synthetic biology and biological engineering (191, 192).

2 AIMS OF THE STUDY

A. vinelandii is a very interesting candidate for the development of bacterial strains producing specialized alginates, due to its natural ability to synthesize alginates with varying G-content. Thus, substantial efforts have been put into investigating the roles and functions of the epimerases responsible for alginate monomer composition in this bacterium. However, in order to develop bioproduction strains that are also optimized with regard to alginate yield, it is equally important to elucidate which factors influence the overall biosynthetic process *in vivo*. The main motivation of this study was to gain a more global understanding of the genes and cellular processes affecting alginate biosynthesis in *A. vinelandii*. At the onset of the study, the first sequenced *A. vinelandii* genome had recently been assembled and was in the manual annotation phase. Access to this genome sequence and annotation data enabled us to carry out genome-wide analyses of factors affecting alginate production.

Specifically, we wanted to:

- Contribute to generating a manually curated and annotated genome for *A. vinelandii*, as this would be crucial for conducting genome-wide studies of this organism. The annotation and analysis work included in this thesis was mainly aimed at genes involved in carbohydrate metabolism.

- Identify genes, and thus cellular processes, affecting alginate biosynthesis in *A. vinelandii* by constructing and screening a transposon mutant insertion library with regard to alginate production.

- Investigate the global effects of deactivating MucA by monitoring growth, alginate production, respiration and gene transcript levels in batch-fermented *A. vinelandii* wild type and *mucA*$^-$ strains. MucA is the main negative regulator of bacterial alginate biosynthesis and this analysis will aid in identifying the metabolic changes resulting from *mucA*$^-$ mutations that contribute to alginate overproduction.

3 SUMMARY OF RESULTS AND DISCUSSION

3.1 *A. vinelandii* genome annotation and analysis

(Papers I and II, and results not described in papers.)

3.1.1 Summary of genome annotation and analysis

(Paper I)

The publication of the 5.4 Mb genome sequence of the non-alginate producing strain *A. vinelandii* DJ (Paper I) was the result of international community efforts led by key scientists involved in *Azotobacter* research. The original shotgun sequence was generated by the DOE Joint Genome Institute in 2002, and a project to finish, annotate and analyze the genome was initiated in 2006. Approximately 5000 protein-coding genes were identified, out of which ~70% were given a functional assignment (Table 1, Paper I). The *A. vinelandii* DJ genome remained the only published *Azotobacter* genome until the publication of the genomes of *A. vinelandii* CA and its derivative CA6 in June this year (193).

Several previously reported physiological and metabolic properties of *A. vinelandii* were supported by the results from the genome analysis, e.g. obligate aerobic growth, nitrogen fixation, respiratory protection, alginate biosynthesis and the ability to utilize a wide range of carbon sources. The genome analysis also provided new clues about hitherto unknown genes presumed to be involved in several of these cellular processes, either directly or via regulatory mechanisms. Furthermore, several genes encoding verified or predicted oxygen-sensitive enzymes besides the nitrogenases were identified (Figure 3 in Paper I). In early biochemical experiments, *A. vinelandii* was often the organism of choice for producing proteins of high yield and quality. Seen in combination with the organism's innate ability to carry out oxygen-sensitive reactions while at the same time maintaining a strictly aerobic metabolism, *A. vinelandii* shows great promise as a potential host for expression and purification of enzymes that may be damaged by the presence of oxygen.

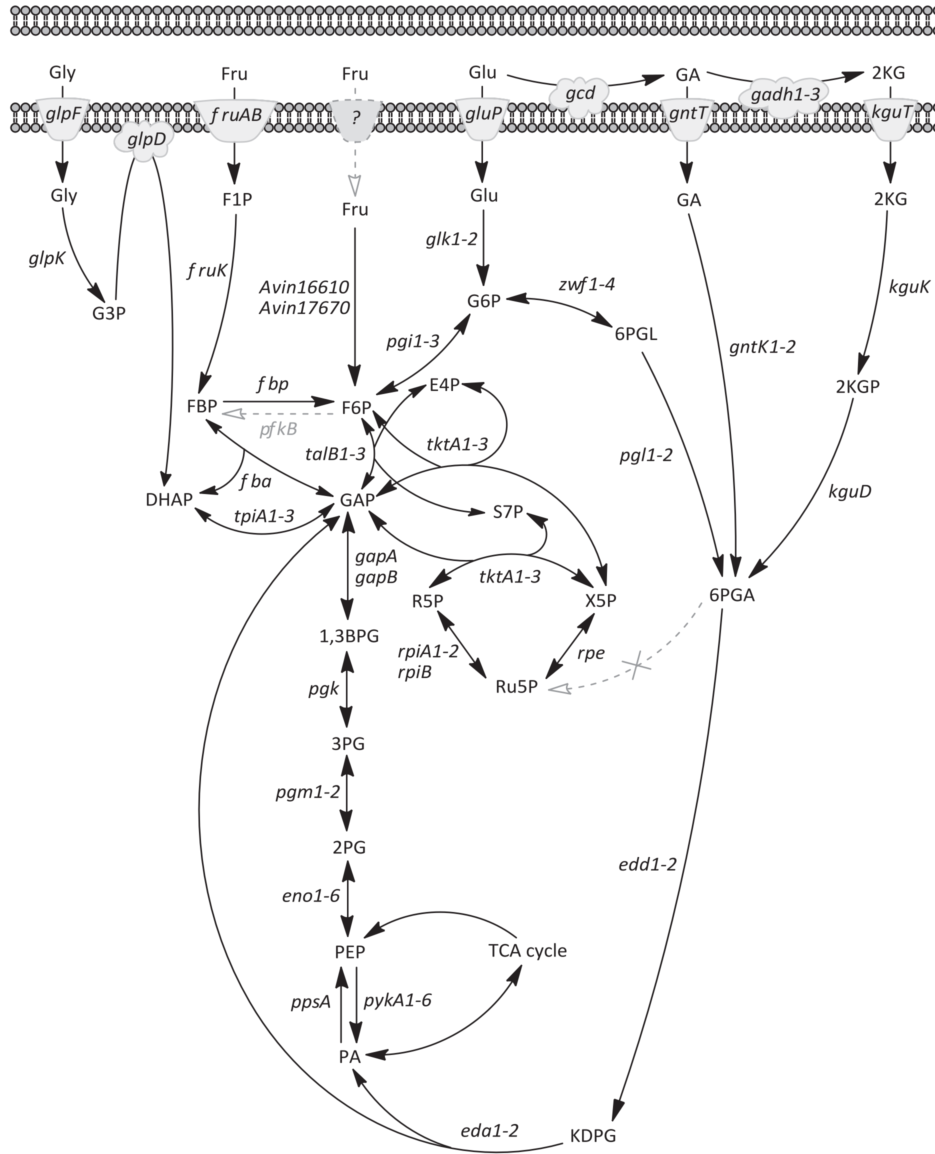


Figure 4: Reconstruction of *Azotobacter vinelandii* core carbohydrate metabolism. Known or putative genes for proteins in the Entner-Doudoroff, gluconeogenetic/glycolytic and pentose phosphate pathways, including glucose, fructose, gluconate and glycerol uptake, identified in the *A. vinelandii* DJ genome. Abbreviations: Gly, glycerol; Fru, fructose; Glu, glucose; GA, gluconate; 2KG, 2-ketogluconate; F1P, fructose 1-phosphate; 6PGL, 6-phosphogluconolactone; G3P, glycerol 3-phosphate; G6P, glucose 6-phosphate; FBP, fructose 1,6-bisphosphate; F6P, fructose 6-phosphate; E4P, erythrose 4-phosphate; 2KGP, 2-keto-6-phosphogluconate; DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde 3-phosphate; S7P, sedoheptulose 7-phosphate; R5P, ribose 5-phosphate; X5P, xylulose 5-phosphate; 6PGA, 6-phosphogluconate; 1,3BPG, 1,3-bisphosphoglycerate; Ru5P, ribulose 5-phosphate; 3PG, 3-

phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; PA, pyruvate; KDPG, 2-keto-3-deoxy-6-phosphogluconate.

3.1.2 Reconstruction of central carbohydrate metabolism

(Results not described in papers.)

The genome annotation and analysis work carried out as part of this thesis was mainly concerned with genes involved in carbohydrate metabolism. It is known that *A. vinelandii* is capable of utilizing a wide range of carbon sources. Accordingly, the strain DJ genome contains genes for the metabolism of several sugars, organic acids and alcohols, including both simple and polymeric compounds. The genome also harbours homologs of many of the genes presumed to be involved in degradation of aromatic compounds in *P. putida* (194). Several of these compounds are derivatives of lignin, and could result from decomposition of plant materials in soil.

In pseudomonads the ED pathway and gluconeogenesis form the core of the central carbohydrate metabolism. These bacteria utilize the cyclic mode of the ED pathway, and have two routes for the conversion of glucose to 6-phosphogluconate; the oxidative and the phosphorylative pathways (62, 195). This is also assumed to be the case for *A. vinelandii*. As expected, the genes required for the ED, gluconeogenetic and pentose phosphate (PP) pathways were identified in the *A. vinelandii* DJ genome, except for the PP pathway enzyme 6-phosphogluconate dehydrogenase (Gnd) (Figure 4). It is thus possible that *A. vinelandii* lacks this enzyme, like *A. beijerinckii* (196) and several pseudomonads (62). Genes enabling a complete tricarboxylic acid (TCA) cycle were found to be present in the genome.

A common feature among *Pseudomonas* and *Azotobacter* spp. is a lack of 6-phosphofructokinase, which renders them unable to utilize the glycolytic pathway (62, 196, 197). Interestingly, the *A. vinelandii* DJ genome encodes a protein (Avin23410) with similarity to the minor 6-phosphofructokinase of *Escherichia coli*, 6-phosphofructokinase II (PfkB) (198, 199). A close homolog of Avin23410 exists in *P. stutzeri*, but none of the other sequenced *Pseudomonas* spp. (61). It could be

interesting to investigate further the possibility that *Avin23410* encodes a functional PfkB, and whether this enables *A. vinelandii* to utilize the glycolytic pathway to some extent. Our transcriptome data indicate that the putative *pfkB* gene is not expressed when *A. vinelandii* is cultivated with fructose as the carbon source (Table S2 in Paper IV). 6-Phosphofructokinase is however not necessary for glycolytic degradation of fructose if the sugar is imported via a fructose PTS (Figure 4).

Despite the highly similar pathways of central carbohydrate metabolism, glucose import in *A. vinelandii* appears to differ from that in *Pseudomonas* spp. Neither *A. vinelandii* nor *P. aeruginosa* utilize a glucose PTS (200), and no genes for such a system could be identified in the *A. vinelandii* DJ genome. In *Pseudomonas* spp. glucose is imported via the OprB porin (201, 202) and an ATP-binding cassette transporter (203, 204), neither of which have any homologs in *A. vinelandii* DJ. *A. vinelandii* has been reported to import glucose via a glucose permease (205), and Avin04150 was identified as a likely candidate due to its similarity to the major facilitator superfamily type glucose/galactose porter GluP in *Brucella abortus* (206). No close homologs of *A. vinelandii* GluP were found in the *Pseudomonas* Genome Database (61). Glucose transport across the outer membrane has not been investigated in *A. vinelandii*, but most likely occurs via general porins or substrate-specific porins other than OprB.

Fructose on the other hand may be imported and phosphorylated by a fructose PTS, similarly to most pseudomonads (62), as genes encoding a putative fructose PTS (FruAB) and 1-phosphofructokinase (FruK) were identified in the *A. vinelandii* DJ genome (*Avin12190-12210*). Results from a previous study have however indicated that fructose is imported without being phosphorylated in *A. vinelandii*, and then converted to fructose 6-phosphate by a fructokinase (65). No putative fructose transporters were identified in our study, but the genome does encode two proteins (Avin16610 and -17670) with homology to the *P. fluorescens* fructose, glucose and glucitol kinase MtlZ (207). It is thus possible that *A. vinelandii* employs two alternative routes for fructose uptake and metabolism (Figure 4), in which case the fructokinase might be the preferred route, since our transcriptome data indicate higher expression

levels for *Avin16610* and *-17670* than the *fruBKA* genes under the conditions studied (Table S2 in Paper IV).

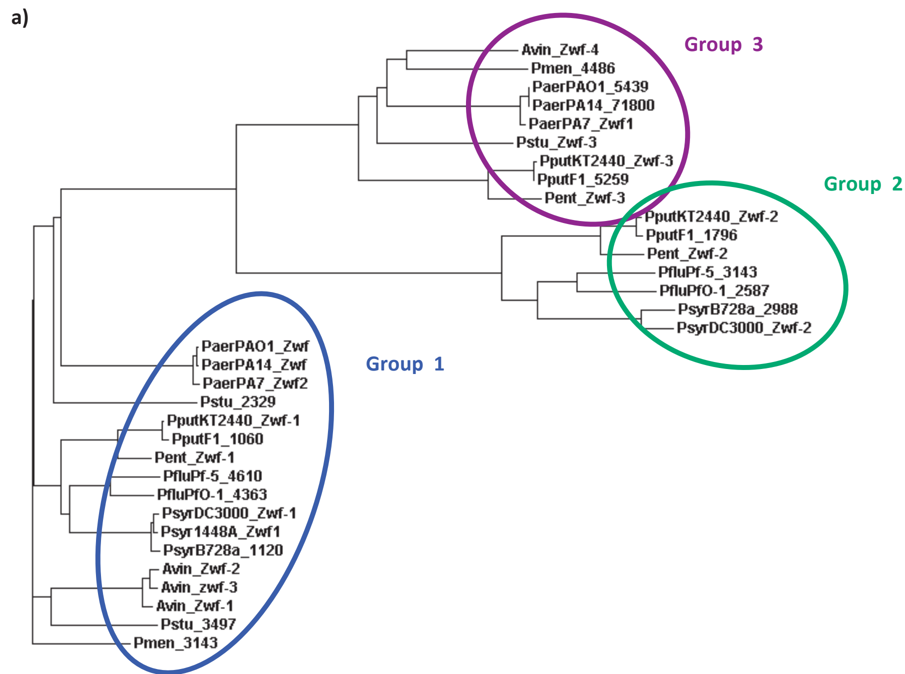
3.1.3 *A. vinelandii* and *Pseudomonas* spp. glucose 6-phosphate dehydrogenase genes

(Results not described in papers.)

Four putative glucose 6-phosphate dehydrogenase (Zwf) genes were identified in the *A. vinelandii* DJ genome (Figure 4), three of which are nearly identical (*zwf1-3*). Glucose 6-phosphate dehydrogenase catalyzes the initial step in the ED and PP pathways, and is thus central to *Pseudomonas* and *Azotobacter* carbohydrate metabolism. In *Pseudomonas* spp. the number of *zwf* genes appears to vary between one and three (61). It seems that *Azotobacter* and *Pseudomonas zwf* genes can be divided into three groups, based on phylogenetic relationships and genomic contexts (Figure 5). Group 1 comprises *Pseudomonas* spp. *zwf* genes which are part of the HexR-regulated *zwf-pgl-eda* operon encoding ED pathway proteins (208, 209), as well as *A. vinelandii zwf1-3* which display different, but related, genomic contexts. Group 2 consists of *zwf* genes that are part of a possible PP pathway *gnd-zwf* operon, and is not present in the *A. vinelandii*, *P. aeruginosa*, *P. stutzeri* and *P. mendocina* strains investigated, all of which also lack *gnd*. Group 3 contains solo *zwf* homologs located next to a divergently transcribed *hexR* homolog.

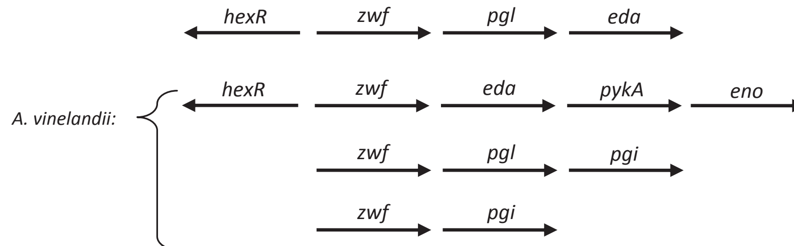
Research on *Pseudomonas* sp. Zwf has mainly focused on the proteins which are here classified as group 1. It has however long been known that two different Zwf proteins, assumed to function primarily in the ED and PP pathway respectively, exist in *P. fluorescens* (210). Based on genomic contexts it seems likely that the *P. fluorescens* group 1 Zwf is the ED enzyme and group 2 Zwf the PP enzyme. No references to the group 3 *zwf* genes could be found in the literature. Thus, this group appears to encompass a set of uninvestigated *zwf* homologs that are present in *A. vinelandii* and several *Pseudomonas* spp.

a)



b)

Group 1: *P. aeruginosa*, *P. entomophila*, *P. fluorescens*, *P. mendocina*, *P. putida*, *P. stutzeri*[1], *P. syringae:*



Group 2: *P. entomophila*, *P. fluorescens*, *P. putida*, *P. syringae:*

Group 3: *A. vinelandii*, *P. aeruginosa*, *P. entomophila*, *P. mendocina*, *P. putida*, *P. stutzeri:*

Figure 5: Phylogram (a) and genomic contexts (b) for *zwf* genes in *Azotobacter vinelandii* and *Pseudomonas* spp. Avin: *A. vinelandii* DJ, Paer: *P. aeruginosa* (strains PAO1, PA7 and PA14), Pent: *P. entomophila* L48, Pflu: *P. fluorescens* (strains PfO-1 and Pf-5), Pmen: *P. mendocina* ymp, Pput: *P. putida* (strains F1 and KT2440), Pstu: *P. stutzeri* A1501, Psyr: *P. syringae* (strains pv. phaseolicola 1448A, pv. syringae B728a and pv. tomato str. DC3000). Phylogram generated using ClustalW2 (211). Gene products: *hexR*; transcriptional repressor, *gnd*; 6-phosphogluconate dehydrogenase. For other gene products see Table 2. [1]*P. stutzeri* has two such clusters, one of which is lacking *hexR*.

3.1.4 Occurrence of highly similar carbohydrate metabolism synologs

(Paper II and results not described in papers.)

The annotation work on the *A. vinelandii* central carbohydrate metabolism genes revealed numerous occurrences of intra-genome homologs (Figure 4). Such homologs can arise by gene duplication (paralogs) or via horizontal gene transfer (xenologs or pseudoparalogs). It is usually not possible to discern between pseudoparalogs and true paralogs in a single-genome analysis (212), so in this work the term synologs, proposed by Lerat *et al*. (213), will be used to describe intra-genome homologs arising from either duplication or horizontal transfer. The *A. vinelandii* DJ central carbohydrate metabolism genes that were found to have synologs are shown in Table 2. The number of synologs and degree of sequence similarity varies, but in many cases the homologous proteins are nearly identical.

A closer investigation of *A. vinelandii* DJ and 15 fully sequenced *Pseudomonas* strains showed that *A. vinelandii* contains a markedly larger number of families (groups of similar proteins) than the pseudomonads among proteins related to carbohydrate metabolism (Figure 1 in Paper II), but not in any other functional category. Furthermore, in this category highly similar synologs (≥90% protein sequence identity) and protein families with more than two members are more common in *A. vinelandii* than in the investigated *Pseudomonas* strains.

To investigate the occurrence of carbohydrate metabolism synologs in a wider range of organisms, relevant gene data from 57 archaeal and 886 bacterial genomes in the SEED database (214) were analyzed. The initial results indicated that highly similar synologs in carbohydrate metabolism are not common in bacterial and archaeal genomes, as the average protein sequence identity between synolog pairs is ≤50% in the vast majority of the analyzed genomes (Figure 2 in Paper II). Filtering the data set to contain only synolog groups (sets of two or more intra-genome homologs) with protein sequence identities ≥90% confirmed that such synologs are indeed rare in bacteria and archaea (Figure 6).

Table 2: Synologs among *Azotobacter vinelandii* central carbohydrate metabolism genes.

Gene name	Protein	Synologs	Protein length (aa)	Protein sequence similarity: Identities (positives)
glk	glucokinase	Avin15690 (Glk1)	322	58% (68%)
		Avin04130 (Glk2)	322	
pgi	glucose 6-phosphate isomerase	Avin17650 (Pgi1)	554	81-95% (90-96%)
		Avin16630 (Pgi2)	554	
		Avin42740 (Pgi3)	544	
tpiA	triosephosphate isomerase	Avin49960 (TpiA1)	251	79-100% (86-100%)
		Avin45770 (TpiA2)	251	
		Avin49910 (TpiA3)	246	
pgm	2,3-bisphosphoglycerate-independent phosphoglycerate mutase	Avin04670 (Pgm1)	511	97% (98%)
		Avin27430 (Pgm2)	511	
eno	enolase	Avin27230 (Eno1)	437	74-100% (84-100%)
		Avin22210 (Eno2)	437	
		Avin43360 (Eno3)	438	
		Avin45760 (Eno4)	439	
		Avin49980 (Eno5)	439	
		Avin38790 (Eno6)	429	
pykA	pyruvate kinase	Avin27240 (PykA1)	481	64-99% (76-99%)
		Avin22180 (PykA2)	481	
		Avin43370 (PykA3)	481	
		Avin51360 (PykA4)	481	
		Avin50000 (PykA5)	481	
		Avin38410 (PykA6)	483	
gntK	gluconokinase	Avin22480 (GntK1)	174	88% (91%)
		Avin27310 (GntK2)	175	
zwf	glucose 6-phosphate dehydrogenase	Avin27260 (Zwf1)	489	53-98% (66-98%)
		Avin17630 (Zwf2)	488	
		Avin16620 (Zwf3)	487	
		Avin02020 (Zwf4)	485	
pgl	6-phosphogluconolactonase	Avin15710 (Pgl1)	238	56% (64%)
		Avin17640 (Pgl2)	237	
edd	6-phosphogluconate dehydratase	Avin15680 (Edd1)	608	75% (85%)
		Avin27280 (Edd2)	608	
eda	KDPG aldolase	Avin27250 (Eda1)	212	68% (81%)
		Avin15720 (Eda2)	212	
rpiA	ribose 5-phosphate isomerase	Avin22160 (RpiA1)	224	87% (93%)
		Avin48280 (RpiA2)	226	
tktA	transketolase	Avin22150 (TktA1)	672	69-100% (79-100%)
		Avin27330 (TktA2)	672	
		Avin05560 (TktA3)	665	
talB	transaldolase	Avin22140 (TalB1)	322	60-100% (74-100%)
		Avin27340 (TalB2)	322	
		Avin32700 (TalB3)	308	

While there was a positive correlation between the number of synolog groups and the total number of carbohydrate metabolism genes or protein-coding genes in the genomes (R^2 = 0.80 and 0.56 respectively for a linear regression), there was practically

no correlation when only synologs with protein sequence identities ≥90% were included (R^2 = 0.14 and 0.15 respectively). This shows that the large amount of highly similar synologs observed in some genomes is not simply due to a high number of genes in general.

The analysis showed that *A. vinelandii* DJ contains 38 highly similar (≥90% identity) carbohydrate metabolism synologs distributed into 15 synolog groups. As can be seen from Figure 6, this is an unusually large number, which places strain DJ among the five

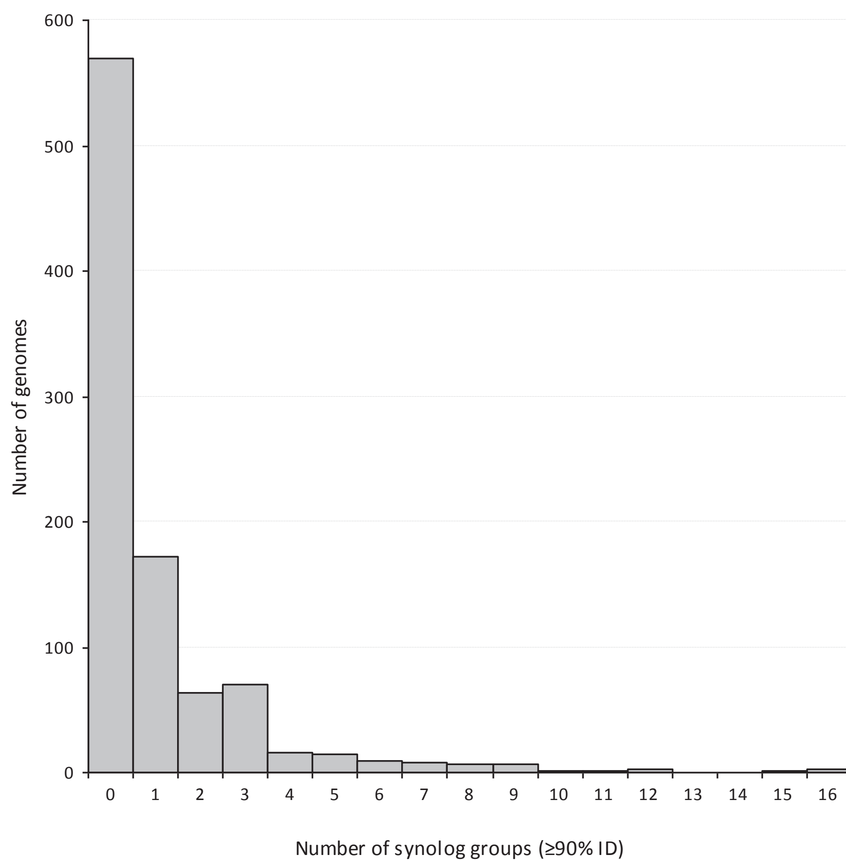


Figure 6: Occurrence of highly similar synolog groups among carbohydrate metabolism proteins in bacteria and archaea. Number of synolog groups with internal protein sequence identity ≥90% observed for the 943 investigated prokaryotic genomes. (Synolog group: set of two or more intra-genome homologs.)

organisms harbouring the highest number of such synologs. The majority of these synolog groups occur in core metabolic processes, and the genes in a given synolog group are generally localized in varying, but metabolically relevant genomic contexts, often as parts of probable operons (Figure 3 in Paper II). Although the SEED subcategory distribution of synologs varies among the investigated genomes containing large numbers of highly similar synologs (Table 2 in Paper II), the differing but metabolically relevant genomic contexts were observed to be common in several of these genomes. Thus, adaptation leading to favourable genomic arrangements, which is assumed to be a very slow process (215, 216), appears to have taken place along with sequence conservation. This indicates that these synologs are not simply the result of very recent duplications and/or horizontal transfer events.

The presence of highly similar synologs may confer adaptive benefits due to increased gene dosage. However, we propose that retaining highly similar synologs in different operonic contexts could also be advantageous with regard to co-regulation of genes, while at the same time leaving more room for fine-tuning expression of certain proteins due to the presence of multiple gene copies. Furthermore, our results suggest that the occurrence of highly similar carbohydrate metabolism synologs could be connected to certain aspects of an organism's lifestyle or environment. The majority of such synologs identified in a given organism are usually clustered in only one or two SEED subcategories (Table 2 in Paper II), and closer investigation revealed several examples of these synologs occurring in metabolic processes that are central for the organism in question.

3.2 Genome-wide studies of factors affecting *A. vinelandii* alginate biosynthesis

(Papers III and IV)

3.2.1 Identification of genes affecting alginate biosynthesis

(Paper III)

SUMMARY OF RESULTS AND DISCUSSION

In order to investigate factors affecting alginate biosynthesis in *A. vinelandii* we constructed a transposon insertion mutant library from the tetracyclin-resistant strain ATCC12518Tc (derived from strain ATCC12518). Cultures of about 4000 independent mutant strains were cultivated in microtiter plates and screened for an increase or decrease in alginate production levels relative to the ATCC12518Tc reference. Cultures that failed to grow were disregarded. Selected mutants were cultivated in deep-well plates, shake flasks and/or batch fermentors to verify the observed phenotypes. The availability of the sequenced and annotated genome of strain *A. vinelandii* DJ (Paper I), which is also an ATCC12518 derivative, enabled easy identification of the affected genes in interesting mutants by sequencing from the transposon insertion. Strain DJ is however non-alginate producing due to a mutation in *algU* (120), and could therefore not be used to construct the transposon insertion library described in this work.

Abolished or decreased alginate production was verified for ~70 transposon insertion mutants, while only two mutants, containing *mucA* and *pgm-2* disruptions respectively, were confirmed to have gained an increase in alginate production (Table 2 in Paper III). The apparent increase in alginate production in the *pgm-2* mutant is likely due to elevated cell culture density. From a bioproduction perspective a mutant with the ability to increase cell yield without reducing alginate production could however still be interesting for further studies. MucA is the main negative regulator of alginate biosynthesis, and inactivation has also previously been shown to increase alginate production in *A. vinelandii* (120, 123) and to activate alginate biosynthesis in initially non-producing *P. aeruginosa* (146, 151) and *P. fluorescens* (217) strains.

The scarcity of alginate overproducing mutants identified in this work indicates that there are very few single gene mutations that will have a profound positive effect on alginate biosynthesis in *A. vinelandii*. Additionally, an earlier study has shown that *A. vinelandii* mutants with substantially increased alginate production are often unstable and quickly revert to normal levels (123). Due to the use of selective media, reversion via loss of the transposon insertion should not occur in our experiments. Suppressor mutations and retention of wild type chromosome copies (see Chapter 1.3) could

however be contributing factors to the low reproducibility observed for apparent alginate overproducing mutant strains in this work.

As could be expected, the transposon insertions in several of the alginate non-producing mutants are located in structural genes directly involved in alginate biosynthesis (Table 2 in Paper III). Additionally, transposon insertions in homologs of *P. aeruginosa amrZ* and *algW*, which encode regulators required for alginate biosynthesis (111, 155), resulted in a loss of alginate production (Table 2 in Paper III). This suggests similar roles for these regulators in *A. vinelandii* alginate biosynthesis. The regulator AlgB is also required for alginate production in *P. aeruginosa* (113), but an *algB*::TnCAM140 mutant identfied in this work has production levels comparable to the reference strain. Thus, it appears the presence of AlgB is not necessary for alginate production in *A. vinelandii*.

Our results revealed two possible new regulators of *A. vinelandii* alginate biosynthesis: the putative IclR family transcriptional regulator Avin13880 and the fructose PTS protein FruA. Transposon insertions in either of these genes abolished alginate biosynthesis (Table 2 in Paper III). Homologs of both genes are present in *Pseudomonas* spp. (61), but nothing has been reported on either gene's potential role in relation to alginate biosynthesis. It is however known that many carbohydrate PTS systems have regulatory roles besides their function as transporters (218). Furthermore, cross-talk has been demonstrated between the PTS^{Fru} and PTS^{Ntr} systems in *P. putida* (219), and PTS^{Ntr} is believed to be involved in coordination of nitrogen and carbon metabolism (218). The *A. vinelandii fruA*::TnCAM140 mutant was unable to produce alginate even if fructose was replaced with glucose, glycerol or sucrose in the growth medium, thus ruling out the possibility that the abolished alginate production is caused by a defective fructose uptake mechanism.

In addition our results showed that gene inactivations affecting iron uptake, peptidoglycan recycling and flagella biosynthesis appear to have a negative effect on alginate production. Connections between alginate biosynthesis and these processes

have also been reported by others. Elevated siderophore production has been shown for mucoid *P. aeruginosa* strains (220), which fits with our observations. *A. vinelandii* and *P. aeruginosa* mutants impaired in peptidoglycan recycling have however been found to have an increase in alginate levels in previous studies (160, 221). Motility is negatively regulated by the alginate sigma factor AlgU in both *P. aeruginosa* and *A. vinelandii* (162, 222, 223), but there is currently no data in the literature that would suggest an alginate negative phenotype for mutants with impaired flagella biosynthesis. It is clear that further investigations will be necessary to elucidate the metabolic and/or regulatory connections between alginate biosynthesis and the above mentioned processes in *A. vinelandii*.

3.2.2 Medium supplements increasing alginate production

(Paper III)

Several *A. vinelandii* mutants with transposon insertions in genes involved in the biosynthesis of cofactors or central metabolites (isoprenoids, pyridoxine, purines, thiamine, biotin and TCA cycle intermediates) produced no or very small amounts of alginate while retaining normal growth levels (Table 2 in Paper III). This indicates that a lack of certain vitamins or key metabolic precursor molecules has a negative effect on alginate biosynthesis. To investigate whether the availability of such compounds is a limiting factor for alginate production, the reference strain ATCC12518Tc was cultivated in liquid media containing different supplements which were selected based on our observations (Paper III) as well as data from the literature: the purine adenine; the vitamins thiamine (B_1), pyridoxal/pyridoxine (B_6) and biotin (B_7); the TCA cycle intermediate succinate; the amino acids lysine, methionine and diaminopimelate. These three amino acids are derived from TCA cycle intermediates, and have furthermore been shown to lower thiamine requirements in *Salmonella typhimurium* (224).

None of the supplements had a significant effect on growth, but addition of thiamine, succinate or a mixture of lysine, methionine and diaminopimelate increased the

amount of measured alginate in the cultures (Figure 1 in Paper III). Thus, our results clearly indicate that alginate biosynthesis in *A. vinelandii* is limited by the availability of thiamine, and possibly certain TCA cycle intermediates. The TCA cycle also appears to be involved in alginate biosynthesis in *P. fluoresecens*, as TCA cycle genes were upregulated in a *mucA*⁻ mutant (217). The same study also reported upregulation of purine biosynthesis genes, and metabolic profiling of *P. fluorescens* have shown that MucA inactivation and alginate biosynthesis affect the purine NTP pools in the cells (225). No significant effect on alginate concentration was however observed in the adenine supplemented *A. vinelandii* cultures (Figure 1 in Paper III), indicating that although purines are important for alginate biosynthesis this nutrient is not a limiting factor under the conditions used in this experiment.

3.2.3 Phenotypic characterization of wild type and *mucA*⁻ strains

(Paper IV)

MucA is the principal negative regulator of the alternative sigma factor σ^{22} (AlgU) which is required for transcription of several genes involved in alginate biosynthesis (see references in Table 1). MucA inactivation has been shown to activate alginate biosynthesis in *P. aeruginosa* (146, 151) and *P. fluorescens* (217), and to result in increased alginate production in *A. vinelandii* (120, 123, Paper III). To investigate the effects of inactivating *mucA* with regard to growth, alginate production, respiration and carbon source utilization, comparative studies of batch fermented *A. vinelandii* ATCC12518 and a *mucA*⁻ derivative were performed. The *mucA* mutant was constructed by the Molecular Microbiology research group at Universidad Nacional Autónoma de México and the fermentation studies were performed by SINTEF Biotechnology and Nanomedicine.

The results clearly showed that the *mucA*⁻ mutant had increased alginate production and lowered growth rate and respiration, which means the *mucA*⁻ mutant converts less fructose to CO_2 and more to alginate compared to the wild type strain (Figures 1 and 2 in Paper IV). Diminished respiration rates have also been observed for *P. fluorescens*

mucA⁻ mutants (217). Alginate production requires a high input of both carbon and energy (226), which could be a direct explanation for the diminished growth in the *A. vinelandii mucA⁻* cultures. *A. vinelandii* is known to have a very high respiration rate, which is assumed to have a role in protection of the oxygen-sensitive nitrogenases (166, 167). From a bioproduction perspective a high respiration rate is not desirable, as it likely contributes to a lower yield of alginate per gram carbon source. The highest alginate yields were obtained in 0.5x PM1 (101) medium (Figure 2 in paper IV). In this medium CO_2 was shown to account for 66% and 40% of the utilized carbon in the wild type and *mucA⁻* fermentations respectively, while alginate accounted for 14% and 32%. In comparison, xanthan yields of 60-70% have in some cases been reported for batch fermentations of *Xanthomonas campestris*, a bacterium used for commercial bioproduction of this polysaccharide (227).

3.2.4 Comparative transcriptome analyses of wild type and *mucA⁻* strains

(Paper IV)

In addition to regulating alginate biosynthesis, MucA/AlgU also affects other cellular processes in both *A. vinelandii* (120, 121, 223) and *Pseudomonas* spp. (217, 228-233). In order to further investigate the metabolic changes accompanying *mucA* inactivation and alginate overproduction, we compared the transcriptomes of fructose fed-batch fermented *A. vinelandii* ATCC12518 and its *mucA⁻* derivative. Since the emphasis of our studies is on alginate production, 0.5x PM1 medium was chosen for the cultivations (see Chapter 3.2.3). Samples for transcript analyses were collected from duplicate cultivations at five different time points ranging from 9.1 to 32.3 hours after inoculation. There was good reproducibility between the replicas (Figure 3 in Paper IV). Collecting data from several time points makes it possible to discover trends that are different in the two data sets, as well as discern between changes caused by the *mucA* mutation and differences attributed to variations in growth kinetics. As could be expected, most genes showed approximately the same transcript levels and patterns in

both strains. Both transcript level and fold change was considered when evaluating the results, and a twofold difference was regarded as significant.

As could be expected, the majority of the genes directly involved in alginate biosynthesis were significantly upregulated in the *mucA* mutant compared to the wild type strain (Table S5 in Paper IV). Interestingly, this includes the genes encoding the secreted AlgE epimerases, except *algE4*, and their secretion system. As these epimerases introduce the G-blocks required for formation of strong alginate gels (28, 172), *A. vinelandii* AlgU appears to have a role in controlling alginate composition in addition to regulating production levels. The *A. vinelandii alg* gene cluster is divided into several transcriptional units, out of which at least two are AlgU independent (see Chapter 1.2.2). The elevated transcription of the entire cluster in the *mucA*$^-$ strain thus indicates that the *A. vinelandii alg* cluster might be transcribed as one operon, as in *P. aeruginosa*, in addition to the shorter transcriptional units. Since *A. vinelandii* (unlike *Pseudomonas* spp.) produces alginate constitutively, one would expect the biosynthetic proteins to be expressed even in the presence of a functional *mucA*. This is supported by our data, which show that the corresponding genes, despite being upregulated in the *mucA* mutant, also display high transcription levels in the wild type strain (Table S5 in Paper IV). In contrast, transcription of alginate biosynthesis genes was high in a *P. fluorescens mucA*$^-$ mutant, but only low levels were detected in the non-mucoid wild type strain (217).

Genes encoding glycolytic and/or gluconeogenetic enzymes, as well as a putative fructokinase (Avin16610; Figure 4), also appear to be upregulated in the *A. vinelandii mucA* mutant (Table S2 and Figure 4 in Paper IV). This could reflect increased channeling of carbon towards the alginate precursor molecule fructose 6-phosphate. Similarly, the elevated transcription of the oxidative phosphorylation pathway genes encoding the ATP-coupled NADH dehydrogenase and cytochrome *c* oxidases (Figure 4 and Table S3 in Paper IV) in the *mucA* mutant could reflect an increased energy need in order to carry out a higher level of alginate biosynthesis. Genes involved in energy generation were also recently found to be upregulated in an alginate-producing *P.*

fluorescens mucA mutant (217). The high respiration rate of *A. vinelandii* is partly caused by the activity of oxidative phosphorylation components that are partially uncoupled from ATP synthesis, like the cytochrome *bd* oxidase (CydAB) (234). The *cydAB* genes displayed significantly lower transcript levels in the *mucA*$^-$ mutant, which in combination with the increased transcription of genes involved in ATP-coupled oxidative phosphorylation could explain the reduced respiration rates observed for this mutant strain (Figures 1 and 2 in Paper IV).

Expression of *cydAB* is negatively regulated by CydR (235), and *cydR* transcription has previously been shown to be elevated in the presence of AlgU (223). Our data however showed comparable *cydR* transcription levels in the *mucA*$^-$ and wild type strains (Table S3 in Paper IV). Nevertheless the *flhCD* genes, which are known to be negatively regulated by AlgU via CydR (223), also displayed significant transcriptional downregulation in the *mucA* mutant (Table S6 in Paper IV). Our data thus suggest that AlgU controls *cydR* expression, and hence CydAB and respiration rate, in an indirect and possibly post-transcriptional manner.

Interestingly, *fruR* transcription appears to be downregulated in the *mucA*$^-$ strain (Table S2 in Paper IV). FruR (Cra) has a key role in regulation of carbohydrate metabolism in enteric bacteria (236-238). Furthermore, in *P. putida* FruR controls expression of the *fruBKA* genes encoding the fructose PTS and has been implicated in regulation of cross-talk between PTS^{Fru} and PTS^{Ntr} (239), the latter of which has a role in control of carbon utilization (240). Reduced transcription of *fruR* could thus be involved in the changes in carbon metabolism observed for the *A. vinelandii mucA* mutant, possibly via the regulatory functions of PTS systems. We also found that alginate production was abolished in an *A. vinelandii fruA* mutant (Paper III), which further supports a regulatory role for PTS^{Fru} in this organism.

Somewhat surprisingly, the genes encoding the Nif and Vnf nitrogenases displayed high transcript levels in both strains (Table S4 in Paper IV) despite the presence of fixed nitrogen in the PM1 medium. It is generally assumed that the nitrogenase genes are

only active when the supply of fixed nitrogen is limiting. If so, it appears the supplemented peptone is only sufficient to cover the organism's needs for the first few hours. This should be taken into consideration when developing media for alginate bioproduction, since the high respiration rate of *A. vinelandii* is believed to be coupled to protection of the oxygen-sensitive nitrogenases (166, 167). It should also be noted that the *vnf* genes have previously been reported to be transcribed only under molybdenum deficient conditions (241, 242), but this was not the case in our cultivations.

Expression of the genes encoding the Nif nitrogenase requires the transcriptional activator NifA (243). Genome annotation of the *A. vinelandii* DJ genome revealed a *nifA* homolog, designated *nifA2* (Paper I). Interestingly, *nifA2* had higher transcript levels than *nifA* in both strains. Furthermore, *nifA2* transcription was significantly higher in the wild type than the *mucA*$^{-}$ strain (Table S4 in Paper IV), which suggests that AlgU regulates NifA2 expression negatively. Our data thus indicate a possible regulatory connection between MucA/AlgU and nitrogen fixation. It is however not known whether NifA2 has a function in regulation of the nitrogenase genes, and all previous studies of *nif* gene regulation have been carried out in strains lacking a functional AlgU.

Lastly, two genes (*Avin44240* and *ispD*) involved in biosynthesis of biotin and the isoprenoid precursor isopentylpyrophosphate were observed to be transcriptionally upregulated in the *mucA*$^{-}$ mutant (Table S7 in Paper IV). This is quite interesting, as we have also observed that mutants with transposon insertions in biotin and isoprenoid biosynthesis genes are unable to produce alginate (Paper III). A total of 54 genes showed significant upregulation at all five time points in the *mucA* mutant, while 175 genes were significantly downregulated at all time points (Tables S6 and S7 in Paper IV). The majority of these genes were annotated as hypothetical or with unknown function. Our results nonetheless indicate that MucA/AlgU activity potentially affects a large regulon encompassing several different cellular processes.

4 CONCLUDING REMARKS

In order to successfully develop strains and protocols for bacterial bioproduction of alginate it is important to gain extensive knowledge about the factors that affect the biosynthetic process. The availability of a sequenced and thoroughly annotated *A. vinelandii* genome (Paper I) has opened up a range of new possibilities for addressing these issues on a global scale. In this work we have established and put to use protocols for two such high-throughput approaches, transcriptional profiling (Paper IV) and transposon insertion library screening (Paper III), optimized for use with alginate-producing *A. vinelandii* strains. These established and demonstrated protocols will be useful for future investigations of this bacterium.

The precursor for alginate biosynthesis is fructose 6-phosphate, which renders this biosynthetic pathway closely tied to the central carbohydrate metabolism (Figure 4). This work has provided a solid foundation for further investigations of sugar uptake and metabolism in *A. vinelandii*, and can be used to design metabolic engineering strategies aimed at optimizing the channeling of imported sugars towards alginate biosynthesis. An unusual and interesting feature of the *A. vinelandii* core carbohydrate metabolism is the abundance of apparently conserved highly similar synologs (Table 2 and Figure 4). We propose that the retention of such synologs may confer adaptive benefits via increased gene dosage and/or increased regulatory flexibility (Paper II), but further studies will be necessary to elucidate the biological roles of these intra-genome homologs.

Our results have provided new insights with regard to regulation of alginate biosynthesis in *A. vinelandii*, including the discovery of new putative regulators (Paper III). Furthermore, we have revealed new roles for the key regulators MucA/AlgU with regard to control of alginate composition, cell mass production, carbon metabolism, respiration and possibly nitrogen fixation (Paper IV). Regulation of alginate biosynthesis has so far mainly been studied in *P. aeruginosa* (244), and our results have underlined both similarities and differences (Paper III and IV) besides those

already reported in the literature (245). Considering the dissimilar life styles and the constitutive versus circumstantial modes of alginate production, it is hardly surprising that regulation operates differently in these organisms. Nevertheless, this highlights the need for further investigations of the alginate regulatory mechanisms in *A. vinelandii* if this organism is to be used for alginate bioproduction.

Alginate biosynthesis appears to be connected to a variety of cellular processes (Paper III and IV), which adds to the complexity of predicting the effects of both genetic and environmental perturbations. Elucidating the exact nature of all these connections would likely be very demanding. Nevertheless, we have demonstrated that even sparse information about metabolic relationships can aid in identifying nutrients that are limiting alginate production (Paper III), and thus be useful for development and optimization of bioproduction media.

Inactivation of the anti-sigma factor MucA resulted in increased alginate production and decreased respiration and growth rate. The *A. vinelandii mucA*$^-$ mutant thus converts less of the supplemented fructose to CO_2 and more to alginate relative to the wild type strain (Paper IV). From a bioproduction perspective this makes the *mucA* mutant an improved production strain candidate due to the increased alginate yield. The high respiration rate of *A. vinelandii* is however a challenge with regard to alginate yield, even in the *mucA* mutant (Paper IV). Engineering mutant strains with reduced respiration rates, for example via inactivation of decoupled oxidative phosphorylation components, could be an interesting experimental strategy for the development of bioproduction strains.

In summary, we have brought forth important results with regard to the long term goal of bacterial strain engineering and cultivation optimization for commercial alginate bioproduction purposes. Our studies underline that in order to improve alginate yield in *A. vinelandii* the relationships between alginate biosynthesis, cell mass production, nitrogen fixation and respiration, and thus carbon utilization, should be investigated further.

REFERENCES

1. **Stanford ECC.** 1881. Manufacture of useful products from seaweeds. Great Britain.
2. **Haug A.** 1964. Report No. 30. Composition and properties of alginates. PhD thesis. Norwegian Institute of Technology, Trondheim.
3. **Linker A, Jones RS.** 1964. A polysaccharide resembling alginic acid from a *Pseudomonas* micro-organism. Nature **204:**187-188.
4. **Carlson DM, Matthews LW.** 1966. Polyuronic acids produced by *Pseudomonas aeruginosa*. Biochemistry **5:**2817-2822.
5. **Gorin PAJ, Spencer JFT.** 1966. Exocellular alginic acid from *Azotobacter vinelandii*. Can J Chem **44:**993-998.
6. **Govan JR, Fyfe JA, Jarman TR.** 1981. Isolation of alginate-producing mutants of *Pseudomonas fluorescens*, *Pseudomonas putida* and *Pseudomonas mendocina*. J Gen Microbiol **125:**217-220.
7. **Fett WF, Osman SF, Fishman ML, Siebles TS.** 1986. Alginate production by plant-pathogenic pseudomonads. Appl Environ Microbiol **52:**466-473.
8. **Osman SF, Fett WF, Fishman ML.** 1986. Exopolysaccharides of the phytopathogen *Pseudomonas syringae* pv. glycinea. J Bacteriol **166:**66-71.
9. **Cote GL, Krull LH.** 1988. Characterization of the exocellular polysaccharides from *Azotobacter chroococcum*. Carbohydr Res **181:**143-152.
10. **Straus DC, Lonon MK, Woods DE, Garner CW.** 1989. Production of an extracellular toxic complex by various strains of *Pseudomonas cepacia*. J Med Microbiol **30:**17-22.
11. **Likhosherstov LM, Senchenkova SN, Shashkov AS, Derevitskaya VA, Danilova IV, Botvinko IV.** 1991. Structure of the major exopolysaccharide produced by *Azotobacter beijerinckii* B-1615. Carbohydr Res **222:**233-238.
12. **Sadoff HL.** 1975. Encystment and germination in *Azotobacter vinelandii*. Bacteriol Rev **39:**516-539.
13. **Sabra W, Zeng AP, Lünsdorf H, Deckwer WD.** 2000. Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. Appl Environ Microbiol **66:**4037-4044.
14. **Halverson LJ.** 2009. Role of Alginate in Bacterial Biofilms, p. 135-151. *In* Rehm BHA (ed.), Alginates: Biology and Applications. Springer-Verlag Berlin, Heidelberg.
15. **Haug A, Larsen B, Smidsrød O.** 1966. A study of the constitution of alginic acid by partial acid hydrolysis. Acta Chem Scand **20:**183-190.
16. **Davidson IW, Sutherland IW, Lawson CJ.** 1977. Localization of *O*-acetyl groups of bacterial alginate. J Gen Microbiol **98:**603-606.
17. **Skjåk-Bræk G, Larsen B, Grasdalen H.** 1985. The role of *O*-acetyl groups in the biosynthesis of alginate by *Azotobacter vinelandii*. Carbohydr Res **145:**169-174.
18. **Bjerkan TM, Bender CL, Ertesvåg H, Drabløs F, Fakhr MK, Preston LA, Skjåk-Bræk G, Valla S.** 2004. The *Pseudomonas syringae* genome encodes a combined mannuronan C-5-epimerase and *O*-acetylhydrolase, which strongly

enhances the predicted gel-forming properties of alginates. J Biol Chem **279:**28920-28929.

19. **Chitnis CE, Ohman DE.** 1990. Cloning of *Pseudomonas aeruginosa algG*, which controls alginate structure. J Bacteriol **172:**2894-2900.
20. **Gimmestad M, Sletta H, Ertesvåg H, Bakkevig K, Jain S, Suh SJ, Skjåk-Bræk G, Ellingsen TE, Ohman DE, Valla S.** 2003. The *Pseudomonas fluorescens* AlgG protein, but not its mannuronan C-5-epimerase activity, is needed for alginate polymer formation. J Bacteriol **185:**3515-3523.
21. **Donati I, Paoletti S.** 2009. Material Properties of Alginates, p. 1-53. *In* Rehm BHA (ed.), Alginates: Biology and Applications. Springer-Verlag Berlin, Heidelberg.
22. **Smidsrød O, Haug A.** 1971. Estimation of the relative stiffness of the molecular chain in polyelectrolytes from measurements of viscosity at different ionic strengths. Biopolymers **10:**1213-1227.
23. **Smidsrød O, Haug A.** 1965. The effect of divalent metal ions on the properties of alginate solutions I. Calcium ions. Acta Chem Scand **19:**329-340.
24. **Haug A, Smidsrød O.** 1965. The effect of divalent metals on he properties of alginate solutions II. Comparison of different metal ions. Acta Chem Scand **19:**341-351.
25. **Grant GT, Morris ER, Rees DA, Smith PJC, Thom D.** 1973. Biological interactions between polysaccharides and divalent cations: The egg-box model. FEBS Letters **32:**195-198.
26. **Smidsrød O, Draget KI.** 1996. Chemistry and physical properties of alginates. Carbohydr Eur **14:**6-13.
27. **Skjåk-Bræk G, Zanetti F, Paoletti S.** 1989. Effect of acetylation on some solution and gelling properties of alginates. Carbohydr Res **185:**131-138.
28. **Draget KI, Skjåk-Bræk G, Smidsrød O.** 1994. Alginic acid gels: The effect of alginate chemical composition and molecular weight. Carbohydr Polym **25:**31-38.
29. **Onsøyen E.** 1996. Commercial applications of alginates. Carbohydr Eur **14:**26-31.
30. **Tønnesen HH, Karlsen J.** 2002. Alginate in drug delivery systems. Drug Dev Ind Pharm **28:**621-630.
31. **Smidsrød O, Skjåk-Bræk G.** 1990. Alginate as immobilization matrix for cells. Trends Biotechnol **8:**71-78.
32. **Chang PL.** 1997. Microcapsules as bio-organs for somatic gene therapy. Ann N Y Acad Sci **831:**461-473.
33. **Zimmermann H, Zimmermann D, Reuss R, Feilen PJ, Manz B, Katsen A, Weber M, Ihmig FR, Ehrhart F, Geßner P, Behringer M, Steinbach A, Wegner LH, Sukhorukov VL, Vásquez JA, Schneider S, Weber MM, Volke F, Wolf R, Zimmermann U.** 2005. Towards a medically approved technology for alginate-based microcapsules allowing long-term immunoisolated transplantation. J Mater Sci Mater Med **16:**491-501.

34. **Gutowska A, Jeong B, Jasionowski M.** 2001. Injectable gels for tissue engineering. Anat Rec **263:**342-349.
35. **Hunt NC, Grover LM.** 2010. Cell encapsulation using biopolymer gels for regenerative medicine. Biotechnol Lett **32:**733-742.
36. **Gandhi JK, Opara EC, Brey EM.** 2013. Alginate-based strategies for therapeutic vascularization. Ther Deliv **4:**327-341.
37. **Torre ML, Faustini M, Attilio KM, Vigo D.** 2007. Cell encapsulation in mammal reproduction. Recent Pat Drug Deliv Formul **1:**81-85.
38. **Kailasapathy K.** 2002. Microencapsulation of probiotic bacteria: technology and potential applications. Curr Issues Intest Microbiol **3:**39-48.
39. **Jain S, Mondal K, Gupta MN.** 2006. Applications of alginate in bioseparation of proteins. Artif Cells Blood Substit Immobil Biotechnol **34:**127-144.
40. **Indergaard M, Skjåk-Bræk G.** 1987. Characteristics of alginate from *Laminaria digitata* cultivated in a high-phosphate environment. Hydrobiologia **151-152:**541-549.
41. **Draget KI, Moe ST, Skjåk-Bræk G, Smidsrød O.** 2006. Alginates, p. 289-334. *In* Stephen AM, Phillips GO, Williams PA (ed.), Food Polysaccharides and Their Applications, 2 ed. CRC Press, Boca Raton.
42. **Valla S, Ertesvåg H, Skjåk-Bræk G.** 1996. Genetics and biosynthesis of alginates. Carbohydrates in Europe **14:**14-18.
43. **Rehm BHA.** 2010. Bacterial polymers: biosynthesis, modifications and applications. Nat Rev Micro **8:**578-592.
44. **Rehm BHA, Valla S.** 1997. Bacterial alginates: biosynthesis and applications. Appl Microbiol Biotechnol **48:**281-288.
45. **Salzig M, Rehm BHA.** 2006. Microbial production of alginates: self-assembly and applications, p. 125-152. *In* Rehm BHA (ed.), Microbial Bionanotechnology: Biological Self-assembly Systems and Biopolymer-based Nanostructures. Horizon Bioscience, Norfolk.
46. **Ertesvåg H, Høidal HK, Hals IK, Rian A, Doseth B, Valla S.** 1995. A family of modular type mannuronan C-5-epimerase genes controls alginate structure in *Azotobacter vinelandii*. Mol Microbiol **16:**719-731.
47. **Svanem BIG, Skjåk-Bræk G, Ertesvåg H, Valla S.** 1999. Cloning and expression of three new *Azotobacter vinelandii* genes closely related to a previously described gene family encoding mannuronan C-5-epimerases. J Bacteriol **181:**68-77.
48. **Gimmestad M, Steigedal M, Ertesvåg H, Moreno S, Christensen BE, Espín G, Valla S.** 2006. Identification and characterization of an *Azotobacter vinelandii* type I secretion system responsible for export of the AlgE-type mannuronan C-5-epimerases. J Bacteriol **188:**5551-5560.
49. **Lebrun L, Junter GA, Jouenne T, Mignot L.** 1994. Exopolysaccharide production by free and immobilized microbial cultures. Enzyme Microb Technol **16:**1048-1054.

50. **Clementi F, Fantozzi P, Mancini F, Moresi M.** 1995. Optimal conditions for alginate production by *Azotobacter vinelandii*. Enzyme Microb Technol **17:**983-988.
51. **Sabra W, Zeng A-P, Deckwer W-D.** 2001. Bacterial alginate: physiology, product quality and process aspects. Appl Microbiol Biotechnol **56:**315-325.
52. **Priego-Jimenéz R, Peña C, Ramírez OT, Galindo E.** 2005. Specific growth rate determines the molecular mass of the alginate produced by *Azotobacter vinelandii*. Biochem Eng J **25:**187-193.
53. **Galindo E, Peña C, Núñez C, Segura D, Espín G.** 2007. Molecular and bioengineering strategies to improve alginate and polydydroxyalkanoate production by *Azotobacter vinelandii*. Microb Cell Fact **6:**7.
54. **Díaz-Barrera A, Soto E, Altamirano C.** 2012. Alginate production and *alg8* gene expression by *Azotobacter vinelandii* in continuous cultures. J Ind Microbiol Biotechnol **39:**613-621.
55. **May TB, Chakrabarty AM.** 1994. *Pseudomonas aeruginosa*: genes and enzymes of alginate synthesis. Trends Microbiol **2:**151-157.
56. **Rehm BH, Ertesvåg H, Valla S.** 1996. A new *Azotobacter vinelandii mannuronan* C-5-epimerase gene (*algG*) is part of an *alg* gene cluster physically organized in a manner similar to that in *Pseudomonas aeruginosa*. J Bacteriol **178:**5884-5889.
57. **Mejía-Ruíz H, Moreno S, Guzmán J, Nájera R, León R, Soberón-Chávez G, Espín G.** 1997. Isolation and characterization of an *Azotobacter vinelandii algK* mutant. FEMS Microbiol Lett **156:**101-106.
58. **Vazquez A, Moreno S, Guzmán J, Alvarado A, Espín G.** 1999. Transcriptional organization of the *Azotobacter vinelandii algGXLVIFA* genes: characterization of *algF* mutants. Gene **232:**217-222.
59. **Peñaloza-Vázquez A, Kidambi SP, Chakrabarty AM, Bender CL.** 1997. Characterization of the alginate biosynthetic gene cluster in *Pseudomonas syringae* pv. syringae. J Bacteriol **179:**4464-4472.
60. **Morea A, Mathee K, Franklin MJ, Giacomini A, O'Regan M, Ohman DE.** 2001. Characterization of *algG* encoding C5-epimerase in the alginate biosynthetic gene cluster of *Pseudomonas fluorescens*. Gene **278:**107-114.
61. **Winsor GL, Lam DK, Fleming L, Lo R, Whiteside MD, Yu NY, Hancock RE, Brinkman FS.** 2011. *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. Nucleic Acids Res **39:**D596-600.
62. **Lessie TG, Phibbs PV, Jr.** 1984. Alternative pathways of carbohydrate utilization in pseudomonads. Annu Rev Microbiol **38:**359-387.
63. **Lynn AR, Sokatch JR.** 1984. Incorporation of isotope from specifically labeled glucose into alginates of *Pseudomonas aeruginosa* and *Azotobacter vinelandii*. J Bacteriol **158:**1161-1162.
64. **Banerjee PC, Vanags RI, Chakrabarty AM, Maitra PK.** 1985. Fructose 1,6-bisphosphate aldolase activity is essential for synthesis of alginate from glucose by *Pseudomonas aeruginosa*. J Bacteriol **161:**458-460.

65. **Anderson AJ, Hacking AJ, Dawes EA.** 1987. Alternative pathways for the biosynthesis of alginate from fructose and glucose in *Pseudomonas mendocina* and *Azotobacter vinelandii*. J Gen Microbiol **133:**1045-1052.
66. **Narbad A, Russell NJ, Gacesa P.** 1988. Radiolabelling patterns in alginate of *Pseudomonas aeruginosa* synthesized from specifically-labelled 14C-monosaccharide precursors. Microbios **54:**171-179.
67. **Beale JM, Foster JL.** 1996. Carbohydrate fluxes into alginate biosynthesis in *Azotobacter vinelandii* NCIB 8789: NMR investigations of the triose pool. Biochemistry **35:**4492-4501.
68. **Narbad A, Hewlins MJ, Gacesa P, Russell NJ.** 1990. The use of 13C-n.m.r. spectroscopy to monitor alginate biosynthesis in mucoid *Pseudomonas aeruginosa*. Biochem J **267:**579-584.
69. **Pindar DF, Bucke C.** 1975. The biosynthesis of alginic acid by *Azotobacter vinelandii*. Biochem J **152:**617-622.
70. **Piggott NH, Sutherland IW, Jarman TR.** 1981. Enzymes involved in the biosynthesis of alginate by *Pseudomonas aeruginosa*. Eur J Appl Microbiol Biotechnol **13:**179-183.
71. **Roychoudhury S, May TB, Gill JF, Singh SK, Feingold DS, Chakrabarty AM.** 1989. Purification and characterization of guanosine diphospho-D-mannose dehydrogenase. A key enzyme in the biosynthesis of alginate by *Pseudomonas aeruginosa*. J Biol Chem **264:**9380-9385.
72. **Campos M, Martínez-Salazar JM, Lloret L, Moreno S, Núñez C, Espín G, Soberón-Chávez G.** 1996. Characterization of the gene coding for GDP-mannose dehydrogenase (*algD*) from *Azotobacter vinelandii*. J Bacteriol **178:**1793-1799.
73. **Shinabarger D, Berry A, May TB, Rothmel R, Fialho A, Chakrabarty AM.** 1991. Purification and characterization of phosphomannose isomerase-guanosine diphospho-D-mannose pyrophosphorylase. A bifunctional enzyme in the alginate biosynthetic pathway of *Pseudomonas aeruginosa*. J Biol Chem **266:**2080-2088.
74. **Lloret L, Barreto R, León R, Moreno S, Martínez-Salazar J, Espín G, Soberón-Chávez G.** 1996. Genetic analysis of the transcriptional arrangement of *Azotobacter vinelandii* alginate biosynthetic genes: identification of two independent promoters. Mol Microbiol **21:**449-457.
75. **Zielinski NA, Chakrabarty AM, Berry A.** 1991. Characterization and regulation of the *Pseudomonas aeruginosa algC* gene encoding phosphomannomutase. J Biol Chem **266:**9754-9763.
76. **Gaona G, Núñez C, Goldberg JB, Linford AS, Nájera R, Castañeda M, Guzmán J, Espín G, Soberón-Chávez G.** 2004. Characterization of the *Azotobacter vinelandii algC* gene involved in alginate and lipopolysaccharide production. FEMS Microbiol Lett **238:**199-206.
77. **Steigedal M.** 2006. The *Azotobacter vinelandii* mannuronan C5-epimerases: their biological functions and new tools useful for their future *in vivo*

biotechnological application. Doctoral thesis. Norwegian University of Science and Technology, Trondheim.

78. **Hay ID, Ur Rehman Z, Ghafoor A, Rehm BHA.** 2010. Bacterial biosynthesis of alginates. J Chem Technol Biotechnol **85:**752-759.
79. **Mejía-Ruíz H, Guzmán J, Moreno S, Soberón-Chávez G, Espín G.** 1997. The *Azotobacter vinelandii alg8* and *alg44* genes are essential for alginate synthesis and can be transcribed from an *algD*-independent promoter. Gene **199:**271-277.
80. **Remminghorst U, Rehm BH.** 2006. Alg44, a unique protein required for alginate biosynthesis in *Pseudomonas aeruginosa*. FEBS Lett **580:**3883-3888.
81. **Remminghorst U, Rehm BH.** 2006. In vitro alginate polymerization and the functional role of Alg8 in alginate production by *Pseudomonas aeruginosa*. Appl Environ Microbiol **72:**298-305.
82. **Merighi M, Lee VT, 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*. Mol Microbiol **65:**876-895.
83. **Oglesby LL, Jain S, Ohman DE.** 2008. Membrane topology and roles of *Pseudomonas aeruginosa* Alg8 and Alg44 in alginate polymerization. Microbiology **154:**1605-1615.
84. **Franklin MJ, Ohman DE.** 1993. Identification of *algF* in the alginate biosynthetic gene cluster of *Pseudomonas aeruginosa* which is required for alginate acetylation. J Bacteriol **175:**5057-5065.
85. **Franklin MJ, Ohman DE.** 1996. Identification of *algI* and *algJ* in the *Pseudomonas aeruginosa* alginate biosynthetic gene cluster which are required for alginate O acetylation. J Bacteriol **178:**2186-2195.
86. **Franklin MJ, Ohman DE.** 2002. Mutant analysis and cellular localization of the AlgI, AlgJ, and AlgF proteins required for O acetylation of alginate in *Pseudomonas aeruginosa*. J Bacteriol **184:**3000-3007.
87. **Franklin MJ, Chitnis CE, Gacesa P, Sonesson A, White DC, Ohman DE.** 1994. *Pseudomonas aeruginosa* AlgG is a polymer level alginate C5-mannuronan epimerase. J Bacteriol **176:**1821-1830.
88. **Schiller NL, Monday SR, Boyd CM, Keen NT, Ohman DE.** 1993. Characterization of the *Pseudomonas aeruginosa* alginate lyase gene (*algL*): cloning, sequencing, and expression in *Escherichia coli*. J Bacteriol **175:**4780-4789.
89. **Ertesvåg H, Erlien F, Skjåk-Bræk G, Rehm BHA, Valla S.** 1998. Biochemical properties and substrate specificities of a recombinantly produced *Azotobacter vinelandii* alginate lyase. J Bacteriol **180:**3779-3784.
90. **Bakkevig K, Sletta H, Gimmestad M, Aune R, Ertesvåg H, Degnes K, Christensen BE, Ellingsen TE, Valla S.** 2005. Role of the *Pseudomonas fluorescens* alginate lyase (AlgL) in clearing the periplasm of alginates not exported to the extracellular environment. J Bacteriol **187:**8375-8384.
91. **Jain S, Ohman DE.** 2005. Role of an alginate lyase for alginate transport in mucoid *Pseudomonas aeruginosa*. Infect Immun **73:**6429-6436.

92. **Rehman ZU, Wang Y, Moradali MF, Hay ID, Rehm BH.** 2013. Insights into the assembly of the alginate biosynthesis machinery in *Pseudomonas aeruginosa*. Appl Environ Microbiol **79:**3264-3272.
93. **Rehm BH.** 1996. The *Azotobacter vinelandii* gene *algJ* encodes an outer-membrane protein presumably involved in export of alginate. Microbiology **142:**873-880.
94. **Jain S, Ohman DE.** 1998. Deletion of *algK* in mucoid *Pseudomonas aeruginosa* blocks alginate polymer formation and results in uronic acid secretion. J Bacteriol **180:**634-641.
95. **Jain S, Franklin MJ, Ertesvåg H, Valla S, Ohman DE.** 2003. The dual roles of AlgG in C-5-epimerization and secretion of alginate polymers in *Pseudomonas aeruginosa*. Mol Microbiol **47:**1123-1133.
96. **Robles-Price A, Wong TY, Sletta H, Valla S, Schiller NL.** 2004. AlgX is a periplasmic protein required for alginate biosynthesis in *Pseudomonas aeruginosa*. J Bacteriol **186:**7369-7377.
97. **Albrecht MT, Schiller NL.** 2005. Alginate lyase (AlgL) activity is required for alginate biosynthesis in *Pseudomonas aeruginosa*. J Bacteriol **187:**3869-3872.
98. **Hay ID, Rehman ZU, Rehm BH.** 2010. Membrane topology of outer membrane protein AlgE, which is required for alginate production in *Pseudomonas aeruginosa*. Appl Environ Microbiol **76:**1806-1812.
99. **Rehm BH, Boheim G, Tommassen J, Winkler UK.** 1994. Overexpression of *algE* in *Escherichia coli*: subcellular localization, purification, and ion channel properties. J Bacteriol **176:**5639-5647.
100. **Ertesvåg H, Doseth B, Larsen B, Skjåk-Bræk G, Valla S.** 1994. Cloning and expression of an *Azotobacter vinelandii* mannuronan C-5-epimerase gene. J Bacteriol **176:**2846-2853.
101. **Steigedal M, Sletta H, Moreno S, Mærk M, Christensen BE, Bjerkan T, Ellingsen TE, Espín G, Ertesvåg H, Valla S.** 2008. The *Azotobacter vinelandii* AlgE mannuronan C-5-epimerase family is essential for the in vivo control of alginate monomer composition and for functional cyst formation. Environ Microbiol **10:**1760-1770.
102. **Chitnis CE, Ohman DE.** 1993. Genetic analysis of the alginate biosynthetic gene cluster of *Pseudomonas aeruginosa* shows evidence of an operonic structure. Mol Microbiol **8:**583-593.
103. **Paletta JL, Ohman DE.** 2012. Evidence for two promoters internal to the alginate biosynthesis operon in *Pseudomonas aeruginosa*. Curr Microbiol **65:**770-775.
104. **Martin DW, Holloway BW, Deretic V.** 1993. Characterization of a locus determining the mucoid status of *Pseudomonas aeruginosa*: AlgU shows sequence similarities with a *Bacillus* sigma factor. J Bacteriol **175:**1153-1164.
105. **Deretic V, Schurr MJ, Boucher JC, Martin DW.** 1994. Conversion of *Pseudomonas aeruginosa* to mucoidy in cystic fibrosis: environmental stress and regulation of bacterial virulence by alternative sigma factors. J Bacteriol **176:**2773-2780.

106. **Wozniak DJ, Ohman DE.** 1994. Transcriptional analysis of the *Pseudomonas aeruginosa* genes *algR*, *algB*, and *algD* reveals a hierarchy of alginate gene expression which is modulated by *algT*. J Bacteriol **176:**6007-6014.
107. **Hershberger CD, Ye RW, Parsek MR, Xie ZD, Chakrabarty AM.** 1995. The *algT* (*algU*) gene of *Pseudomonas aeruginosa*, a key regulator involved in alginate biosynthesis, encodes an alternative σ factor (σ^E). Proc Natl Acad Sci U S A **92:**7941-7945.
108. **Deretic V, Dikshit R, Konyecsni WM, Chakrabarty AM, Misra TK.** 1989. The *algR* gene, which regulates mucoidy in *Pseudomonas aeruginosa*, belongs to a class of environmentally responsive genes. J Bacteriol **171:**1278-1283.
109. **Kato J, Chakrabarty AM.** 1991. Purification of the regulatory protein AlgR1 and its binding in the far upstream region of the *algD* promoter in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **88:**1760-1764.
110. **Baynham PJ, Wozniak DJ.** 1996. Identification and characterization of AlgZ, an AlgT-dependent DNA-binding protein required for *Pseudomonas aeruginosa algD* transcription. Mol Microbiol **22:**97-108.
111. **Baynham PJ, Brown AL, Hall LL, Wozniak DJ.** 1999. *Pseudomonas aeruginosa* AlgZ, a ribbon-helix-helix DNA-binding protein, is essential for alginate synthesis and *algD* transcriptional activation. Mol Microbiol **33:**1069-1080.
112. **Goldberg JB, Ohman DE.** 1987. Construction and characterization of *Pseudomonas aeruginosa algB* mutants: role of *algB* in high-level production of alginate. J Bacteriol **169:**1593-1602.
113. **Goldberg JB, Dahnke T.** 1992. *Pseudomonas aeruginosa* AlgB, which modulates the expression of alginate, is a member of the NtrC subclass of prokaryotic regulators. Mol Microbiol **6:**59-66.
114. **Martin DW, Schurr MJ, Yu H, Deretic V.** 1994. Analysis of promoters controlled by the putative sigma factor AlgU regulating conversion to mucoidy in *Pseudomonas aeruginosa*: relationship to σ^E and stress response. J Bacteriol **176:**6688-6696.
115. **DeVries CA, Ohman DE.** 1994. Mucoid-to-nonmucoid conversion in alginate-producing *Pseudomonas aeruginosa* often results from spontaneous mutations in *algT*, encoding a putative alternate sigma factor, and shows evidence for autoregulation. J Bacteriol **176:**6677-6687.
116. **Zielinski NA, Maharaj R, Roychoudhury S, Danganan CE, Hendrickson W, Chakrabarty AM.** 1992. Alginate synthesis in *Pseudomonas aeruginosa*: environmental regulation of the *algC* promoter. J Bacteriol **174:**7680-7688.
117. **Deretic V, Konyecsni WM.** 1989. Control of mucoidy in *Pseudomonas aeruginosa*: transcriptional regulation of *algR* and identification of the second regulatory gene, *algQ*. J Bacteriol **171:**3680-3688.
118. **Kato J, Misra TK, Chakrabarty AM.** 1990. AlgR3, a protein resembling eukaryotic histone H1, regulates alginate synthesis in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **87:**2887-2891.
119. **Konyecsni WM, Deretic V.** 1990. DNA sequence and expression analysis of *algP* and *algQ*, components of the multigene system transcriptionally regulating

mucoidy in *Pseudomonas aeruginosa*: *algP* contains multiple direct repeats. J Bacteriol **172:**2511-2520.

120. **Martínez-Salazar JM, Moreno S, Nájera R, Boucher JC, Espín G, Soberón-Chávez G, Deretic V.** 1996. Characterization of the genes coding for the putative sigma factor AlgU and its regulators MucA, MucB, MucC, and MucD in *Azotobacter vinelandii* and evaluation of their roles in alginate biosynthesis. J Bacteriol **178:**1800-1808.
121. **Moreno S, Nájera R, Guzmán J, Soberón-Chávez G, Espín G.** 1998. Role of alternative σ factor *algU* in encystment of *Azotobacter vinelandii*. J Bacteriol **180:**2766-2769.
122. **Castañeda M, Guzmán J, Moreno S, Espín G.** 2000. The GacS sensor kinase regulates alginate and poly-β-hydroxybutyrate production in *Azotobacter vinelandii*. J Bacteriol **182:**2624-2628.
123. **Núñez C, León R, Guzmán J, Espín G, Soberón-Chávez G.** 2000. Role of *Azotobacter vinelandii mucA* and *mucC* gene products in alginate production. J Bacteriol **182:**6550-6556.
124. **Núñez C, Moreno S, Soberón-Chávez G, Espín G.** 1999. The *Azotobacter vinelandii* response regulator AlgR is essential for cyst formation. J Bacteriol **181:**141-148.
125. **Kato J, Chu L, Kitano K, DeVault JD, Kimbara K, Chakrabarty AM, Misra TK.** 1989. Nucleotide sequence of a regulatory region controlling alginate synthesis in *Pseudomonas aeruginosa*: characterization of the *algR2* gene. Gene **84:**31-38.
126. **Yin Y, Withers TR, Wang X, Yu HD.** 2013. Evidence for sigma factor competition in the regulation of alginate production by *Pseudomonas aeruginosa*. PLoS One **8:**e72329.
127. **Wozniak DJ, Ohman DE.** 1991. Pseudomonas aeruginosa AlgB, a two-component response regulator of the NtrC family, is required for algD transcription. J Bacteriol **173:**1406-1413.
128. **Qiu D, Eisinger VM, Rowen DW, Yu HD.** 2007. Regulated proteolysis controls mucoid conversion in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A **104:**8107-8112.
129. **Cezairliyan BO, Sauer RT.** 2009. Control of *Pseudomonas aeruginosa* AlgW protease cleavage of MucA by peptide signals and MucB. Mol Microbiol **72:**368-379.
130. **Wood LF, Ohman DE.** 2009. Use of cell wall stress to characterize σ^{22} (AlgT/U) activation by regulated proteolysis and its regulon in *Pseudomonas aeruginosa*. Mol Microbiol **72:**183-201.
131. **Damron FH, Yu HD.** 2011. *Pseudomonas aeruginosa* MucD regulates the alginate pathway through activation of MucA degradation via MucP proteolytic activity. J Bacteriol **193:**286-291.
132. **Qiu D, Eisinger VM, Head NE, Pier GB, Yu HD.** 2008. ClpXP proteases positively regulate alginate overexpression and mucoid conversion in *Pseudomonas aeruginosa*. Microbiology **154:**2119-2130.

133. **Wozniak DJ, Ohman DE.** 1993. Involvement of the alginate *algT* gene and integration host factor in the regulation of the *Pseudomonas aeruginosa algB* gene. J Bacteriol **175:**4145-4153.
134. **Delic-Attree I, Toussaint B, Froger A, Willison JC, Vignais PM.** 1996. Isolation of an IHF-deficient mutant of a *Pseudomonas aeruginosa* mucoid isolate and evaluation of the role of IHF in *algD* gene expression. Microbiology **142:**2785-2793.
135. **Toussaint B, Delic-Attree I, Vignais PM.** 1993. *Pseudomonas aeruginosa* contains an IHF-like protein that binds to the *algD* promoter. Biochem Biophys Res Commun **196:**416-421.
136. **Schlictman D, Kubo M, Shankar S, Chakrabarty AM.** 1995. Regulation of nucleoside diphosphate kinase and secretable virulence factors in *Pseudomonas aeruginosa*: roles of *algR2* and *algH*. J Bacteriol **177:**2469-2474.
137. **Delic-Attree I, Toussaint B, Garin J, Vignais PM.** 1997. Cloning, sequence and mutagenesis of the structural gene of *Pseudomonas aeruginosa* CysB, which can activate *algD* transcription. Mol Microbiol **24:**1275-1284.
138. **Boucher JC, Schurr MJ, Deretic V.** 2000. Dual regulation of mucoidy in *Pseudomonas aeruginosa* and sigma factor antagonism. Mol Microbiol **36:**341-351.
139. **Damron FH, Qiu D, Yu HD.** 2009. The *Pseudomonas aeruginosa* sensor kinase KinB negatively controls alginate production through AlgW-dependent MucA proteolysis. J Bacteriol **191:**2285-2295.
140. **Reiling SA, Jansen JA, Henley BJ, Singh S, Chattin C, Chandler M, Rowen DW.** 2005. Prc protease promotes mucoidy in *mucA* mutants of *Pseudomonas aeruginosa*. Microbiology **151:**2251-2261.
141. **Hay ID, Remminghorst U, Rehm BH.** 2009. MucR, a novel membrane-associated regulator of alginate biosynthesis in *Pseudomonas aeruginosa*. Appl Environ Microbiol **75:**1110-1120.
142. **Manzo J, Cocotl-Yañez M, Tzontecomani T, Martinez VM, Bustillos R, Velásquez C, Goiz Y, Solís Y, López L, Fuentes LE, Nuñez C, Segura D, Espín G, Castañeda M.** 2011. Post-transcriptional regulation of the alginate biosynthetic gene *algD* by the Gac/Rsm system in *Azotobacter vinelandii*. J Mol Microbiol Biotechnol **21:**147-159.
143. **Suh SJ, Silo-Suh L, Woods DE, Hassett DJ, West SE, Ohman DE.** 1999. Effect of *rpoS* mutation on the stress response and expression of virulence factors in *Pseudomonas aeruginosa*. J Bacteriol **181:**3890-3897.
144. **Schuster M, Hawkins AC, Harwood CS, Greenberg EP.** 2004. The *Pseudomonas aeruginosa* RpoS regulon and its relationship to quorum sensing. Mol Microbiol **51:**973-985.
145. **Castañeda M, Sánchez J, Moreno S, Núñez C, Espín G.** 2001. The global regulators GacA and σ^S form part of a cascade that controls alginate production in *Azotobacter vinelandii*. J Bacteriol **183:**6787-6793.
146. **Schurr MJ, Yu H, Martinez-Salazar JM, Boucher JC, Deretic V.** 1996. Control of AlgU, a member of the σ^E-like family of stress sigma factors, by the negative

regulators MucA and MucB and *Pseudomonas aeruginosa* conversion to mucoidy in cystic fibrosis. J Bacteriol **178:**4997-5004.

147. **Xie ZD, Hershberger CD, Shankar S, Ye RW, Chakrabarty AM.** 1996. Sigma factor-anti-sigma factor interaction in alginate synthesis: inhibition of AlgT by MucA. J Bacteriol **178:**4990-4996.
148. **Boucher JC, Schurr MJ, Yu H, Rowen DW, Deretic V.** 1997. *Pseudomonas aeruginosa* in cystic fibrosis: role of *mucC* in the regulation of alginate production and stress sensitivity. Microbiology **143:**3473-3480.
149. **Boucher JC, Martinez-Salazar J, Schurr MJ, Mudd MH, Yu H, Deretic V.** 1996. Two distinct loci affecting conversion to mucoidy in *Pseudomonas aeruginosa* in cystic fibrosis encode homologs of the serine protease HtrA. J Bacteriol **178:**511-523.
150. **Balasubramanian D, Kong KF, Jayawardena SR, Leal SM, Sautter RT, Mathee K.** 2011. Co-regulation of β-lactam resistance, alginate production and quorum sensing in *Pseudomonas aeruginosa*. J Med Microbiol **60:**147-156.
151. **Martin DW, Schurr MJ, Mudd MH, Govan JR, Holloway BW, Deretic V.** 1993. Mechanism of conversion to mucoidy in *Pseudomonas aeruginosa* infecting cystic fibrosis patients. Proc Natl Acad Sci U S A **90:**8377-8381.
152. **Martin DW, Schurr MJ, Mudd MH, Deretic V.** 1993. Differentiation of *Pseudomonas aeruginosa* into the alginate-producing form: inactivation of *mucB* causes conversion to mucoidy. Mol Microbiol **9:**497-506.
153. **Mathee K, McPherson CJ, Ohman DE.** 1997. Posttranslational control of the *algT* (*algU*)-encoded σ^{22} for expression of the alginate regulon in *Pseudomonas aeruginosa* and localization of its antagonist proteins MucA and MucB (AlgN). J Bacteriol **179:**3711-3720.
154. **Damron FH, Goldberg JB.** 2012. Proteolytic regulation of alginate overproduction in *Pseudomonas aeruginosa*. Mol Microbiol **84:**595-607.
155. **Wood LF, Leech AJ, Ohman DE.** 2006. Cell wall-inhibitory antibiotics activate the alginate biosynthesis operon in *Pseudomonas aeruginosa*: Roles of σ^{22} (AlgT) and the AlgW and Prc proteases. Mol Microbiol **62:**412-426.
156. **Setubal JC, dos Santos P, Goldman BS, Ertesvåg H, Espìn G, Rubio LM, Valla S, Almeida NF, Balasubramanian D, Cromes L, Curatti L, Du Z, Godsy E, Goodner B, Hellner-Burris K, Hernandez JA, Houmiel K, Imperial J, Kennedy C, Larson TJ, Latreille P, Ligon LS, Lu J, Mærk M, Miller NM, Norton S, O'Carroll IP, Paulsen I, Raulfs EC, Roemer R, Rosser J, Segura D, Slater S, Stricklin SL, Studholme DJ, Sun J, Viana CJ, Wallin E, Wang B, Wheeler C, Zhu H, Dean DR, Dixon R, Wood D.** 2009. Genome sequence of *Azotobacter vinelandii*, an obligate aerobe specialized to support diverse anaerobic metabolic processes. J Bacteriol **191:**4534-4545.
157. **Liao CH, McCallus DE, Wells JM, Tzean SS, Kang GY.** 1996. The *repB* gene required for production of extracellular enzymes and fluorescent siderophores in *Pseudomonas viridiflava* is an analog of the *gacA* gene of *Pseudomonas syringae*. Can J Microbiol **42:**177-182.

158. **Willis DK, Holmstadt JJ, Kinscherf TG.** 2001. Genetic evidence that loss of virulence associated with *gacS* or *gacA* mutations in *Pseudomonas syringae* B728a does not result from effects on alginate production. Appl Environ Microbiol **67:**1400-1403.
159. **Parkins MD, Ceri H, Storey DG.** 2001. *Pseudomonas aeruginosa* GacA, a factor in multihost virulence, is also essential for biofilm formation. Mol Microbiol **40:**1215-1226.
160. **Núñez C, Moreno S, Cárdenas L, Soberón-Chávez G, Espín G.** 2000. Inactivation of the *ampDE* operon increases transcription of *algD* and affects morphology and encystment of *Azotobacter vinelandii*. J Bacteriol **182:**4829-4835.
161. **Ambrosi C, Tiburzi F, Imperi F, Putignani L, Visca P.** 2005. Involvement of AlgQ in transcriptional regulation of pyoverdine genes in *Pseudomonas aeruginosa* PAO1. J Bacteriol **187:**5097-5107.
162. **Baynham PJ, Ramsey DM, Gvozdyev BV, Cordonnier EM, Wozniak DJ.** 2006. The *Pseudomonas aeruginosa* ribbon-helix-helix DNA-binding protein AlgZ (AmrZ) controls twitching motility and biogenesis of type IV pili. J Bacteriol **188:**132-140.
163. **Galán-Vásquez E, Luna B, Martínez-Antonio A.** 2011. The regulatory network of *Pseudomonas aeruginosa*. Microb Inform Exp **1:**3.
164. **Jones CJ, Ryder CR, Mann EE, Wozniak DJ.** 2013. AmrZ modulates *Pseudomonas aeruginosa* biofilm architecture by directly repressing transcription of the *psl* operon. J Bacteriol **195:**1637-1644.
165. **Kennedy C, Rudnick P, MacDonald ML, Melton T.** 2005. Genus III. *Azotobacter* Beijerinck 1901, 567[AL], p. 384-402. *In* Brenner DJ, Krieg NR, Staley JT, Garrity GM (ed.), Bergey's manual of systematic bacteriology. Volume Two: The Proteobacteria, 2 ed. Springer, New York.
166. **Poole RK, Hill S.** 1997. Respiratory protection of nitrogenase activity in *Azotobacter vinelandii*—roles of the terminal oxidases. Biosci Rep **17:**303-317.
167. **Bertsova YV, Bogachev AV, Skulachev VP.** 2001. Noncoupled NADH:ubiquinone oxidoreductase of *Azotobacter vinelandii* is required for diazotrophic growth at high oxygen concentrations. J Bacteriol **183:**6869-6874.
168. **Moshiri F, Kim JW, Fu C, Maier RJ.** 1994. The FeSII protein of *Azotobacter vinelandii* is not essential for aerobic nitrogen fixation, but confers significant protection to oxygen-mediated inactivation of nitrogenase *in vitro* and *in vivo*. Mol Microbiol **14:**101-114.
169. **Thorneley RN, Ashby GA.** 1989. Oxidation of nitrogenase iron protein by dioxygen without inactivation could contribute to high respiration rates of *Azotobacter* species and facilitate nitrogen fixation in other aerobic environments. Biochem J **261:**181-187.
170. **Oelze J.** 2000. Respiratory protection of nitrogenase in *Azotobacter* species: is a widely held hypothesis unequivocally supported by experimental evidence? FEMS Microbiol Rev **24:**321-333.
171. **Socolofsky MD, Wyss O.** 1962. Resistance of the *Azotobacter* cyst. J Bacteriol **84:**119-124.

172. **Ertesvåg H, Høidal HK, Schjerven H, Svanem BIG, Valla S.** 1999. Mannuronan C-5-epimerases and their application for *in vitro* and *in vivo* design of new alginates useful in biotechnology. Metab Eng **1:**262-269.
173. **Hoidal HK, Glaerum Svanem BI, Gimmestad M, Valla S.** 2000. Mannuronan C-5 epimerases and cellular differentiation of Azotobacter vinelandii. Environ Microbiol **2:**27-38.
174. **Maldonado R, Jiménez J, Casadesús J.** 1994. Changes of ploidy during the *Azotobacter vinelandii* growth cycle. J Bacteriol **176:**3911-3919.
175. **Rediers H, Vanderleyden J, De Mot R.** 2004. *Azotobacter vinelandii*: a *Pseudomonas* in disguise? Microbiology **150:**1117-1119.
176. **Özen AI, Ussery DW.** 2012. Defining the *Pseudomonas* genus: where do we draw the line with *Azotobacter*? Microb Ecol **63:**239-248.
177. **Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, McKenney K, Sutton G, FitzHugh W, Fields C, Gocayne JD, Scott J, Shirley R, Liu L-I, Glodek A, Kelley JM, Weidman JF, Phillips CA, Spriggs T, Hedblom E, Cotton MD, Utterback TR, Hanna MC, Nguyen DT, Saudek DM, Brandon RC, Fine LD, Fritchman JL, Fuhrmann JL, Geoghagen NSM, Gnehm CL, McDonald LA, Small KV, Fraser CM, Smith HO, Venter JC.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496-512.
178. **MacLean D, Jones JD, Studholme DJ.** 2009. Application of 'next-generation' sequencing technologies to microbial genetics. Nat Rev Microbiol **7:**287-296.
179. **Loman NJ, Constantinidou C, Chan JZ, Halachev M, Sergeant M, Penn CW, Robinson ER, Pallen MJ.** 2012. High-throughput bacterial genome sequencing: an embarrassment of choice, a world of opportunity. Nat Rev Microbiol **10:**599-606.
180. **Pagani I, Liolios K, Jansson J, Chen IM, Smirnova T, Nosrat B, Markowitz VM, Kyrpides NC.** 2012. The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res **40:**D571-579.
181. **Bentley SD, Parkhill J.** 2004. Comparative genomic structure of prokaryotes. Annu Rev Genet **38:**771-792.
182. **Gregory TR, DeSalle R.** 2005. Comparative genomics in prokaryotes, p. 586-675. *In* Gregory TR (ed.), The Evolution of the Genome. Elsevier Academic Press, New York.
183. **Hardison RC.** 2003. Comparative genomics. PLoS Biol **1:**E58.
184. **Binnewies TT, Motro Y, Hallin PF, Lund O, Dunn D, La T, Hampson DJ, Bellgard M, Wassenaar TM, Ussery DW.** 2006. Ten years of bacterial genome sequencing: comparative-genomics-based discoveries. Funct Integr Genomics **6:**165-185.
185. **Abby S, Daubin V.** 2007. Comparative genomics and the evolution of prokaryotes. Trends Microbiol **15:**135-141.
186. **Lockhart DJ, Winzeler EA.** 2000. Genomics, gene expression and DNA arrays. Nature **405:**827-836.

187. **Sorek R, Cossart P.** 2010. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. Nat Rev Genet **11:**9-16.
188. **Gagarinova A, Emili A.** 2012. Genome-scale genetic manipulation methods for exploring bacterial molecular biology. Mol Biosyst **8:**1626-1638.
189. **Dove A.** 2007. High-throughput screening goes to school. Nat Meth **4:**523-532.
190. **Kong F, Yuan L, Zheng YF, Chen W.** 2012. Automatic liquid handling for life science: a critical review of the current state of the art. J Lab Autom **17:**169-185.
191. **Yadav VG, De Mey M, Lim CG, Ajikumar PK, Stephanopoulos G.** 2012. The future of metabolic engineering and synthetic biology: towards a systematic practice. Metab Eng **14:**233-241.
192. **Esvelt KM, Wang HH.** 2013. Genome-scale engineering for systems and synthetic biology. Mol Syst Biol **9:**641.
193. **Noar JD, Bruno-Bárcena JM.** 2013. Complete genome sequences of *Azotobacter vinelandii* wild-type strain CA and tungsten-tolerant mutant strain CA6. Genome Announc **1:**e00313-00313.
194. **Nelson KE, Weinel C, Paulsen IT, Dodson RJ, Hilbert H, dos Santos VAPM, Fouts DE, Gill SR, Pop M, Holmes M, Brinkac L, Beanan M, DeBoy RT, Daugherty S, Kolonay J, Madupu R, Nelson W, White O, Peterson J, Khouri H, Hance I, Lee PC, Holtzapple E, Scanlan D, Tran K, Moazzez A, Utterback T, Rizzo M, Lee K, Kosack D, Moestl D, Wedler H, Lauber J, Stjepandic D, Hoheisel J, Straetz M, Heim S, Kiewitz C, Eisen J, Timmis KN, Düsterhöft A, Tümmler B, Fraser CM.** 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ Microbiol **4:**799-808.
195. **Conway T.** 1992. The Entner-Doudoroff pathway: History, physiology and molecular biology. FEMS Microbiol Rev **103:**1-28.
196. **Stephenson MP, Jackson FA, Dawes EA.** 1978. Further observations on carbohydrate metabolism and its regulation in *Azotobacter beijerinckii*. J Gen Microbiol **109:**89-96.
197. **Steinbüchel A.** 1986. Expression of the *Escherichia coli pfkA* gene in *Alcaligenes eutrophus* and in other gram-negative bacteria. J Bacteriol **166:**319-327.
198. **Daldal F.** 1984. Nucleotide sequence of the gene *pfkB* encoding the minor phosphofructokinase of *Escherichia coli* K-12. Gene **28:**337-342.
199. **Lovingshimer MR, Siegele D, Reinhart GD.** 2006. Construction of an inducible, *pfkA* and *pfkB* deficient strain of *Escherichia coli* for the expression and purification of phosphofructokinase from bacterial sources. Protein Expr Purif **46:**475-482.
200. **Romano AH, Eberhard SJ, Dingle SL, McDowell TD.** 1970. Distribution of the phosphoenolpyruvate:glucose phosphotransferase system in bacteria. J Bacteriol **104**.
201. **Adewoye LO, Tschetter L, O'Neil J, Worobec EA.** 1998. Channel specificity and secondary structure of the glucose-inducible porins of *Pseudomonas* spp. J Bioenerg Biomembr **30:**257-267.

202. **van den Berg B.** 2012. Structural basis for outer membrane sugar uptake in pseudomonads. J Biol Chem **287:**41044-41052.
203. **del Castillo T, Ramos JL, Rodríguez-Herva JJ, Fuhrer T, Sauer U, Duque E.** 2007. Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. J Bacteriol **189:**5142-5152.
204. **Adewoye LO, Worobec EA.** 2000. Identification and characterization of the *gltK* gene encoding a membrane-associated glucose transport protein of *Pseudomonas aeruginosa*. Gene **253:**323-330.
205. **Wong TY, Murdock CA, Concannon SP, Lockey TD.** 1991. Simultaneous uptake of galactose and glucose by *Azotobacter vinelandii*. Biochem Cell Biol **69:**711-714.
206. **Essenberg RC, Candler C, Nida SK.** 1997. *Brucella abortus* strain 2308 putative glucose and galactose transporter gene: cloning and characterization. Microbiology **143:**1549-1555.
207. **Brünker P, Altenbuchner J, Mattes R.** 1998. Structure and function of the genes involved in mannitol, arabitol and glucitol utilization from *Pseudomonas fluorescens* DSM50106. Gene **206:**117-126.
208. **Hager PW, Calfee MW, Phibbs PV.** 2000. The *Pseudomonas aeruginosa devB/SOL* homolog, *pgl*, is a member of the *hex* regulon and encodes 6-phosphogluconolactonase. J Bacteriol **182:**3934-3941.
209. **Daddaoua A, Krell T, Ramos JL.** 2009. Regulation of glucose metabolism in *Pseudomonas*: the phosphorylative branch and Entner-Doudoroff enzymes are regulated by a repressor containing a sugar isomerase domain. J Biol Chem **284:**21360-21368.
210. **Lessmann D, Schimz K-L, Kurz G.** 1975. D-Glucose-6-phosphate dehydrogenase (Entner-Doudoroff enzyme) from *Pseudomonas fluorescens*. Eur J Biochem **59:**545-559.
211. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947-2948.
212. **Koonin EV.** 2005. Orthologs, paralogs, and evolutionary genomics. Annu Rev Genet **39:**309-338.
213. **Lerat E, Daubin V, Ochman H, Moran NA.** 2005. Evolutionary origins of genomic repertoires in bacteria. PLoS Biol **3:**e130.
214. **Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crecy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Ruckert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V.** 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. Nucleic Acids Res **33:**5691-5702.
215. **Rocha EP.** 2006. Inference and analysis of the relative stability of bacterial chromosomes. Mol Biol Evol **23:**513-522.

216. **Ballouz S, Francis AR, Lan R, Tanaka MM.** 2010. Conditions for the evolution of gene clusters in bacterial genomes. PLoS Comput Biol **6:**e1000672.
217. **Borgos SE, Bordel S, Sletta H, Ertesvåg H, Jakobsen Ø, Bruheim P, Ellingsen TE, Nielsen J, Valla S.** 2013. Mapping global effects of the anti-sigma factor MucA in *Pseudomonas fluorescens* SBW25 through genome-scale metabolic modeling. BMC Syst Biol **7:**19.
218. **Deutscher J, Francke C, Postma PW.** 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol Mol Biol Rev **70:**939-1031.
219. **Pflüger-Grau K, Chavarría M, de Lorenzo V.** 2011. The interplay of the $EIIA^{Ntr}$ component of the nitrogen-related phosphotransferase system (PTS^{Ntr}) of *Pseudomonas putida* with pyruvate dehydrogenase. Biochim Biophys Acta **1810:**995-1005.
220. **Hassett DJ, Howell ML, Sokol PA, Vasil ML, Dean GE.** 1997. Fumarase C activity is elevated in response to iron deprivation and in mucoid, alginate-producing *Pseudomonas aeruginosa*: cloning and characterization of *fumC* and purification of native FumC. J Bacteriol **179:**1442-1451.
221. **Bagge N, Schuster M, Hentzer M, Ciofu O, Givskov M, Greenberg EP, Høiby N.** 2004. *Pseudomonas aeruginosa* biofilms exposed to imipenem exhibit changes in global gene expression and β-lactamase and alginate production. Antimicrob Agents Chemother **48:**1175-1187.
222. **Tart AH, Blanks MJ, Wozniak DJ.** 2006. The AlgT-dependent transcriptional regulator AmrZ (AlgZ) inhibits flagellum biosynthesis in mucoid, nonmotile *Pseudomonas aeruginosa* cystic fibrosis isolates. J Bacteriol **188:**6483-6489.
223. **León R, Espín G.** 2008. *flhDC*, but not *fleQ*, regulates flagella biogenesis in *Azotobacter vinelandii*, and is under AlgU and CydR negative control. Microbiology **154:**1719-1728.
224. **Enos-Berlage JL, Downs DM.** 1997. Mutations in *sdh* (succinate dehydrogenase genes) alter the thiamine requirement of *Salmonella typhimurium*. J Bacteriol **179:**3989-3996.
225. **Lien SK, Sletta H, Ellingsen TE, Valla S, Correa E, Goodacre R, Vernstad K, Borgos SEFB, Bruheim P.** 2013. Investigating alginate production and carbon utilization in *Pseudomonas fluorescens* SBW25 using mass spectrometry-based metabolic profiling. Metabolomics **9:**403-417.
226. **Jarman TR, Pace GW.** 1984. Energy requirements for microbial exopolysaccharide synthesis. Arch Microbiol **137:**231-235.
227. **García-Ochoa F, Santos VE, Casas JA, Gómez E.** 2000. Xanthan gum: production, recovery, and properties. Biotechnol Adv **18:**549-579.
228. **Schurr MJ, Deretic V.** 1997. Microbial pathogenesis in cystic fibrosis: co-ordinate regulation of heat-shock response and conversion to mucoidy in *Pseudomonas aeruginosa*. Mol Microbiol **24:**411-420.
229. **Keith LM, Bender CL.** 1999. AlgT (σ^{22}) controls alginate production and tolerance to environmental stress in *Pseudomonas syringae*. J Bacteriol **181:**7176-7184.

230. **Firoved AM, Boucher JC, Deretic V.** 2002. Global genomic analysis of AlgU (σ^E)-dependent promoters (sigmulon) in *Pseudomonas aeruginosa* and implications for inflammatory processes in cystic fibrosis. J Bacteriol **184:**1057-1064.
231. **Firoved AM, Deretic V.** 2003. Microarray analysis of global gene expression in mucoid *Pseudomonas aeruginosa*. Journal of Bacteriology **185:**1071-1081.
232. **Jones AK, Fulcher NB, Balzer GJ, Urbanowski ML, Pritchett CL, Schurr MJ, Yahr TL, Wolfgang MC.** 2010. Activation of the *Pseudomonas aeruginosa* AlgU regulon through *mucA* mutation inhibits cyclic AMP/Vfr signaling. J Bacteriol **192:**5709-5717.
233. **Wood LF, Ohman DE.** 2012. Identification of genes in the σ^{22} regulon of *Pseudomonas aeruginosa* required for cell envelope homeostasis in either the planktonic or the sessile mode of growth. MBio **3**.
234. **Bertsova YV, Bogachev AV, Skulachev VP.** 1997. Generation of protonic potential by the *bd*-type quinol oxidase of *Azotobacter vinelandii*. FEBS Lett **414:**369-372.
235. **Wu G, Cruz-Ramos H, Hill S, Green J, Sawers G, Poole RK.** 2000. Regulation of cytochrome *bd* expression in the obligate aerobe *Azotobacter vinelandii* by CydR (Fnr). Sensitivity to oxygen, reactive oxygen species, and nitric oxide. J Biol Chem **275:**4679-4686.
236. **Saier MH, Jr., Ramseier TM.** 1996. The catabolite repressor/activator (Cra) protein of enteric bacteria. J Bacteriol **178:**3411-3417.
237. **Sarkar D, Siddiquee KA, Araúzo-Bravo MJ, Oba T, Shimizu K.** 2008. Effect of *cra* gene knockout together with *edd* and *iclR* genes knockout on the metabolism in *Escherichia coli*. Arch Microbiol **190:**559-571.
238. **Shimada T, Yamamoto K, Ishihama A.** 2011. Novel members of the Cra regulon involved in carbon metabolism in *Escherichia coli*. J Bacteriol **193:**649-659.
239. **Chavarría M, Fuhrer T, Sauer U, Pflüger-Grau K, de Lorenzo V.** 2013. Cra regulates the cross-talk between the two branches of the phosphoenolpyruvate : phosphotransferase system of *Pseudomonas putida*. Environ Microbiol **15:**121-132.
240. **Chavarría M, Kleijn RJ, Sauer U, Pflüger-Grau K, de Lorenzo V.** 2012. Regulatory tasks of the phosphoenolpyruvate-phosphotransferase system of *Pseudomonas putida* in central carbon metabolism. MBio **3**.
241. **Jacobson MR, Premakumar R, Bishop PE.** 1986. Transcriptional regulation of nitrogen fixation by molybdenum in *Azotobacter vinelandii*. J Bacteriol **167:**480-486.
242. **Hamilton TL, Ludwig M, Dixon R, Boyd ES, Dos Santos PC, Setubal JC, Bryant DA, Dean DR, Peters JW.** 2011. Transcriptional profiling of nitrogen fixation in *Azotobacter vinelandii*. J Bacteriol **193:**4477-4486.
243. **Dixon R, Kahn D.** 2004. Genetic regulation of biological nitrogen fixation. Nat Rev Microbiol **2:**621-631.
244. **Ohman DE.** 2009. Alginate gene regulation, p. 117-133. *In* Rehm BHA (ed.), Alginates: Biology and Applications. Springer-Verlag Berlin, Heidelberg.

245. **Muhammadi, Ahmed N.** 2007. Genetics of bacterial alginate: alginate genes distribution, organization and biosynthesis in bacteria. Curr Genomics **8:**191-202.

Paper I

Journal of Bacteriology, July 2009, p. 4534–4545 Vol. 191, No. 14
0021-9193/09/$08.00+0 doi:10.1128/JB.00504-09
Copyright © 2009, American Society for Microbiology. All Rights Reserved.

Genome Sequence of *Azotobacter vinelandii*, an Obligate Aerobe Specialized To Support Diverse Anaerobic Metabolic Processes[▽]†

João C. Setubal,[1,2]* Patricia dos Santos,[3] Barry S. Goldman,[4] Helga Ertesvåg,[5] Guadelupe Espin,[6] Luis M. Rubio,[7] Svein Valla,[5] Nalvo F. Almeida,[1,8] Divya Balasubramanian,[9]‡ Lindsey Cromes,[9]§ Leonardo Curatti,[10] Zijin Du,[4] Eric Godsy,[4] Brad Goodner,[9] Kaitlyn Hellner-Burris,[9] José A. Hernandez,[11] Katherine Houmiel,[12] Juan Imperial,[20] Christina Kennedy,[13] Timothy J. Larson,[14] Phil Latreille,[4] Lauren S. Ligon,[14]¶ Jing Lu,[4] Mali Mærk,[5] Nancy M. Miller,[4] Stacie Norton,[4] Ina P. O'Carroll,[14]‖ Ian Paulsen,[15] Estella C. Raulfs,[14] Rebecca Roemer,[9] James Rosser,[12]# Daniel Segura,[6] Steve Slater,[16] Shawn L. Stricklin,[4] David J. Studholme,[17] Jian Sun,[1]†† Carlos J. Viana,[8] Erik Wallin,[9] Baomin Wang,[13] Cathy Wheeler,[9]‡‡ Huijun Zhu,[4] Dennis R. Dean,[14] Ray Dixon,[18] and Derek Wood[12,19]

Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061[1]; Department of Computer Science, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061[2]; Department of Chemistry, Wake Forest University, Winston-Salem, North Carolina 27109[3]; Monsanto Company, St. Louis, Missouri[4]; Department of Biotechnology, Norwegian University of Science and Technology, N-7491 Trondheim, Norway[5]; Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico[6]; IMDEA Energia, Centro de Biotecnologia y Genomica de Plantas, Madrid 28223, Spain[7]; Departmento de Computacao e Estatistica, Universidade Federal do Mato Grosso do Sul, Campo Grande, Brazil[8]; Department of Biology, Hiram College, Ohio[9]; CONICET-FIBA, Centro de Estudios de Biodiversidad y Biotecnologia, Mar del Plata 7600, Argentina[10]; Department of Biochemistry, Midwestern University, Glendale, Arizona 85308[11]; Seattle Pacific University, Seattle, Washington 98026[12]; Department of Plant Sciences, University of Arizona, Tucson, Arizona 85721[13]; Department of Biochemistry, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061[14]; Macquarie University, Sydney, Australia[15]; Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, Wisconsin[16]; The Sainsbury Laboratory, Colney Lane, Norwich NR4 7UH, United Kingdom[17]; Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, United Kingdom[18]; Department of Microbiology, University of Washington, Seattle, Washington 98119[19]; and Centro de Biotecnologia y Genomica de Plantas (CBGP, UPM-INIA), Madrid 28223, Spain[20]

Received 11 April 2009/Accepted 30 April 2009

Azotobacter vinelandii **is a soil bacterium related to the *Pseudomonas* genus that fixes nitrogen under aerobic conditions while simultaneously protecting nitrogenase from oxygen damage. In response to carbon availability, this organism undergoes a simple differentiation process to form cysts that are resistant to drought and other physical and chemical agents. Here we report the complete genome sequence of *A. vinelandii* DJ, which has a single circular genome of 5,365,318 bp. In order to reconcile an obligate aerobic lifestyle with exquisitely oxygen-sensitive processes, *A. vinelandii* is specialized in terms of its complement of respiratory proteins. It is able to produce alginate, a polymer that further protects the organism from excess exogenous oxygen, and it has multiple duplications of alginate modification genes, which may alter alginate composition in response to oxygen availability. The genome analysis identified the chromosomal locations of the genes coding for the three known oxygen-sensitive nitrogenases, as well as genes coding for other oxygen-sensitive enzymes, such as carbon monoxide dehydrogenase and formate dehydrogenase. These findings offer new prospects for the wider application of *A. vinelandii* as a host for the production and characterization of oxygen-sensitive proteins.**

Azotobacter vinelandii is a free-living nitrogen-fixing bacterium of the gammaproteobacteria. It is found in soils worldwide, with features of nitrogen and energy metabolism relevant to agriculture (41, 42). This organism has been studied for more than 100 years by numerous scientists throughout the world. Prior to Joshua Lederberg's discovery of sexuality in *Escherichia coli* (47), *A. vinelandii* was the experimental organism of choice for many investigators during the emergence of biochemistry as a dominant discipline within the life sciences. Examples include the classical Lineweaver-Burk kinetic parameters, developed using enzymes from *A. vinelandii* (51), and the isolation by Severo Ochoa of polynucleotide phosphor-

* Corresponding author. Mailing address: Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061. Phone: (540) 231-9464. Fax: (540) 231-2606. E-mail: setubal@vbi.vt.edu.

† Supplemental material for this article may be found at http://jb.asm.org/.

‡ Present address: Department of Microbiology, University of Illinois, Urbana, IL.

§ Present address: School of Veterinary Medicine, Ohio State University, Columbus, OH.

¶ Present address: Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC 27599-7290.

‖ Present address: HIV Drug Resistance Program, National Cancer Institute, Frederick, MD.

Present address: School of Molecular Biosciences, Washington State University, Pullman, WA 99164.

†† Present address: La Jolla Institute for Allergy & Immunology, La Jolla, CA 92037.

‡‡ Present address: Department of Biology, John Carroll University, Cleveland, OH.

▽ Published ahead of print on 8 May 2009.

ylase from *A. vinelandii*, which was used in studies that contributed to the elucidation of the genetic code (62).

A. vinelandii is able to adapt its metabolism to diverse sources of nutrients. If no carbon source is present, *A. vinelandii* will undergo a differentiation process to form cysts that are resistant to desiccation and other chemical and physical challenges (74). While the process of encystment has been known for many years and studied at the physiological and morphological levels, there is little knowledge about the unique biosynthetic pathways that are involved and how they are regulated. Previous work has implicated the alternative sigma factors AlgU and RpoS in the differentiation process (13, 57, 64). Alginate polymers with different monomer compositions are an important structural component of the cyst, and at the end of exponential growth, *A. vinelandii* cells accumulate poly-beta-hydroxybutyrate (PHB) as a reserve carbon and energy source (81). The physiology of PHB formation has been well studied in a variety of different systems, and the PHB biosynthetic operon has been described (67, 77). *A. vinelandii* can also produce copolymers of hydroxybutyrate and hydroxyvalerate, known to improve the flexibility and stretch of bioplastics (63).

A. vinelandii has long served as a model for biochemical and genetic studies of biological nitrogen fixation, the conversion of N_2 into NH_3 by a nitrogenase enzyme. The best-studied nitrogenase consists of two oxygen-sensitive metalloproteins that, in the case of the molybdenum nitrogenase, are denominated the Fe protein and the MoFe protein. *A. vinelandii* is unusual in that it is one of the few bacteria that contain three nitrogenases with different subunit and metal cofactor compositions, namely, the molybdenum nitrogenase, the vanadium nitrogenase, and the iron-only nitrogenase. Expression of these nitrogenases is differentially regulated by metal availability from the medium (27).

Here we present the complete genome sequence of *A. vinelandii* DJ and discuss what the genome has revealed about the organism's ability to protect oxygen-sensitive processes. *A. vinelandii* has been cited as having one of the highest respiratory rates of any known bacterium (10). Diazotrophic growth under aerobic conditions is possible because *A. vinelandii* can adjust oxygen consumption rates to help maintain low levels of cytoplasmic oxygen, which is otherwise detrimental not only to nitrogenase but also to other oxygen-sensitive enzymes expressed by *A. vinelandii*. This phenomenon is called respiratory protection. In this work, we identify unique features of the *A. vinelandii* genome that help to explain the coexistence of oxygen-sensitive reactions and strict aerobic metabolism. The genome sequence and annotation allowed identification of the genes involved in respiration, including key players in respiratory protection. In addition, we have identified unexpected gene clusters encoding a carbon monoxide dehydrogenase (CODH), a formate dehydrogenase (FDH), and a second hydrogenase, all of which are also oxygen-sensitive enzymes.

MATERIALS AND METHODS

Strain description. *A. vinelandii* O is a strain that forms gummy, slimy colonies of pale color. The earliest report of this strain dates from 1952 (92). *Azotobacter agilis* (later named *Azotobacter vinelandii* strain O) was part of the bacterial collection at the University of Wisconsin, Madison. In 1959, Bush and Wilson (11) reported the isolation of a nongummy chromogenic isolate of *A. vinelandii* O, which they named *A. vinelandii* OP. This new strain has a fluorescent color and well-defined colony shape. *A. vinelandii* strain OP has also been referred to as UW and CA. *A. vinelandii* strain DJ is a high-frequency transforming variant of *A. vinelandii* UW generated in 1984 through multiple rounds of transformation, using chromosomal DNAs from rifampin (rifampicin)-resistant and -sensitive strains (Dennis Dean, personal communication), and is available from the American Type Culture Collection under number ATCC BAA-1303.

Genome sequencing and assembly. A total of 87,000 reads were generated by the Joint Genome Institute for a draft sequence released prior to this project. These reads came from two DNA libraries (insert sizes of 2 to 4 kb and 4 to 8 kb) obtained using mechanical shearing of DNA and cloning into pUC18, followed by a shotgun sequencing approach. The reads and clones were sent to Monsanto, where finishing occurred. The genome was then assembled and edited using Phred, Phrap, and Consed (23, 24, 32). Finishing was completed by generating an optical map (46) cut with the restriction enzymes BamHI and BsiWI and aligning the assembled sequences to the map. Gaps were closed by sequencing specific products. All rRNA operons were amplified with specific flanking primers, sequenced, and assembled individually. All positions with Phred scores of <40 were resequenced using an independent PCR fragment as a template. The error rate is estimated to be less than 1:10,000 bp.

Annotation and comparative genomics. Genome annotation was done by supervised teams of undergraduate students using a web-based system over a preliminary automated annotation, both developed by J. C. Setubal. Comparison to other genomes was done using MUMmer (45) for alignments and orthoMCL (50) for protein families.

Phylogenetic analysis. A total of 18 genomes (see Table S1 in the supplemental material) were used to build the phylogenetic tree. *Chromohalobacter selexigens*, *Hahella chejuensis*, and *Marinomonas* sp. strain MWYL1 were used as outgroup species. OrthoMCL (50) provided 1,445 protein families containing only one representative member of each ingroup genome. These protein sequences were aligned with Muscle (20), and the noninformative columns of the resulting (concatenated) alignment were removed by Gblocks (14). The tree was built using RAxML (79) with the PROTGAMMAWAGF model. Individual gene phylogenies were obtained by aligning protein sequences from *A. vinelandii* and *Pseudomonas* species, using the same method as that described above, but using MrBayes (35) for tree building.

Prediction of σ^{54}- and NifA-binding sites. Multiple alignments of known σ^{54}- and NifA-binding-site DNA sequences were constructed from the literature (3). Using HMMER 1.8.5 (25), we generated profile hidden Markov models representing the binding sequence motifs from the alignments and used HMMER to search the genome for matches.

CydR binding site prediction. We used the sequence proposed by Wu et al. (93) as a query against the genome in a BLASTN (1) search.

Nucleotide sequence accession number. The *A. vinelandii* DJ sequence and its annotation are available at GenBank under accession number CP001157 and at www.azotobacter.org.

RESULTS

General features and phylogeny. The general features of the *A. vinelandii* genome are shown in Table 1. Based on the sequence and its annotation, we have obtained a phylogeny for *A. vinelandii* (Fig. 1) that indicates that it groups most closely with the pseudomonads (family *Pseudomonadaceae*). Among the pseudomonads, the closest relative is the nitrogen-fixing strain *Pseudomonas stutzeri* A1501 (96). Table 1 shows that *A. vinelandii* and *P. stutzeri* A1501 share 46% and 56% of their respective protein coding gene complements. Figure S1 in the supplemental material presents a whole-chromosome alignment between *A. vinelandii* DJ and *P. stutzeri* A1501 in which large-scale conservation of gene order can readily be observed. The "X" pattern seen in this comparison is likely due to inversions around the origin of replication (22) and suggests that many such inversions have taken place since these two species diverged.

Respiration and respiratory protection genes. *A. vinelandii* is an obligately aerobic bacterium that uses the electron transport chain, with molecular oxygen as the final electron acceptor. This strict aerobic metabolism is supported by our analysis

TABLE 1. General features of *Azotobacter vinelandii* DJ and *Pseudomonas stutzeri* A1501 genomes[a]

Parameter	Value	
	Azotobacter vinelandii DJ	*Pseudomonas stutzeri* A1501
Chromosome features		
Size (bp)	5,365,318	4,567,418
% GC	65.7	63.9
Protein-encoding genes		
Total no.	5,051	4,128
No. (%) with functional assignment	3,561 (70.5)	3,180 (77.0)
No. (%) of conserved hypothetical protein genes	739 (14.6)	0 (0.0)
No. (%) of hypothetical protein genes	751 (14.9)	948 (23.0)
No. of pseudogenes	66	11
No. (%) shared protein-encoding genes	2,332 (44.1)	2,298 (55.7)
RNAs		
No. of rRNA operons	6	4
No. of tRNAs	64	61
No. of other RNAs	18	0

[a] Data for *P. stutzeri* were taken from reference 96. The number of shared protein-encoding genes was computed using orthoMCL (50).

of the genome, which lacks genes for complete systems involved in anaerobic respiration using alternative terminal electron acceptors or fermentation processes. At the interface of catabolism and respiration, NADH and succinate are the major electron carriers that feed the electron transport chain, via complex I and complex II, respectively. The genome encodes four NADH-ubiquinone oxidoreductases, including one ATP-coupled NADH oxidoreductase of the Nuo type (Avin28440 to Avin28560) and three ATP-uncoupled NADH-ubiquinone oxidoreductases. The latter are Ndh (Avin12000) and two membrane-associated enzymes involved in the transport of cations across the membrane, i.e., the Nqr type (Avin14590 to Avin14640) and the Sha type (Avin19530 to Avin19580). While there is no association of Nuo, Nqr, and Sha with oxygen protection, the *ndh* gene product has been shown to be important for aerobic nitrogen fixation (5). It is possible that NADH-driven protection of nitrogenase is dependent on Ndh and CydAB oxidase via a ubiquinol-dependent electron transfer pathway. Despite its apparent involvement in nitrogen fixation (5), a copy of *ndh* is found in several *Pseudomonas* species, some of which are not known to fix nitrogen. In addition, *A. vinelandii* carries two copies of the Rnf system (Rnf1 [Avin50930 to Avin50980] and Rnf2 [Avin19220 to Avin 19270]), which shows sequence similarity to a sodium-dependent NADH-ubiquinone oxidoreductase (18). Although there is no change in respiration upon inactivation of either one or both Rnf systems, Rnf1 is required for accumulation of nitrogenase Fe protein (18). An ortholog of Rnf1 exists in the diazotroph *P. stutzeri* A1501 (YP_001171857 to YP_00117852) but not in any other members of the *Pseudomonas* genus.

Ubiquinol-oxygen oxido-reduction can occur either through a two-step pathway, via cytochrome *c* reductase (complex III; Avin13060 to Avin13080) followed by a cytochrome terminal oxidase (complex IV), or in a single-step process, via a ubiquinol-dependent cytochrome terminal oxidase. The genome annotation identified the catalytic and biosynthetic genes of the following five terminal oxidases: cytochrome *c* oxidase (Cdt oxidase; Avin00950 to Avin01020), cytochrome *o* oxidase (Cox; Avin11170 to Avin11180), cytochrome *bd* copy I (CydAB I; Avin19880 to Avin19890), cytochrome *bd* copy II (CydAB II; Avin11040 to Avin11050), and cytochrome cbb_3 (Cco; Avin19940 to Avin20010). While the Cox, CydAB II, and Cco oxidases are encoded in all *Pseudomonas* species analyzed to date, the Cdt oxidase gene is found only in *P. stutzeri* A1501 and *Pseudomonas mendocina* ymp.

The presence of two CydAB oxidases in *A. vinelandii* was unexpected. Phylogenetic analysis revealed that CydAB oxidase I (encoded by *cydAB* gene copy I) has its closest orthologs in *P. stutzeri* A1501 but groups with similar genes from members of the *Acinetobacter* and *Shewanella* genera rather than with those from other members of the *Pseudomonas* genus. Copy II, on the other hand, groups with similar genes from several members of the *Pseudomonas* genus (see Fig. S2 to S5 in the supplemental material). CydAB oxidase I has been characterized extensively (21, 39, 40, 94). The precise function of CydAB oxidase II remains to be elucidated.

Oxygen consumption by the aforementioned terminal oxidases is not the only factor responsible for oxygen protection of nitrogenase. The FeSII protein, known as the Shethna protein and encoded by Avin01520, forms a protective complex with nitrogenase when the enzyme is exposed to oxygen (58). In addition, it has been determined that cellular levels of ATP also contribute to protection of nitrogenase against oxygen damage (53). High concentrations of ATP directly correlate with high electron flux to nitrogenase, which influences the dissociation rate constant of the nitrogenase components (84) and, consequently, the susceptibility of the Fe protein to oxygen damage (83). The *A. vinelandii* genome encodes two sets of ATP synthase machineries. The first complex (Avin52150 to Avin52230) is the ortholog of the *Pseudomonas* ATP synthase and is located close to the putative origin of replication. The second copy (Avin19670 to Avin19750) is located 8 genes upstream of the *sha* operon (complex I-like) and 10 genes downstream of *cydAB* I (complex IV). This copy does not seem to be related phylogenetically to ATP synthases from the *Pseudomonas* genus, and the only orthologs are found in *Burkholderia* species, most of which are obligate aerobes, although very few can fix nitrogen.

Genomic analyses of *A. vinelandii* and microaerobic diazotrophs such as *P. stutzeri* show that they have similar respiratory complexes, suggesting that the regulation of these machineries, especially at the transcriptional level, is very important for adjusting rates of oxygen consumption in order to protect oxygen-sensitive processes. The transcriptional regulator CydR (Avin19910) seems to regulate various physiological processes associated with respiratory protection. Wu et al. (94) showed that CydR can coordinate an oxygen-labile [Fe-S] cluster, which provides a mechanism for CydR to sense subtle changes in the intercellular oxygen concentration and to regulate the expression of respiratory genes. During nitrogen fixation, CydR repression is presumably relieved, resulting in increased expression of uncoupled NADH dehydrogenase (Ndh) and CydAB I. A series of reports have suggested a role

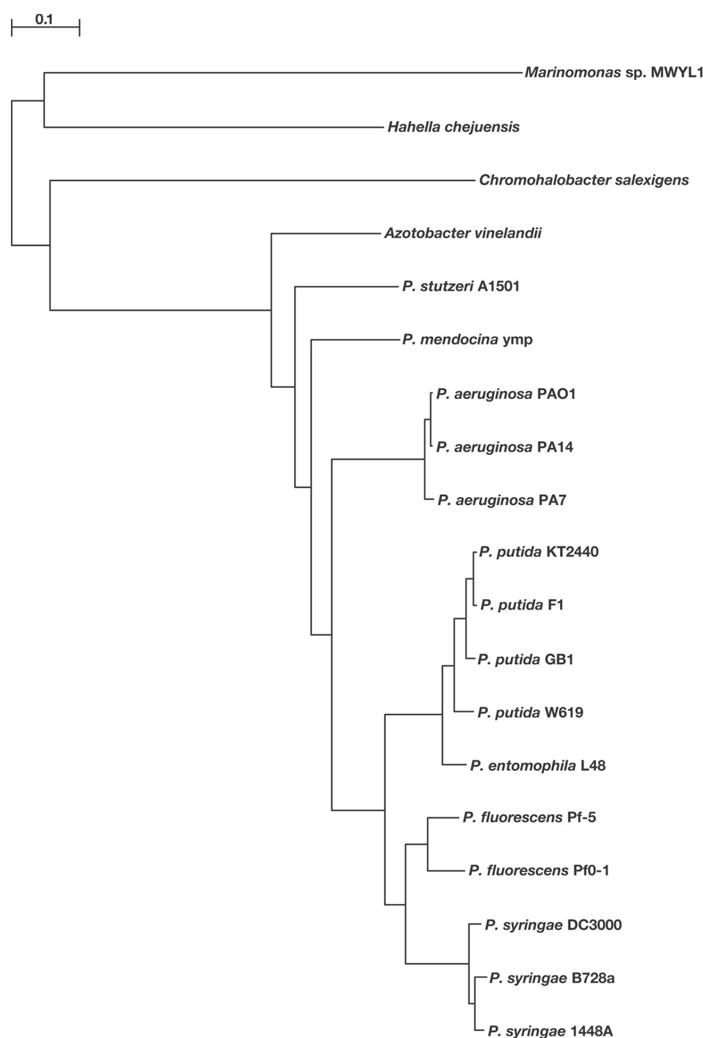


FIG. 1. Phylogenetic tree for *A. vinelandii*, all fully sequenced members of the *Pseudomonas* genus, and three outgroups. Accession numbers for the genomes used are given in Table S1 in the supplemental material. Bootstrap support for all branches was 100%.

for CydR in several other metabolic processes, including synthesis of PHB (95) and flagellar motility (48). Although the role of CydR in respiration is not completely defined, its direct or indirect participation in other cellular processes is even more obscure.

The genomic location of *cydR* (Avin19910) downstream of the *cco* region (containing the *cco* genes [Avin19920 to Avin20010]) suggests that it is involved in the regulation of this system. In support of this, a putative CydR binding site was identified upstream of the *cco* genes, and expression of a *ccoN*::*lacZ* fusion is CydR dependent (data not shown). Similar regulation of Cco by FNR (a CydR ortholog) is seen in other organisms (16). CydR also binds upstream of the CydAB oxidase I promoter and is known to repress transcription of this operon (93). While *cydR* in *A. vinelandii* is adjacent to the *cydAB* operon (Avin19890 to Avin19880) encoding CydAB oxidase I, this organization is not seen in other organisms. Our promoter analysis has further identified a CydR binding site immediately upstream of Avin11170, suggesting that CydR also regulates the *cox* genes in *A. vinelandii.*

A putative CydR binding site is also located upstream of the iron-sulfur cluster biosynthetic genes (*isc* genes; Avin40380 to Avin40410). Cross regulation of Fe-S cluster biosynthesis and respiration makes biochemical sense for at least the following two reasons: (i) a large number of Fe-S proteins are involved in respiration, and (ii) elevated oxygen concentrations result in

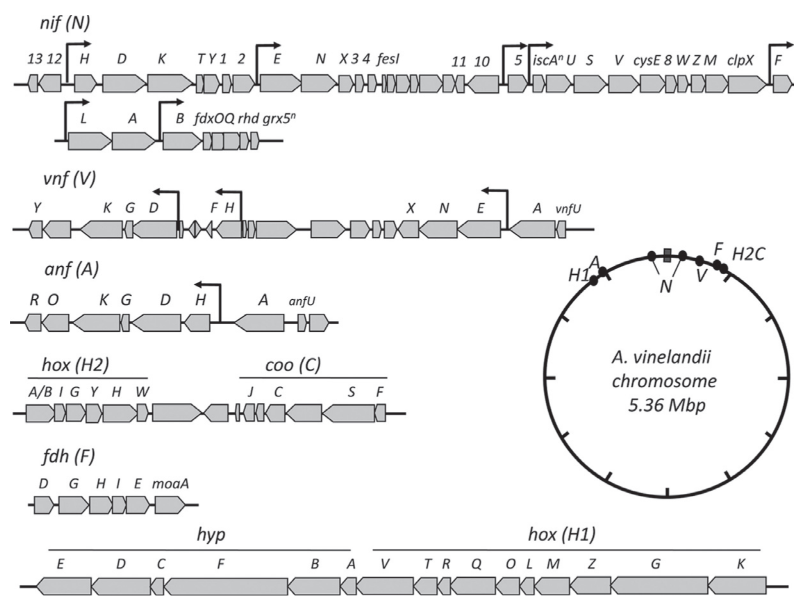


FIG. 2. Oxygen-sensitive protein-encoding genes in *A. vinelandii* and their genomic locations.

increased damage to Fe-S proteins under conditions in which the respiratory proteins are more highly expressed. Therefore, under conditions of high oxygen concentrations, CydR regulation would increase the expression of respiratory components and increase the capacity of the Fe-S cluster biosynthetic apparatus to supply the temporary higher demand for Fe-S clusters.

Nitrogen fixation. *A. vinelandii* expresses three oxygen-sensitive nitrogenase enzymes with different structural subunits and metal cofactor dependencies. Genes encoding the well-studied molybdenum-dependent nitrogenase and its assembly machinery and regulation (*nif*) are located in two regions of the chromosome, adjacent to and equidistant from the origin of replication. Proximity to the origin might result in a higher gene dosage during active growth, which can contribute to the high expression levels of the Nif components. The major *nif* region (comprising genes Avin01360 to Avin01710, oriented away from the putative origin, on the plus strand) encodes the structural subunits and the majority of the assembly machinery. The minor *nif* region (genes Avin50990 to Avin51060, also on the plus strand) contains the regulatory genes *nifL* and *nifA* and genes required for Mo trafficking and nitrogenase cofactor biosynthesis. Adjacent to the minor *nif* region are the *rnf1* region (upstream) and genes for a putative rhodanese (rhd_{nif}; Avin51050) and a monothiol glutaredoxin (grx_{nif}; Avin51060) (downstream). The *rnf1* gene products are associated with the accumulation of active nitrogenase component 2, also referred to as Fe protein (18). Likewise, the inactivation of the *nif*-associated glutaredoxin resulted in a 50% loss of Fe protein activity (data not shown). The dramatic effect on Fe protein activity and the less pronounced effect on nitrogenase component 1, the MoFe protein, could be attributed to an inherent susceptibility to oxygen damage and/or deficiency in repair of the [Fe-S] cofactor in strains lacking these genes.

The genome sequence reveals that the region containing the structural genes for the vanadium nitrogenase (Avin02650 to Avin02660) is in close proximity to the gene cluster carrying genes (Avin02740 to Avin02780) involved in its assembly and regulation (Fig. 2). However, these two *vnf* clusters are separated by three genes predicted to participate in molybdopterin biosynthesis (Avin02700 to Avin02720) and a gene, *pcaK*, encoding a hydroxybenzoate transporter (Avin02690). In contrast, the *anf* genes, encoding the iron-only nitrogenase, are located in a single gene region (Avin48950 to Avin49000), which is regulated by the AnfA transcriptional activator (Avin49020) (Fig. 2).

Although no additional nitrogenase genes were identified in the genome sequence, genes with sequence similarity to the *nif*, *vnf*, and *anf* genes were found scattered throughout the genome. Whether or not these other genes are involved in nitrogen fixation remains to be determined. Interestingly, a gene with sequence similarity to *nifD* was identified (Avin39870), but its inactivation did not affect growth under standard diazotrophic or nondiazotrophic conditions (data not shown). Phylogenetic analysis of the NifD-like sequence in comparison with its paralogs in *A. vinelandii* and *P. stutzeri* reinforced the idea that the three structural and accessory nitrogenase genes were derived from a common ancestor and that the *nifD*-like sequence resembles the ancestral gene (see Fig. S6 in the supplemental material).

A paralog of *nifA*, Avin26490, was identified as encoding a protein that contains an almost identical DNA recognition helix in the C-terminal DNA-binding domain, implying that this protein binds similar enhancer sequences to those recog-

nized by NifA (Avin51000). A search for σ^{54} promoters and NifA upstream activator sequences detected the experimentally determined NifA-dependent promoters upstream of the *nif* and *rnf* genes and additional putative NifA-dependent σ^{54} promoters upstream of genes encoding ModE (Avin50680), a putative molybdate binding protein, ModA3 (Avin50730), and a putative McbC-like oxidoreductase (Avin48680). In view of the similarity in DNA recognition motifs, it is possible that these promoters are activated by either NifA or its paralog, Avin26490.

The genome sequence revealed two genes (Avin33440 and Avin47100) with 78% and 71% identity to *vnfA* (Avin02780), respectively. These VnfA paralogs have two of the three cysteines present in the proposed metal-binding cluster in the amino-terminal GAF domain of VnfA, similar to that of AnfA, which also has two cysteine residues implicated in metal or redox sensing (37, 69).

The *nfuA* gene (Avin28760) encodes a protein with a NifU-like C-terminal domain that shares sequence similarity with the products of *vnfU* (Avin02790) and *anfU* (Avin49030), which are located adjacent to *vnfA* (Avin02780) and *anfA* (Avin49020), respectively. Biochemical experiments suggest that NfuA represents an intermediate [Fe-S] cluster carrier involved in [Fe-S] protein maturation (2). It is likely that VnfU and AnfU play a role in the maturation of clusters present in vanadium- and iron-only nitrogenases.

In addition to paralogs of nitrogen fixation genes, other systems critical to nitrogen fixation appear to have been duplicated. Positioned 35 kb upstream of the minor *nif* gene cluster is the high-affinity molybdate transport system, encoded by *modG* and *modEA1B1C1* (Avin50650 to Avin50690), which supports the expression of Mo-dependent nitrogenase under molybdenum-limiting conditions (65). The second known molybdate transport locus, *modA2B2C2* (Avin01280 to Avin01300), is located 10 kb upstream of the major *nif* gene cluster. Interestingly, we identified a third putative molybdate transport system (Avin50700 to Avin50730), located directly next to the *modEA1B1C1* operon.

Other oxygen-sensitive processes. (i) CODH. The genome of *A. vinelandii* DJ contains genes with sequence similarity to membrane-bound Ni-dependent anaerobic CODH genes. CODH is an α_2 homodimer of the *cooS* product (Avin04490) that carries out the reversible oxidation of CO to CO_2 (71). Electrons extracted from CO are transferred to CooF (Avin04500), a hydrophobic FeS protein similar to the electron transfer subunits of oxidoreductases, such as the FDH beta subunit FdhH (Avin03820).

CODH contains a $NiFe_4S_4$ center, known as the C cluster, at its active site. The assembly of the CODH active site, which has been well studied in *Rhodospirillum rubrum*, requires the activities of three proteins, encoded by the *cooJ*, *cooC*, and *cooT* genes (43), involved in Ni storage and insertion (36, 91). Whereas one gene similar to *cooC* is present in the *A. vinelandii* genome (Avin04470), the genome does not code for proteins similar to CooT or CooJ. Avin04460, located downstream of *cooC*, encodes a 73-amino-acid histidine-rich protein with sequence similarity to the N-terminal portion (first 20%) of HypB proteins in *R. rubrum* (E value, 10^{-7}) and *Ralstonia eutropha* (10^{-5}), which are involved in Ni processing for hydrogenase maturation. Thus, it is likely that the product of Avin04460 substitutes for the role of CooJ in the maturation of CODH. There is a gene similar to *hupE/ureJ* genes (Avin04450) downstream of Avin04460. HupE/UreJ proteins are secondary Ni transporters, and their coding genes are widespread among bacteria, normally clustered with urease or hydrogenase genes. There is another gene in the *hupE/ureJ* family (Avin50400), located near the *hypEDCFB* gene cluster, that probably encodes a HupE protein. We hypothesize that Avin04450 and Avin50400 may encode two different Ni transport systems, specific for CODH and hydrogenase maturation, respectively.

The *A. vinelandii coo* gene cluster includes a gene for a flavin adenine dinucleotide-dependent pyridine nucleotide-disulfide oxidoreductase (Avin04480) inserted between the *cooS* and *cooC* genes. A similar protein is also encoded downstream of *cooFS* in *Carboxydothermus hydrogenoformans.* Proteins of this family utilize flavin adenine dinucleotide to shuttle electrons from NADH to a redox-active disulfide bridge. The product of Avin04480 might be involved in CODH maturation. It has been proposed that Ni insertion into the C cluster involves reversible Ni binding to an Fe_3S_4 center followed by coordination to a specific cysteinyl residue (Cys531 in *R. rubrum* CooS) (36). It is also noteworthy that in vitro Ni insertion into CODH requires dithionite as a reductant. Thus, the Avin04480 gene product could be involved in reduction of a cysteine residue in *A. vinelandii* apo-CODH to facilitate Ni insertion.

The genome sequence reveals the presence of a *cooA* gene (Avin47010), as previously reported (98). The *cooA* gene encodes a transcriptional activator distantly related to the CRP family. CooA is a heme-containing homodimeric protein that functions as a CO sensor and controls the expression of the *coo* genes (72). In *R. rubrum*, CooA activates transcription of two contiguous gene clusters, namely, one that encodes CODH and accessory proteins (*cooFSCTJ*) and another encoding a membrane-bound CO-tolerant NiFe hydrogenase and its accessory proteins (*cooMKLXUH*). Although *A. vinelandii* lacks the genes encoding the CO-tolerant hydrogenase, it contains a gene cluster encoding a chimerical soluble NiFe hydrogenase in the corresponding locus (see below).

(ii) FDH. FDHs combine heterogeneous groups of enzymes found in both prokaryotes and eukaryotes that catalyze the oxidation of formate to CO_2 and H^+. In aerobic organisms, the FDHs are mostly NAD^+-dependent FDHs. Many prokaryotes, however, thrive in anoxic environments, where FDHs are NAD^+-independent enzymes containing a variety of redox centers with oxygen-sensitive active sites composed of molybdenum or tungsten cofactors. These organisms utilize formate (produced from pyruvate) as a main electron donor for a variety of inducible anaerobic respiratory pathways (49).

The *A. vinelandii* genome contains a gene cluster that encodes a NAD^+-independent FDH (*fdhDGHIE*; genes Avin03800 to Avin03840). While best BLAST (1) hits of *fdhD* are against *Pseudomonas* spp., including *P. stutzeri* A1501 (E value, 10^{-84}), the catalytic subunit of FDH encoded by *fdhGHI* shows significant similarity to the alpha, beta, and gamma subunits of the well-studied nitrate-inducible Fdh-N enzyme from enterobacteria, including *E. coli* (85). Similarity E values against *E. coli* O157 were as follows: for *fdhG*, 0.0; for *fdhH*, 10^{-129}; and for *fdhI*, 10^{-48}. In the catalytic site of the alpha subunit (FdhG), the molybdopterin-guanine dinucleotide co-

factor extracts electrons from formate. These electrons are transferred first to the [FeS]-containing beta subunit (FdhH) and finally to the heme-containing gamma subunit (FdhI). Although in *E. coli* Fdh-N function is associated with the activity of a respiratory nitrate reductase complex (NarGHI) located in the inner membrane, *narGHI* genes were not found in the genome of *A. vinelandii*. Interestingly, FdnG from *Klebsiella pneumoniae* was suggested to participate in relieving NifL inhibition of NifA (33), since *K. pneumoniae* strains carrying null mutations of *fdnG* (or the NADH-ubiquinone oxidoreductase genes *nuoCD*) showed significantly reduced *nif* induction under nitrogen-fixing conditions. The *moaA* gene (Avin03850), which encodes a SAM-dependent radical enzyme involved in the first step of the biosynthesis of the molybdenum cofactor (34), is located directly downstream of the *fdhDGHIE* gene cluster. Avin30330 encodes another MoaA homolog.

(iii) Hydrogenases. Prokaryotes, mostly from anaerobic ecosystems, have the ability to use H_2 by employing uptake hydrogenases or to produce H_2 by the activity of H_2-evolving hydrogenases. About one thousand hydrogenase sequences have been identified, many by genome sequencing, and more than 100 have been characterized genetically and/or biochemically (89). Three phylogenetically distinct classes of hydrogenases have been described, namely, [NiFe] hydrogenases, [FeFe] hydrogenases, and [Fe] hydrogenases ([Fe-S] cluster-free hydrogenases) (89, 90). Hydrogenases are usually sensitive to oxygen. For example, while [FeFe] hydrogenases are irreversibly destroyed by oxygen, the catalytic function of [NiFe] hydrogenases is reversibly inactivated by oxygen, but their structural integrity remains stable (8). The *A. vinelandii* genome contains genes encoding two [NiFe] hydrogenases, in membrane-bound and water-soluble forms (see below). Genes predicted to encode [FeFe] or [Fe] hydrogenases were not found.

The membrane-bound [NiFe] hydrogenase of *A. vinelandii* has been characterized extensively at both the biochemical (44, 75) and genetic (15, 31, 56) levels. The structural genes for this enzyme (*hoxKG*) (56) are clustered together with genes coding for hydrogenase accessory proteins, *hoxKGZMLOQRTV* (Avin50500 to Avin50590) and *hypABFCDE* (Avin50440 to Avin50480), encoding the components of the hydrogenase electron transport chain and proteins for the biosynthesis of the hydrogenase cofactors and hydrogenase maturation (15, 31, 56). Mutational analysis has shown the requirement of these genes for H_2 oxidation activity (15, 31, 56).

It has generally been assumed that the presence of an uptake hydrogenase in nitrogen-fixing bacteria provides the advantage of recycling H_2 produced by nitrogenase, thereby yielding extra reductant and/or ATP for N_2 reduction and also providing respiratory protection for nitrogenase. Although competitive fitness experiments under carbon-limited conditions supported that notion for the related strain *Azotobacter chroococcum* (97), no significant difference was observed in ATP accumulation or in H_2-dependent O_2 uptake in a *hoxKG*-deficient mutant of *A. vinelandii* (52). Thus, the physiological role of the uptake hydrogenase in N_2 fixation is not completely understood.

Close to *hypE* (Avin50440) there is a gene, Avin50400, which codes for a protein that belongs to the HupE/UreJ family of proteins involved in Ni transport or processing. Two adjacent open reading frames (Avin50410 and Avin50420), encoding putative membrane proteins, are unique to the *A. vinelandii* genome. Most pseudomonads possess a single copy of this membrane protein-encoding gene, but 12 copies are predicted for the *A. vinelandii* genome (Avin00560, Avin00570, Avin01720, Avin04580, Avin28060, Avin31780, Avin32050, Avin33410, Avin35450, Avin38580, Avin50420, and Avin50410). Whether these genes encode previously unknown accessory proteins for the *A. vinelandii* uptake hydrogenase and/or display an *A. vinelandii*-specific trait needs to be addressed.

A soluble NAD^+-reducing [NiFe] hydrogenase is predicted to be encoded by Avin04360 to Avin04410. This putative multisubunit enzyme shares characteristics with both the soluble sensing hydrogenase HoxFUYHWI from *Ralstonia eutropha* (9) and the soluble H_2-evolving, S^0-reducing sulfhydrogenase ShyBCDA from *Pyrococcus furiosus* (54). There is weak but still significant similarity (E value, 10^{-4}) between Avin04410 and *R. eutropha hoxW* (CAA63575), encoding the hydrogenase peptidase, and between Avin04370 and *hoxI* (NP_942732), encoding the NADPH-binding subunit (E value, 10^{-6}) (9). These two *A. vinelandii* genes appear to be interspersed among the *P. furiosus shyBCDA*-like genes (Avin04360 and Avin04370 to Avin04390) for the hydrogenase structural genes and the accessory subunits encoded by *shyBC*. Thus, *A. vinelandii* would encode a putative soluble [NiFe] hydrogenase of chimerical nature among other previously characterized hydrogenases. A similar arrangement of genes to those of *A. vinelandii* is present in the genomes of other bacteria, such as the aerobic, nitrogen-fixing alphaproteobacterium *Beijerinckia indica*, the aerobic ammonia-oxidizing gammaproteobacterium *Nitrosococcus mobilis* Nb-231, and others, which have been uncovered by genome sequencing projects and represent the most similar relatives to the putative *A. vinelandii* soluble hydrogenase. The corresponding amino acid sequence identities to *A. vinelandii shyB-hoxI-shyCDA-hoxW*-like genes of these putative hydrogenase subunits are 66%, 75%, 63%, 73%, 76%, and 52% for *B. indica* and 58%, 60%, 65%, 71%, 65%, and 45% for *N. mobilis*, respectively. Interestingly, this gene cluster appears to have undergone extensive rearrangements during the evolution of hydrogenases. While this cluster is adjacent to the putative CODH gene cluster in *A. vinelandii* (Avin04450 to Avin04490), in *N. mobilis* it is adjacent to the *hypBAEDCF*-like genes for the maturation of the [NiFe] hydrogenase, and in *B. indica* they do not appear to be linked physically to hydrogen metabolism genes. Mutagenesis studies of these genes would shed light on hydrogen metabolism in *A. vinelandii* and other bacteria.

Alginate. There are two possible strategies for keeping the cytoplasm anaerobic: either the bacterium can remove the oxygen after it has entered the cell, or it can create a barrier to impede O_2 transfer into the cell. Sabra et al. showed that the alginate capsule of *A. vinelandii* was affected by the oxygen tension (73). Alginate is a linear copolymer of 1→4-linked β-D-mannuronic acid and α-L-guluronic acid, where some of the mannuronic acid residues may be acetylated. The alginate biosynthetic gene set consists of 12 genes (Avin10860 to Avin10970). The physical organization of this cluster is highly conserved in *A. vinelandii* and *Pseudomonas* species able to produce alginate, including *Pseudomonas aeruginosa*, *Pseudomonas fluorescens*, *Pseudomonas syringae*, *Pseudomonas putida*, *Pseudomonas entomophila*, and *P. mendocina* (see Fig. S7 in

the supplemental material). In addition, *A. vinelandii* encodes a set of seven secreted mannuronan C-5 epimerases (AlgE1-7) that modify the polymer outside the cells (*algE1*, Avin51190; *algE2*, Avin51180; *algE3*, Avin51170; *algE4*, Avin51200; *algE5*, Avin33710; *algE6*, Avin51230; and *algE7*, Avin51250). Some of these epimerases are responsible for introducing consecutive guluronic acid residues into the polymer, and this structural feature is necessary for forming a gel with divalent cations. Close homologs of the epimerase genes *algE1-7* are not described for other species except for *P. syringae*, which encodes a single bifunctional alginate epimerase/deacetylase, PsmE (7).

Wild-type strains of *Pseudomonas* usually do not produce alginate under laboratory conditions, while most strains of *A. vinelandii* produce alginate constitutively and use the polymer both for making the vegetative state capsule and for the cyst coat (see below). This probably explains why the regulation of alginate synthesis in *A. vinelandii* differs from that in *Pseudomonas*. Alginate production in *P. aeruginosa* is controlled by a complex regulatory system that includes AlgR, AlgB, AlgQ, AmrZ, and AlgP (30, 59). Orthologs of the genes encoding these regulators are present in the genome of *A. vinelandii* DJ, as follows: *algR* (Avin47610), *algB* (Avin11120), *algP* (Avin47540), *algQ* (Avin47550), and *amrZ* (Avin34410). While in *Pseudomonas* AlgR is required for alginate production, *algR* mutants of *A. vinelandii* produce alginate but are unable to form mature cysts (60). The participation of AlgB, AlgP, AlgQ, and AmrZ in the regulation of alginate synthesis has not been studied in *A. vinelandii*. Orthologs of these genes are present, however, in *P. stutzeri* A1501, which does not produce alginate.

Other regulators of alginate synthesis, including the *algU-mucABCD* operon (Avin13660 to Avin13730), the protease genes *algW* (Avin12950), *mucP* (Avin38920), and *prc* (Avin35170), and the diguanylate cyclase gene *mucR* (Avin49140), are also conserved in *A. vinelandii* and *Pseudomonas* species. *A. vinelandii* does not carry any close homolog to *mucE*, which in *P. aeruginosa* is necessary for the proteolytic events that degrade the anti-sigma factor MucA enabling alginate production (70). Strain DJ does not produce alginate due to the presence of an insertion sequence that splits *algU* (Avin13660) in two (55).

Polymer production and encystment. A distinctive characteristic of *A. vinelandii* is the formation of a desiccation-resistant cell, described as a cyst, upon encountering adverse growth conditions or upon induction of vegetative cells with specific reagents. Only a few bacterial species have been observed to make cysts. Initial studies of cyst formation in *Rhodospirillum centenum* and *Azospirillum brasiliense* identified a few genes (4, 88) that are involved in the regulation of encystment in these bacteria. Genes similar to these regulators, however, have not been found in the *A. vinelandii* genome. Alginate, in addition to its role in respiratory protection, is a major component of the cyst capsule, and alginate with consecutive guluronic residues is essential for the formation of mature cysts (12, 80). During encystment, *A. vinelandii* was shown to incorporate phenolic lipids (alkylresorcinols and alkylpyrones) into the membrane. The proteins involved in the biosynthesis of these lipids have been characterized in *A. vinelandii* (29), and it has been shown that these lipids play a structural role in the cyst capsule, although they are not essential for desiccation resistance (78).

Even though *A. vinelandii* seems unable to utilize alginate as an energy source, it carries at least six enzymes with alginate lyase activities, including AlyA1-3 (Avin31810, Avin23960, and Avin13810), AlgE7 (Avin51250), Avin46500 (82; our unpublished data), and AlgL (Avin10900), the alginate lyase encoded in the alginate biosynthetic gene cluster that is also present in *Pseudomonas* species. One of these, AlyA3, is required for cyst germination, probably helping the germinating cells to escape from the cyst coat by depolymerizing the alginate (H. Ertesvåg, personal communication). In contrast, AlgL is not required for cyst germination (86).

Abundant PHB granules accumulate during encystment and are a major component of the central body of the cyst. Conditions or strains favoring greater polymer accumulation also produce more mature cysts (81), although PHB synthesis is not essential for encystment (76). The PHB biosynthetic genes *phbB* (Avin23650), *phbA* (Avin23640), and *phaC* (Avin23630), the regulatory genes *phbR* (Avin23660) and *phbF* (Avin23680), and the phasin gene *phbP* (Avin23670) were previously identified (66, 76, 77). The biosynthetic genes seem to be under CydR control, since the PhbB and PhbA proteins are increased in a *cydR* mutant (94) and a putative CydR binding site is present upstream of *phbB* (66). Genome analysis revealed additional genes probably involved in PHB metabolism, including genes similar to *phbB* and *phbA* and genes encoding putative phasins, PHB synthases, PHB depolymerases, and PHB oligomer hydrolases (see Table S6 in the supplemental material).

PHB is the polyhydroxyalkanoate (PHA) usually produced by *A. vinelandii*. However, the addition of valerate, heptanoate, or nonanoate to a culture of *A. vinelandii* UWD grown in glucose allows the synthesis of a copolymer, poly-(hydroxybutyrate-cohydroxyvalerate) (PHB-co-HV) (63). In the pseudomonads belonging to rRNA homology group I, an (*R*)-specific enoyl-coenzyme A hydratase, the product of the *phaJ* gene, is responsible for the channeling of β-oxidation intermediates to PHA synthesis (28). In accordance with the PHB-co-HV synthetic capacity of *A. vinelandii*, a *phaJ* gene (Avin30160) is present in its genome.

DISCUSSION

While *P. stutzeri* is the closest relative to *A. vinelandii* among fully sequenced prokaryotes, there are marked physiological and metabolic differences between these organisms. Unlike *A. vinelandii*, *P. stutzeri* can fix nitrogen only under microaerobic conditions, but in contrast to *A. vinelandii*, this organism can grow under anaerobic conditions by utilizing nitrate as a terminal electron acceptor. The classical high rates of respiration exhibited by *A. vinelandii* are apparently bestowed by the provision of five terminal oxidases (Fig. 3), supported by a number of NADH- and other substrate-driven respiratory chains. Unexpectedly, the genome encodes two cytochrome *bd*-type terminal oxidases. One of these, CydAB I, is the well-studied low-affinity *bd* oxidase known to be involved in respiratory protection of nitrogenase (68). This terminal oxidase appears to be essential for aerotolerant nitrogen fixation in *A. vinelandii*. The second *bd*-type terminal oxidase, CydAB II, resembles the cyanide-tolerant Cio terminal oxidases, which differ in their heme complement from the canonical cytochrome *bd* (17).

The distinctive ability of *A. vinelandii* to carry out nitrogen fixation under aerobic conditions is subject to regulation by the

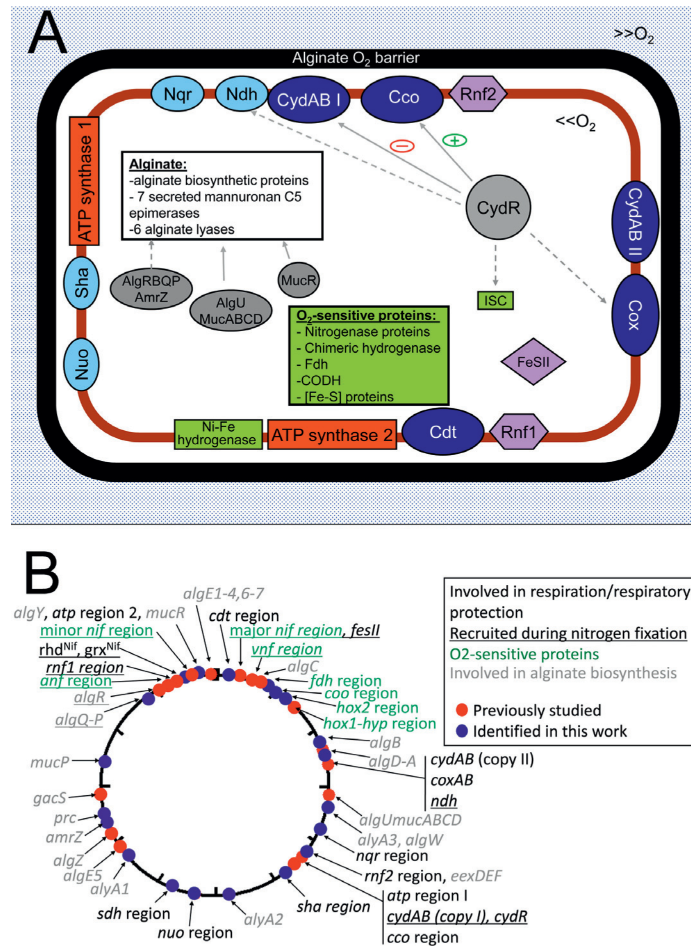


FIG. 3. Respiratory protection in *Azotobacter vinelandii*. (A) Illustration of *A. vinelandii* cell and the key proteins involved in respiratory protection. A number of alginate biosynthetic proteins are produced that direct the formation of an alginate coating around the cell, acting as a physical O_2 barrier (black lining). Respiratory proteins, including five terminal oxidases (dark blue), four NADH-ubiquinone oxidoreductases (light blue), two ATP synthases (orange), and other proteins involved in respiratory protection (lavender), provide the second line of defense against O_2. Proteins sensitive to O_2 exposure are shown in green boxes. Solid line arrows represent regulatory events for which experimental evidence exists. Dashed line arrows represent putative regulation on the basis of promoter binding regions. Regulatory proteins are depicted in gray. (B) Locations of respiratory and respiratory protection genes in the *A. vinelandii* genome. Proteins are color coded according to their role in respiration, respiratory protection, and nitrogen fixation. Genes whose protein products have been identified previously are represented by red dots. The genes identified in this work (represented by blue dots) encode the following proteins: cytochrome *o* oxidase (Avin11160 to Avin11180), Rnf2 complex (Avin19220 to Avin19270), Sdh complex (Avin29780 to Avin29810), Sha complex (Avin19530 to Avin19580), Nuo complex (Avin28440 to Avin28560), Nqr complex (Avin14590 to Avin14640), Rhd^{Nif} (Avin51050), Grx^{Nif} (Avin51060), cytochrome *c* oxidase (Avin00940 to Avin01020), chimeric soluble Ni-Fe hydrogenase (Avin04360 to Avin04410), Fdh (Avin03820), CODH (Avin04450 to Avin04490), ATP synthase 2 complex (Avin52150 to Avin52230), AlgB (Avin11120), AlgQ (Avin47550), AmrZ (Avin34410), AlgP (Avin47540), AlgW (Avin12950), MucP (Avin38920), Prc (Avin35170), AlyA1-3 (Avin13810, Avin23960, and Avin31810), AlgY (Avin49400), and the alginate lyase (Avin 46500).

oxygen-responsive transcriptional regulator CydR, which like Fnr contains a $[4Fe\text{-}4S]^{2+}$ cluster and negatively regulates expression of the *cydAB* genes that encode the low-affinity CydAB I enzyme (93). CydR also appears to control expression of the uncoupled NADH-ubiquinol dehydrogenase (Ndh) (6), which is thought to supply electrons to CydAB I oxidase and is also essential for aerotolerant nitrogen fixation (5). Our studies suggest that CydR activates expression of Cco, which may explain why CydR is required for growth under microaerobic conditions (94). Overall, CydR may have a role analogous to that of Anr in controlling the expression of terminal oxidases in *Pseudomonas* (87), but its involvement in the regulation of PHB metabolism (95) implies a wider role in the integration of carbon source and oxygen availability (Fig. 3).

Mechanisms for respiratory protection are complemented by the barrier to oxygen diffusion into the cell provided by the alginate capsule, which appears to play a major role in protecting nitrogenase from oxygen damage, particularly under phosphate-limiting conditions (73). It is noteworthy that alginate biosynthesis is constitutive in *A. vinelandii* and that a large number of secreted mannuronan C-5 epimerases are encoded in the genome by *algE1-7*. These enzymes may be required to increase the L-guluronic acid content of the alginate when the organism is grown at high oxygen tensions (73). Some of the six alginate lyase enzymes encoded by the genome may be responsible for altering the composition of alginate in response to oxygen availability (Fig. 3). *P. stutzeri* does not possess genes for alginate biosynthesis and therefore is not able to take advantage of this mechanism for reducing oxygen diffusion into the cell.

In addition to the oxygen diffusion barrier and respiratory removal of oxygen at the cell surface, a number of other strategies may be used by *A. vinelandii* to protect oxygen-sensitive enzymes in the cytoplasm. These include maintenance of a low redox state and efficient energy metabolism (61). In the latter context, it is interesting that *A. vinelandii* contains a second operon encoding an F-type ATP synthase (Avin19670 to Avin19750) that is not closely related to the ancestral operon, with its closest relative being present in the genus *Burkholderia*. The redundancy of oxygen protection mechanisms may explain why *A. vinelandii* DJ is able to fix nitrogen under standard atmospheric conditions in spite of its inability to produce alginate.

The *A. vinelandii* genome encodes several unexpected anaerobic enzymes in addition to the three well-characterized oxygen-sensitive nitrogenases and the membrane-bound [NiFe] hydrogenase (Fig. 3). The presence of NAD^+-independent FDH is unusual for a strict aerobe that does not possess a respiratory nitrate reductase. Even more intriguing is the presence of genes predicted to encode a Ni-dependent CODH, although a homolog of the heme-containing CO sensor, CooA, has been found previously (98). These proteins are characteristic of anaerobic CO metabolism and contrast with the molybdopterin-containing CODHs found in aerobes (71). CO oxidation by *A. vinelandii* has not been reported previously, although it may be an important detoxification process given the CO sensitivity of hydrogenase and respiratory enzymes. The electron acceptor for CO oxidation could potentially be provided by the gene cluster encoding the soluble chimeric NiFe hydrogenase (Avin04360 to Avin04410), which intriguingly is positioned in the corresponding locus to that of the membrane-bound CO-tolerant NiFe hydrogenase in *R. rubrum*. It has been suggested that this energy-conserving hydrogenase produces a proton gradient to drive ATP synthesis by coupling CO-dependent H_2 production and proton translocation in *R. rubrum* (26). Clearly, the role of CO in *A. vinelandii* metabolism requires further investigation.

The ability of *A. vinelandii* to reconcile an obligate aerobic lifestyle with the maintenance of fundamental oxygen-sensitive processes, such as nitrogen fixation, is a remarkable metabolic accomplishment that has important implications for biotechnological exploitation. *A. vinelandii* is a model organism for biochemical studies on the basis of the high yield and quality of enzymes that can be prepared from it. Moreover, powerful genetic approaches facilitated by homologous recombination are readily available, as are a stringent system for controlled expression of proteins encoded within the genome (38) and a means to achieve high-level protein expression (19). Given the high quality of proteins purified from *A. vinelandii*, these combined genetic and biochemical tools make this an ideal organism for the production of enzymes, particularly those that are oxygen sensitive.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant 0523357 to D.W.W. and by M. J. Murdock Charitable Trust Life Sciences Program grants (2004262:JVZ and 2006245:JVZ) to D.W.W.

We thank Paul Rudnick, Nirav Merchant, and other contributors at the University of Arizona, the Joint Genome Institute, and the Oak Ridge National Laboratory who developed the first draft version of the *A. vinelandii* genome. We also thank the students at Hiram College (50 students), Virginia Tech (9 students), and Seattle Pacific University (94 students) who participated in the deep annotation of this genome as part of their undergraduate coursework.

This report is dedicated to the memory of Christina Kennedy (1945–2009). Christina made numerous contributions to our understanding of nitrogen fixation and *Azotobacter* biology over her nearly 4-decade-long career, culminating in the publication of its genome.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Bandyopadhyay, S., S. G. Naik, I. P. O'Carroll, B. H. Huynh, D. R. Dean, M. K. Johnson, and P. C. Dos Santos.** 2008. A proposed role for the *Azotobacter vinelandii* NfuA protein as an intermediate iron-sulfur cluster carrier. J. Biol. Chem. **283:**14092–14099.
3. **Barrios, H., B. Valderrama, and E. Morett.** 1999. Compilation and analysis of sigma(54)-dependent promoter sequences. Nucleic Acids Res. **27:**4305–4313.
4. **Berleman, J. E., B. M. Hasselbring, and C. E. Bauer.** 2004. Hypercyst mutants in *Rhodospirillum centenum* identify regulatory loci involved in cyst cell differentiation. J. Bacteriol. **186:**5834–5841.
5. **Bertsova, Y. V., A. V. Bogachev, and V. P. Skulachev.** 2001. Noncoupled NADH:ubiquinone oxidoreductase of *Azotobacter vinelandii* is required for diazotrophic growth at high oxygen concentrations. J. Bacteriol. **183:**6869–6874.
6. **Bertsova, Y. V., A. V. Bogachev, and V. P. Skulachev.** 1998. Two NADH: ubiquinone oxidoreductases of *Azotobacter vinelandii* and their role in the respiratory protection. Biochim. Biophys. Acta **1363:**125–133.
7. **Bjerkan, T. M., C. L. Bender, H. Ertesvag, F. Drablos, M. K. Fakhr, L. A. Preston, G. Skjak-Braek, and S. Valla.** 2004. The *Pseudomonas syringae* genome encodes a combined mannuronan C-5-epimerase and O-acetylhydrolase, which strongly enhances the predicted gel-forming properties of alginates. J. Biol. Chem. **279:**28920–28929.
8. **Buhrke, T., O. Lenz, N. Krauss, and B. Friedrich.** 2005. Oxygen tolerance of the H_2-sensing [NiFe] hydrogenase from *Ralstonia eutropha* H16 is based on limited access of oxygen to the active site. J. Biol. Chem. **280:**23791–23796.
9. **Burgdorf, T., E. van der Linden, M. Bernhard, Q. Y. Yin, J. W. Back, A. F. Hartog, A. O. Muijsers, C. G. de Koster, S. P. Albracht, and B. Friedrich.** 2005. The soluble NAD^+-reducing [NiFe]-hydrogenase from *Ralstonia eutropha* H16 consists of six subunits and can be specifically activated by NADPH. J. Bacteriol. **187:**3122–3132.
10. **Burk, D.** 1930. The influence of oxygen gas upon the organic catalysis of nitrogen fixation by *Azotobacter*. J. Phys. Chem. **34:**1195–1209.
11. **Bush, J. A., and P. W. Wilson.** 1959. A non-gummy chromogenic strain of Azotobacter vinelandii. Nature **184:**381.
12. **Campos, M., J. M. Martinez-Salazar, L. Lloret, S. Moreno, C. Nunez, G. Espin, and G. Soberon-Chavez.** 1996. Characterization of the gene coding for GDP-mannose dehydrogenase (*algD*) from *Azotobacter vinelandii*. J. Bacteriol. **178:**1793–1799.
13. **Castaneda, M., J. Sanchez, S. Moreno, C. Nunez, and G. Espin.** 2001. The global regulators GacA and sigma(S) form part of a cascade that controls alginate production in *Azotobacter vinelandii*. J. Bacteriol. **183:**6787–6793.
14. **Castresana, J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17:**540–552.
15. **Chen, J. C., and L. E. Mortenson.** 1992. Identification of six open reading frames from a region of the Azotobacter vinelandii genome likely involved in dihydrogen metabolism. Biochim. Biophys. Acta **1131:**199–202.

16. **Cosseau, C., and J. Batut.** 2004. Genomics of the *ccoNOQP*-encoded *cbb3* oxidase complex in bacteria. Arch. Microbiol. **181:**89–96.
17. **Cunningham, L., M. Pitt, and H. D. Williams.** 1997. The *cioAB* genes from *Pseudomonas aeruginosa* code for a novel cyanide-insensitive terminal oxidase related to the cytochrome bd quinol oxidases. Mol. Microbiol. **24:**579–591.
18. **Curatti, L., C. S. Brown, P. W. Ludden, and L. M. Rubio.** 2005. Genes required for rapid expression of nitrogenase activity in *Azotobacter vinelandii*. Proc. Natl. Acad. Sci. USA **102:**6291–6296.
19. **Dos Santos, P. C., D. C. Johnson, B. E. Ragle, M. C. Unciuleac, and D. R. Dean.** 2007. Controlled expression of *nif* and *isc* iron-sulfur protein maturation components reveals target specificity and limited functional replacement between the two systems. J. Bacteriol. **189:**2854–2862.
20. **Edgar, R. C.** 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinform. **5:**113.
21. **Edwards, S. E., C. S. Loder, G. Wu, H. Corker, B. W. Bainbridge, S. Hill, and R. K. Poole.** 2000. Mutation of cytochrome bd quinol oxidase in reduced stationary phase survival, iron deprivation, metal toxicity, and oxidative stress in *Azotobacter vinelandii*. FEMS Microbiol. Lett. **185:**71–77.
22. **Eisen, J. A., J. F. Heidelberg, O. White, and S. L. Salzberg.** 2000. Evidence for symmetric chromosomal inversions around the replication origin in bacteria. Genome Biol. **1:**RESEARCH0011.
23. **Ewing, B., and P. Green.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8:**186–194.
24. **Ewing, B., L. Hillier, M. C. Wendl, and P. Green.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8:**175–185.
25. **Finn, R. D., J. Tate, J. Mistry, P. C. Coggill, S. J. Sammut, H. R. Hotz, G. Ceric, K. Forslund, S. R. Eddy, E. L. Sonnhammer, and A. Bateman.** 2008. The Pfam protein families database. Nucleic Acids Res. **36:**D281–D288.
26. **Fox, J. D., Y. He, D. Shelver, G. P. Roberts, and P. W. Ludden.** 1996. Characterization of the region encoding the CO-induced hydrogenase of *Rhodospirillum rubrum*. J. Bacteriol. **178:**6200–6208.
27. **Frazzon, J., and D. R. Dean.** 2002. Biosynthesis of the nitrogenase iron-molybdenum-cofactor from Azotobacter vinelandii. Met. Ions Biol. Syst. **39:**163–186.
28. **Fukui, T., N. Shiomi, and Y. Doi.** 1998. Expression and characterization of (*R*)-specific enoyl coenzyme A hydratase involved in polyhydroxyalkanoate biosynthesis by *Aeromonas caviae*. J. Bacteriol. **180:**667–673.
29. **Funa, N., H. Ozawa, A. Hirata, and S. Horinouchi.** 2006. Phenolic lipid synthesis by type III polyketide synthases is essential for cyst formation in Azotobacter vinelandii. Proc. Natl. Acad. Sci. USA **103:**6356–6361.
30. **Galindo, E., C. Pena, C. Nunez, D. Segura, and G. Espin.** 2007. Molecular and bioengineering strategies to improve alginate and polydydroxyalkanoate production by Azotobacter vinelandii. Microb. Cell Fact. **6:**7.
31. **Garg, R. P., A. L. Menon, K. Jacobs, R. M. Robson, and R. L. Robson.** 1994. The hypE gene completes the gene cluster for H_2-oxidation in Azotobacter vinelandii. J. Mol. Biol. **236:**390–396.
32. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
33. **Grabbe, R., and R. A. Schmitz.** 2003. Oxygen control of *nif* gene expression in *Klebsiella pneumoniae* depends on NifL reduction at the cytoplasmic membrane by electrons derived from the reduced quinone pool. Eur. J. Biochem. **270:**1555–1566.
34. **Hanzelmann, P., and H. Schindelin.** 2004. Crystal structure of the S-adenosylmethionine-dependent enzyme MoaA and its implications for molybdenum cofactor deficiency in humans. Proc. Natl. Acad. Sci. USA **101:**12870–12875.
35. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754–755.
36. **Jeon, W. B., J. Cheng, and P. W. Ludden.** 2001. Purification and characterization of membrane-associated CooC protein and its functional role in the insertion of nickel into carbon monoxide dehydrogenase from *Rhodospirillum rubrum*. J. Biol. Chem. **276:**38602–38609.
37. **Joerger, R. D., M. R. Jacobson, and P. E. Bishop.** 1989. Two *nifA*-like genes required for expression of alternative nitrogenases by *Azotobacter vinelandii*. J. Bacteriol. **171:**3258–3267.
38. **Johnson, D. C., M.-C. Unciuleac, and D. R. Dean.** 2006. Controlled expression and functional analysis of iron-sulfur cluster biosynthetic components within *Azotobacter vinelandii*. J. Bacteriol. **188:**7551–7561.
39. **Jünemann, S., and J. M. Wrigglesworth.** 1995. Cytochrome bd oxidase from Azotobacter vinelandii. J. Biol. Chem. **270:**16213–16220.
40. **Kelly, M. J., R. K. Poole, M. G. Yates, and C. Kennedy.** 1990. Cloning and mutagenesis of genes encoding the cytochrome *bd* terminal oxidase complex in *Azotobacter vinelandii*: mutants deficient in the cytochrome *d* complex are unable to fix nitrogen in air. J. Bacteriol. **172:**6010–6019.
41. **Kennedy, C., P. Rudnick, T. MacDonald, and T. Melton.** 2005. Genus *Azotobacter*, p. 384–401. *In* G. M. Garrity (ed.), Bergey's manual of sytematic bacteriology, vol. 2, part B. Springer-Verlag, New York, NY.
42. **Kennedy, C., and A. Toukdarian.** 1987. Genetics of azotobacters: applications to nitrogen fixation and related aspects of metabolism. Annu. Rev. Microbiol. **41:**227–258.
43. **Kerby, R. L., P. W. Ludden, and G. P. Roberts.** 1997. In vivo nickel insertion into the carbon monoxide dehydrogenase of *Rhodospirillum rubrum*: molecular and physiological characterization of *cooCTJ*. J. Bacteriol. **179:**2259–2266.
44. **Kow, Y. W., and R. H. Burris.** 1984. Purification and properties of membrane-bound hydrogenase from *Azotobacter vinelandii*. J. Bacteriol. **159:**564–569.
45. **Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and S. L. Salzberg.** 2004. Versatile and open software for comparing large genomes. Genome Biol. **5:**R12.
46. **Latreille, P., S. Norton, B. S. Goldman, J. Henkhaus, N. Miller, B. Barbazuk, H. B. Bode, C. Darby, Z. Du, S. Forst, S. Gaudriault, B. Goodner, H. Goodrich-Blair, and S. Slater.** 2007. Optical mapping as a routine tool for bacterial genome sequence finishing. BMC Genomics **8:**321.
47. **Lederberg, J., and E. L. Tatum.** 1946. Gene recombination in Escherichia coli. Nature **158:**558.
48. **Leon, R., and G. Espin.** 2008. flhDC, but not fleQ, regulates flagella biogenesis in Azotobacter vinelandii, and is under AlgU and CydR negative control. Microbiology **154:**1719–1728.
49. **Leopoldini, M., S. G. Chiodo, M. Toscano, and N. Russo.** 2008. Reaction mechanism of molybdoenzyme formate dehydrogenase. Chemistry **14:**8674–8681.
50. **Li, L., C. J. Stoeckert, Jr., and D. S. Roos.** 2003. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. **13:**2178–2189.
51. **Lineweaver, H., and D. Burk.** 1934. The determination of enzyme dissociation constants. J. Am. Chem. Soc. **56:**658–666.
52. **Linkerhagner, K., and J. Oelze.** 1995. Hydrogenase does not confer significant benefits to *Azotobacter vinelandii* growing diazotrophically under conditions of glucose limitation. J. Bacteriol. **177:**6018–6020.
53. **Linkerhagner, K., and J. Oelze.** 1997. Nitrogenase activity and regeneration of the cellular ATP pool in *Azotobacter vinelandii* adapted to different oxygen concentrations. J. Bacteriol. **179:**1362–1367.
54. **Ma, K., R. Weiss, and M. W. Adams.** 2000. Characterization of hydrogenase II from the hyperthermophilic archaeon *Pyrococcus furiosus* and assessment of its role in sulfur reduction. J. Bacteriol. **182:**1864–1871.
55. **Martinez-Salazar, J. M., S. Moreno, R. Najera, J. C. Boucher, G. Espin, G. Soberon-Chavez, and V. Deretic.** 1996. Characterization of the genes coding for the putative sigma factor AlgU and its regulators MucA, MucB, MucC, and MucD in *Azotobacter vinelandii* and evaluation of their roles in alginate biosynthesis. J. Bacteriol. **178:**1800–1808.
56. **Menon, A. L., L. E. Mortenson, and R. L. Robson.** 1992. Nucleotide sequences and genetic analysis of hydrogen oxidation (*hox*) genes in *Azotobacter vinelandii*. J. Bacteriol. **174:**4549–4557.
57. **Moreno, S., R. Najera, J. Guzman, G. Soberon-Chavez, and G. Espin.** 1998. Role of alternative sigma factor AlgU in encystment of *Azotobacter vinelandii*. J. Bacteriol. **180:**2766–2769.
58. **Moshiri, F., J. Kim, C. Fu, and R. Maier.** 1994. The FeSII protein of Azotobacter vinelandii is not essential for aerobic nitrogen fixation but confers significant protection to oxygen-mediated inactivation of nitrogenase in vitro and in vivo. Mol. Microbiol. **14:**101–114.
59. **Muhammadi, and N. Ahmed.** 2007. Genetics of bacterial alginate: alginate genes distribution, organization and biosynthesis in bacteria. Curr. Genomics **8:**191–202.
60. **Nunez, C., S. Moreno, G. Soberon-Chavez, and G. Espin.** 1999. The *Azotobacter vinelandii* response regulator AlgR is essential for cyst formation. J. Bacteriol. **181:**141–148.
61. **Oelze, J.** 2000. Respiratory protection of nitrogenase in *Azotobacter* species: is a widely held hypothesis unequivocally supported by experimental evidence? FEMS Microbiol. Rev. **24:**321–333.
62. **Ortiz, P. J., and S. Ochoa.** 1959. Studies on polynucleotides synthesized by polynucleotide phosphorylase. IV. P^{32}-labeled ribonucleic acid. J. Biol. Chem. **234:**1208–1212.
63. **Page, W. J., J. Manchak, and B. Rudy.** 1992. Formation of poly(hydroxybutyrate-co-hydroxyvalerate) by *Azotobacter vinelandii* UWD. Appl. Environ. Microbiol. **58:**2866–2873.
64. **Page, W. J., A. Tindale, M. Chandra, and E. Kwon.** 2001. Alginate formation in *Azotobacter vinelandii* UWD during stationary phase and the turnover of poly-beta-hydroxybutyrate. Microbiology **147:**483–490.
65. **Pau, R. N., and D. M. Lawson.** 2002. Transport, homeostasis, regulation, and binding of molybdate and tungstate to proteins. Met. Ions Biol. Syst. **39:**31–74.
66. **Peralta-Gil, M., D. Segura, J. Guzman, L. Servin-Gonzalez, and G. Espin.** 2002. Expression of the *Azotobacter vinelandii* poly-beta-hydroxybutyrate biosynthetic *phbBAC* operon is driven by two overlapping promoters and is dependent on the transcriptional activator PhbR. J. Bacteriol. **184:**5672–5677.
67. **Pettinari, M. J., G. J. Vazquez, D. Silberschmidt, B. Rehm, A. Steinbuchel, and B. S. Mendez.** 2001. Poly(3-hydroxybutyrate) synthesis genes in *Azotobacter* sp. strain FA8. Appl. Environ. Microbiol. **67:**5331–5334.
68. **Poole, R. K., and S. Hill.** 1997. Respiratory protection of nitrogenase activity in *Azotobacter vinelandii*—roles of the terminal oxidases. Biosci. Rep. **17:** 303–317.

69. **Premakumar, R., T. M. Loveless, and P. E. Bishop.** 1994. Effect of amino acid substitutions in a potential metal-binding site of AnfA on expression from the *anfH* promoter in *Azotobacter vinelandii*. J. Bacteriol. **176:**6139–6142.
70. **Qiu, D., V. M. Eisinger, D. W. Rowen, and H. D. Yu.** 2007. Regulated proteolysis controls mucoid conversion in Pseudomonas aeruginosa. Proc. Natl. Acad. Sci. USA **104:**8107–8112.
71. **Ragsdale, S. W.** 2004. Life with carbon monoxide. Crit. Rev. Biochem. Mol. Biol. **39:**165–195.
72. **Roberts, G. P., R. L. Kerby, H. Youn, and M. Conrad.** 2005. CooA, a paradigm for gas sensing regulatory proteins. J. Inorg. Biochem. **99:**280–292.
73. **Sabra, W., A. P. Zeng, H. Lunsdorf, and W. D. Deckwer.** 2000. Effect of oxygen on formation and structure of *Azotobacter vinelandii* alginate and its role in protecting nitrogenase. Appl. Environ. Microbiol. **66:**4037–4044.
74. **Sadoff, H. L.** 1975. Encystment and germination in *Azotobacter vinelandii*. Bacteriol. Rev. **39:**516–539.
75. **Seefeldt, L. C., and D. J. Arp.** 1986. Purification to homogeneity of Azotobacter vinelandii hydrogenase: a nickel and iron containing alpha beta dimer. Biochimie **68:**25–34.
76. **Segura, D., T. Cruz, and G. Espin.** 2003. Encystment and alkylresorcinol production by *Azotobacter vinelandii* strains impaired in poly-beta-hydroxybutyrate synthesis. Arch. Microbiol. **179:**437–443.
77. **Segura, D., E. Vargas, and G. Espin.** 2000. Beta-ketothiolase genes in *Azotobacter vinelandii*. Gene **260:**113–120.
78. **Segura, D., O. Vite, Y. Romero, S. Moreno, M. Castaneda, and G. Espin.** 2009. Isolation and characterization of *Azotobacter vinelandii* mutants impaired in alkylresorcinol synthesis: alkylresorcinols are not essential for cyst desiccation resistance. J. Bacteriol. **191:**3142–3148.
79. **Stamatakis, A.** 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:**2688–2690.
80. **Steigedal, M., H. Sletta, S. Moreno, M. Maerk, B. E. Christensen, T. Bjerkan, T. E. Ellingsen, G. Espin, H. Ertesvag, and S. Valla.** 2008. The *Azotobacter vinelandii* AlgE mannuronan C-5-epimerase family is essential for the in vivo control of alginate monomer composition and for functional cyst formation. Environ. Microbiol. **10:**1760–1770.
81. **Stevenson, L. H., and M. D. Socolofsky.** 1966. Cyst formation and poly-beta-hydroxybutyric acid accumulation in *Azotobacter*. J. Bacteriol. **91:**304–310.
82. **Svanem, B. I., W. I. Strand, H. Ertesvag, G. Skjak-Braek, M. Hartmann, T. Barbeyron, and S. Valla.** 2001. The catalytic activities of the bifunctional Azotobacter vinelandii mannuronan C-5-epimerase and alginate lyase AlgE7 probably originate from the same active site in the enzyme. J. Biol. Chem. **276:**31542–31550.
83. **Thorneley, R. N., and G. A. Ashby.** 1989. Oxidation of nitrogenase iron protein by dioxygen without inactivation could contribute to high respiration rates of *Azotobacter* species and facilitate nitrogen fixation in other aerobic environments. Biochem. J. **261:**181–187.
84. **Thorneley, R. N., and D. J. Lowe.** 1983. Nitrogenase of *Klebsiella pneumoniae*. Kinetics of the dissociation of oxidized iron protein from molybdenum-iron protein: identification of the rate-limiting step for substrate reduction. Biochem. J. **215:**393–403.
85. **Tishkov, V. I., and V. O. Popov.** 2004. Catalytic mechanism and application of formate dehydrogenase. Biochemistry (Moscow) **69:**1252–1267.
86. **Trujillo-Roldan, M. A., S. Moreno, D. Segura, E. Galindo, and G. Espin.** 2003. Alginate production by an *Azotobacter vinelandii* mutant unable to produce alginate lyase. Appl. Microbiol. Biotechnol. **60:**733–737.
87. **Ugidos, A., G. Morales, E. Rial, H. D. Williams, and F. Rojo.** 2008. The coordinate regulation of multiple terminal oxidases by the *Pseudomonas putida* ANR global regulator. Environ. Microbiol. **10:**1690–1702.
88. **Valverde, A., Y. Okon, and S. Burdman.** 2006. cDNA-AFLP reveals differentially expressed genes related to cell aggregation of *Azospirillum brasilense*. FEMS Microbiol. Lett. **265:**186–194.
89. **Vignais, P. M.** 2008. Hydrogenases and H(+)-reduction in primary energy conservation. Results Probl. Cell Differ. **45:**223–252.
90. **Vignais, P. M., B. Billoud, and J. Meyer.** 2001. Classification and phylogeny of hydrogenases. FEMS Microbiol. Rev. **25:**455–501.
91. **Wattt, R. K., and P. W. Ludden.** 1999. Nickel-binding proteins. Cell. Mol. Life Sci. **56:**604–625.
92. **Wilson, P. W., and S. G. Knight.** 1952. Experiments in bacterial physiology. Burgess Publishing Co., Minneapolis, MN.
93. **Wu, G., H. Cruz-Ramos, S. Hill, J. Green, G. Sawers, and R. K. Poole.** 2000. Regulation of cytochrome bd expression in the obligate aerobe Azotobacter vinelandii by CydR (Fnr). Sensitivity to oxygen, reactive oxygen species, and nitric oxide. J. Biol. Chem. **275:**4679–4686.
94. **Wu, G., S. Hill, M. J. Kelly, G. Sawers, and R. K. Poole.** 1997. The cydR gene product, required for regulation of cytochrome bd expression in the obligate aerobe Azotobacter vinelandii, is an Fnr-like protein. Microbiology **143:**2197–2207.
95. **Wu, G., A. J. Moir, G. Sawers, S. Hill, and R. K. Poole.** 2001. Biosynthesis of poly-beta-hydroxybutyrate (PHB) is controlled by CydR (Fnr) in the obligate aerobe Azotobacter vinelandii. FEMS Microbiol. Lett. **194:**215–220.
96. **Yan, Y., J. Yang, Y. Dou, M. Chen, S. Ping, J. Peng, W. Lu, W. Zhang, Z. Yao, H. Li, W. Liu, S. He, L. Geng, X. Zhang, F. Yang, H. Yu, Y. Zhan, D. Li, Z. Lin, Y. Wang, C. Elmerich, M. Lin, and Q. Jin.** 2008. Nitrogen fixation island and rhizosphere competence traits in the genome of root-associated Pseudomonas stutzeri A1501. Proc. Natl. Acad. Sci. USA **105:**7564–7569.
97. **Yates, M. G., and F. O. Campbell.** 1989. The effect of nutrient limitation on the competition between H_2-uptake hydrogenase positive (Hup^+) recombinant strain of *Azotobacter chroococcum* and the Hup^- mutant parent in mixed populations. J. Gen. Microbiol. **135:**221–226.
98. **Youn, H., R. L. Kerby, M. Conrad, and G. P. Roberts.** 2004. Functionally critical elements of CooA-related CO sensors. J. Bacteriol. **186:**1320–1329.

Paper II

Is not included due to copyright

Paper III

Is not included due to copyright

Paper IV

Is not included due to copyright