ABSTRACT

NOAR, JESSE DAVID. Aerobic Hydrogen Production in *Azotobacter vinelandii*: Genetics, Physiology, and Optimization. (Under the direction of José M. Bruno-Bárcena).

*Azotobacter vinelandii* is an obligate aerobic soil microbe that is highly adapted to fix nitrogen across a wide variety of oxygen concentrations and other conditions. Part of this adaptation is the possession of three different versions of the nitrogen-fixing enzyme, nitrogenase. A mutant strain, strain CA6, was discovered previously that can fix nitrogen in the presence of tungsten, an element that poisons or induces repression of all three nitrogenases in the wild-type. We analyzed the genome of this mutant and discovered a large deletion of more than 40 open reading frames; this deletion explained the observed phenotypes and suggested others yet undiscovered. One such phenotype, hydrogen production, is caused by deletion of the uptake hydrogenase, allowing the hydrogen that the nitrogenases produce as a byproduct to escape into the atmosphere. Comparison of yields of biomass, carbon dioxide, and hydrogen from CA6 and another strain lacking only a single uptake hydrogenase gene showed equal production in high-aeration, carbon-limiting, iron-sufficient chemostat conditions. However, in the presence of tungsten, CA6’s hydrogen yield increased 4.5-fold, while the other strain was unable to grow. In lower aeration conditions, hydrogen and biomass yields increased due to lower maintenance energy requirements. Some granules of storage polymer were produced when oxygen was limiting, but this helped to increase rather than decrease hydrogen yields. Overall, with low aeration in the presence of tungsten, *A. vinelandii* CA6 achieved an increase over 9-fold in hydrogen yield compared to high-aeration, tungsten-free growth. This aerobic hydrogen production, especially in
combination with the production of fixed nitrogen, could provide value as an industrial process and serve as a model for future research.
Aerobic Hydrogen Production in *Azotobacter vinelandii*: Genetics, Physiology, and Optimization

by

Jesse David Noar

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Microbiology

Raleigh, North Carolina

2016

APPROVED BY:

_______________________________  ______________________________
José M. Bruno-Bárcena               Jonathan W. Olson
Committee Chair

_______________________________  ______________________________
Paul T. Hamilton                   John D. Sheppard
BIOGRAPHY

Jesse Noar grew up in some suburbs of New Jersey, but despite that has always been pretty fond of nature. He spent a good amount of time playing outdoors and reading about how things work.

In high school, an especially good biology teacher introduced Jesse to techniques of microbiological culture and told him to separate and identify the organisms in a mixed culture, which he found so enjoyable that he decided to make a career of studying bacteria.

In his undergraduate years at Cornell University, Jesse did well in his microbiology classes and also worked in several labs, gaining experience and research enough to complete an honors thesis. After graduating, he decided to remain in his current lab and gather enough data to publish his first scientific journal article.

With a few years’ worth of research experience, Jesse decided that it was time to aim for a PhD, and enrolled in the Microbiology graduate program at North Carolina State University, working in the lab of Dr. Bruno-Bárcena. Funded by the College of Agriculture and Life Sciences at NCSU and from July 2011 to 2014 by a NSF graduate research fellowship, he has researched a strain of diazotroph that produces hydrogen gas, as described herein.
ACKNOWLEDGMENTS

I must acknowledge the immense contributions of my advisor, Dr. Bruno-Bárcena: I could not ask for a better mentor, with more dedication to the success of his students. The contributions of teaching, training, and mentorship from the other members of my committee, professors of my classes, and other faculty in the department are also worthy of acknowledgement.

I would also like to thank my classmates in the Microbiology program—those that started before, at the same time, or after I started—for friendship, collaboration, and support. And many thanks to the students and other members of the Bruno lab; it has been a great place to work.

Thanks to the National Science Foundation for their support over the three years of my Graduate Research Fellowship, to NC State for its support of the lab and my education, and to our collaborators for their assistance and expertise.

Finally, but not least, thanks to my family for their support and advice, and to my wife Roslyn for her many valuable suggestions and encouragement throughout the years.
TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................................................ V

LIST OF FIGURES ..................................................................................................................................... VI

CHAPTER 1. AZOTOBACTER VINELANDII: ONCE AND FUTURE BACTERIUM OF INTEREST ............................................................... 1
  Introduction ........................................................................................................................................... 1
  General Characteristics of Azotobacter vinelandii ............................................................................... 2
  Historical Interest ................................................................................................................................. 5
  Promising Areas for A. vinelandii Research ......................................................................................... 18
  Conclusions ......................................................................................................................................... 40

CHAPTER 2. GENOMIC COMPARISON OF AZOTOBACTER VINELANDII MUTANT STRAIN CA6 WITH ITS PARENT STRAIN CA ................................................. 41
  Introduction ......................................................................................................................................... 41
  Materials and Methods ......................................................................................................................... 41
  Results ................................................................................................................................................ 46
  Discussion .......................................................................................................................................... 59

CHAPTER 3. PHYSIOLOGICAL EVALUATION OF AZOTOBACTER VINELANDII CA6 MUTANT PHENOTYPES, ESPECIALLY HYDROGEN PRODUCTION .......... 64
  Introduction ......................................................................................................................................... 64
  Methods ............................................................................................................................................. 65
  Results ................................................................................................................................................ 70
  Discussion .......................................................................................................................................... 75

CHAPTER 4. EVALUATION OF AZOTOBACTER VINELANDII BIOMASS AND HYDROGEN YIELDS GROWING UNDER INCREASING DISSOLVED OXYGEN EXPOSURE ............................................................................................................................. 78
  Introduction ......................................................................................................................................... 78
  Methods ............................................................................................................................................. 79
  Results ................................................................................................................................................ 86
  Discussion .......................................................................................................................................... 95

CHAPTER 5. HYDROGEN AND AZOTOBACTER VINELANDII: FUTURE POTENTIAL DIRECTIONS ............................................................................................. 101

REFERENCES .......................................................................................................................................... 104
LIST OF TABLES

Table 2.1. Primers used ...........................................................................................................44
Table 2.2. Sequencing Statistics .............................................................................................47
Table 2.3. Variations Discovered in Genome Comparison .....................................................52
Table 2.4. Genes and putative genes deleted in CA6, and putative homologs elsewhere in the genome..................................................................................................................................54
Table 3.1. Kinetic and yield parameters from stepwise chemostat experiments, calculated from slopes of lines in Figure 3.4........................................................................................................74
Table 4.1. Strains, plasmids, and primers ..................................................................................80
Table 4.2. Poly-β-hydroxybutyrate (PHB) observed in cells at different agitation speeds ....91
LIST OF FIGURES

Figure 1.1. “Azotobacter” as a percentage of all words in books written in English and published in the United States of America from 1900 to 2016 .........................................................2
Figure 1.2. Three versions of nitrogenase with different central cofactors in wild-type Azotobacter vinelandii, in the presence of different metals .................................................................9
Figure 1.3. Pathway from glucose to PHB in A. vinelandii .................................................................19
Figure 1.4. Nitrogenase produces hydrogen gas by reduction of protons ........................................35
Figure 2.1. FastQC analysis of A. vinelandii CA and CA6 454 reads .................................................45
Figure 2.2. Proportion of genome covered to at least a minimum depth .........................................47
Figure 2.3. False variation detection rates .......................................................................................49
Figure 2.4. Schematic representation of a section of CA’s genome deleted in CA6 ....................56
Figure 2.5. Atlases of variations between strains ............................................................................57
Figure 3.1. Strains of A. vinelandii used in this study ......................................................................65
Figure 3.2. Peak hydrogen concentrations in strains of A. vinelandii after pulse with iron sulfate + sucrose or sucrose alone .......................................................................................71
Figure 3.3. Transient response to nitrogen pulses ............................................................................72
Figure 3.4. Generation rates of biomass (dry weight), CO₂, and hydrogen in response to increases in sucrose feeding rates ........................................................................................73
Figure 4.1. Example of cessation of hydrogen production and concurrent increase in CO₂ production upon increase in agitation ..................................................................................88
Figure 4.2. Oxygen and sucrose utilization in steady states at different agitation speeds .....89
Figure 4.3. Hydrogen, dry weight, and CO₂ yields in steady states at different agitation speeds ..............................................................................................................................90
Figure 4.4. The presence or absence of PHB granules in A. vinelandii strains stained with Nile Blue A ............................................................................................................................92
Figure 4.5. Hydrogen yields in strain CA6 in steady state at each agitation speed, with or without tungsten added .................................................................................................95
Figure 4.6. Ammonia productivity in steady state at each agitation speed in DJ strains .........100
Figure 5.1. Hydrogen yields in A. vinelandii CA6 growing in chemostat culture in different conditions .................................................................101
CHAPTER 1. *Azotobacter vinelandii*: Once and Future Bacterium of Interest

Introduction

Soon after the discovery of *Azotobacter vinelandii* in 1903\(^1\), it was adopted as a model organism for the study of aerobic nitrogen fixation, respiration, microbial physiology, hydrogen production and assimilation, and other enzyme kinetics. Multiple important discoveries have been made studying azotobacters, including Lineweaver-Burk kinetic models and the genetic code\(^2,3\). The genus’s features of most interest, though, have been its nitrogenases and nitrogen-fixing capabilities\(^4-6\) and aerobic physiology\(^7-9\); its respiration rates are some of the highest observed\(^9-12\).

Interest in studying these microbes has waned since the early 1970s (Figure 1.1), but they may be due for a comeback. Effort is being made to sequence *Azotobacter* genomes\(^13-15\) and tap into their biotechnological and agricultural potential\(^16-20\). This review aims to revisit the discoveries of the past and lay out the prospects for discoveries in the future, including the study of hydrogen production that will be the focus of later chapters.
General Characteristics of *Azotobacter vinelandii*

There is some suggestion that the genus *Azotobacter* should be subsumed by the genus *Pseudomonas*, based mainly on sequence similarities and alginate production, though different phylogenies give differing results. The genus’s chief characteristics include obligate aerobic lifestyles, diazotrophy, formation of dessication-resistant cysts, production of polyhydroxyalkanoates (PHAs) as storage polymers, and natural competence. Wild-type strains are often motile, and cell morphology is pleomorphic, varying depending on physiological state and growth conditions.

*A. vinelandii* was isolated first from soil in Vineland, New Jersey. It possesses three versions of nitrogenase with different metals in their central cofactors. These
enzymes can substitute one for another to allow growth when one or another essential metal is absent\textsuperscript{45–48}, or at different temperatures\textsuperscript{49}. Growing in iron-limited conditions, the species produces a characteristic neon-green pigment, a diffusible siderophore called azotobactin\textsuperscript{50–52}; it also produces several other siderophore compounds\textsuperscript{51,53}. \textit{A. vinelandii} can use a great variety of carbon substrates, including alcohols, organic acids, and sugars\textsuperscript{10,54,55}, though it displays a preference for certain compounds over others in the form of diauxic growth\textsuperscript{56–58}. This versatility raises the possibility of low-cost, unrefined carbon substrates for biotechnological processes\textsuperscript{57,59,60}.

As a model organism for research, \textit{A. vinelandii} also displays several favorable characteristics. It is naturally competent in certain conditions: it takes up linear DNA and performs double recombination with homologous portions of its genome, permitting stable modifications\textsuperscript{35–37,61,62}. The optimal conditions for transformation depend on a given strain’s characteristics, especially alginate-producing ability, but generally competence is highest in late exponential phase under iron limitation, with sugars or glycerol rather than organic acids as a carbon source\textsuperscript{35,36}. The production of azotobactin is necessary but not sufficient for competence\textsuperscript{36}. Gummy (alginate-producing) strains of \textit{A. vinelandii} are generally less competent than non-gummy strains\textsuperscript{35,61}, presumably because alginate acts as a barrier between the cells and exogenous DNA, but competence in such strains can be enhanced, such as by replacing glucose with other carbon substrates (especially sugar alcohols)\textsuperscript{61}, by increasing phosphate or decreasing calcium concentration\textsuperscript{35}, or by adding cyclic adenosine monophosphate (cAMP)\textsuperscript{62}. In non-gummy strains, such as \textit{A. vinelandii} strain CA, molybdenum limitation and excess magnesium enhance competence\textsuperscript{35,37,61,62}. When
transformation is performed in liquid rather than on solid medium, adding a limited amount of fixed nitrogen is necessary. Competent cells have been observed to contain greater amounts of poly-β-hydroxybutyrate (PHB) than non-competent cells, but it is not clear whether competence and PHB are causally related. Such competence enhances the value of *A. vinelandii* as a research and biotechnology organism, permitting straightforward knocking out, mutating, and introducing new traits in the organism. A strategy has even been developed for markerless genetic modification.

There has been some concern that *A. vinelandii* cells may be polyploid, containing as many as 80 copies of their genome per cell; this could complicate transformation, making it difficult to obtain homozygous transformants, at least at certain loci. However, others have found that segregation behavior in *A. vinelandii* does not appear consistent with polyploidy. A possible explanation is that in rich medium, freshly-divided cells have one or a few copies of their chromosome, while cells in late-exponential or stationary phase have 10- to 50-fold more. This effect is not observed in minimal media. Therefore this effect should not be an insurmountable obstacle to transformation.

In addition to linear DNA, *A. vinelandii* can maintain and be transformed with plasmids of certain kinds. Plasmids of the IncQ and IncP-1 groups have been successfully transformed into and maintained in *A. vinelandii* using conjugation from *Escherichia coli*, electroporation, and the natural competence method of Page and von Tigerstrom discussed previously. Plasmids can also be used to introduce transposons into *A. vinelandii* for mutagenesis. These possibilities further enhance the organism’s usefulness.
Historical Interest

Nitrogen Fixation

Since its discovery as a nitrogen-fixing soil microbe in 1903, *A. vinelandii* has been studied throughout the decades as an aerobic, diazotrophic model organism\(^1,83–89\). Its largest contribution, from discovery to the present day, has been to the study of enzymes relating to nitrogen fixation. As mentioned, *A. vinelandii* has three separate versions of nitrogenase\(^{27,28,41–44,90,91}\), one of which contains an atom of molybdenum (Mo), one with vanadium (V) in the place of Mo, and a third with Fe in place of the heterometal. They all reduce N\(_2\) to NH\(_3\) but have minor differences from each other. They each have two main components: dinitrogenase reductase (DNR) and dinitrogenase\(^92–95\). The former is a single subunit encoded by one gene and the latter comprises at least two subunits (\(\alpha\) and \(\beta\)) that form an \(\alpha_2\beta_2\) tetramer in the Mo nitrogenase\(^95,96\), while the V and Fe nitrogenases also contain a \(\delta\) subunit\(^97,98\). The DNR protein contains a Fe\(_4\)S\(_4\) cluster necessary for its function\(^99–101\). The dinitrogenase also contains four Fe\(_4\)S\(_4\) clusters, called P-clusters, per tetramer\(^102–104\). These facilitate transfer of electrons to the active site, where the central cofactor is located; this cofactor contains the Mo, V, or Fe atom that defines each isoenzyme, along with another 8 Fe and 6 S atoms arranged as a central trigonal prismatic arrangement of atoms\(^104–107\). The central metal in this cofactor is also ligated to a homocitrate residue\(^108–110\). Some studies have investigated the specific amino acid residues essential for the catalysis, especially in the Mo nitrogenase\(^111,112\). Other necessary enzymes are not involved in catalyzing the reaction, but rather assembling the cofactors and inserting them into the apoenzymes\(^90,113–115\). The nitrogenase as a whole is located in the cytoplasm, not associated
with the cell membrane\textsuperscript{88}, and makes up about 10\% of total cell protein when cells are fixing nitrogen\textsuperscript{116}.

The process of dinitrogen reduction begins when the dinitrogenase reductase (DNR) consumes two ATP molecules, allowing the transfer of an electron from an electron carrier such as flavodoxin\textsuperscript{117–120} to the DNR, “charging” it\textsuperscript{121,122}. This electron is then transferred from the DNR to the dinitrogenase, where a molecule of \( \text{N}_2 \) binds\textsuperscript{93,121}. This process repeats at least seven more times (in the case of the Mo nitrogenase), resulting in the addition of six electrons and protons to the molecule \( \text{N}_2 \) and the reduction of two protons to a molecule of \( \text{H}_2 \)\textsuperscript{123}. Nitrogenases can also reduce other substrates, in place of or in competition with dinitrogen and protons. Hydrogen itself may competitively inhibit nitrogen fixation, but this is only observed at high concentrations of hydrogen (over 20\% of atmosphere) and relatively low concentrations of nitrogen\textsuperscript{124,125}. Carbon monoxide (CO) is reduced to ethylene and other hydrocarbons, especially by the V nitrogenase, as discussed later\textsuperscript{126–128}. Acetylene gas (\( \text{C}_2\text{H}_2 \)) is reduced to ethylene (\( \text{C}_2\text{H}_4 \)), and to ethane (\( \text{C}_2\text{H}_6 \)) in the case of V and Fe nitrogenases\textsuperscript{129,130}; this inhibits Mo nitrogenase hydrogen production more than nitrogen gas does, with 86 to 98\% of the electron flux going to acetylene when present at 10\% of the atmosphere\textsuperscript{41,131}. The proportion is much lower (12 to 35\%) in the V nitrogenase\textsuperscript{41,131}. Other possible substrates that can affect the hydrogen yields are azide (\( \text{N}_3^- \)) and cyanide (\( \text{CN}^- \)), which nitrogenase reduces primarily to ammonia and hydrazine or ammonia and methane, respectively\textsuperscript{132–134}. With azide as a substrate, 30 to 50\% of the electron flux through the Mo nitrogenase goes to hydrogen\textsuperscript{132,133}. Cyanide also inhibits more than 80\% of hydrogen production, but adding carbon monoxide can prevent this inhibition, at least partially\textsuperscript{132,135}. 
These figures are similar in the V nitrogenase\textsuperscript{136}. Carbon dioxide can be reduced to methane and other hydrocarbons\textsuperscript{137,138}, and nitrous oxide (N$_2$O) is also a potential substrate\textsuperscript{139}. The ratio of nitrogenase components (DNR and dinitrogenase) affects the rate of electron flow and the proportion of electrons that go to a given substrate when multiple are available\textsuperscript{93,140–142}, the average ratio of dinitrogenase to DNR \textit{in vivo} is 1 to 1.45\textsuperscript{116,118}. As delicate redox catalysts, these enzymes are sensitive to inactivation by various reactive chemicals though. The DNR can be irreversibly inactivated by oxygen exposure\textsuperscript{143}, and though dinitrogenase can reduce CO, azide, N$_2$O, and cyanide somewhat, these compounds inhibit the ability of cells to fix nitrogen\textsuperscript{144,145}. The binding to and reactions of these substrates with nitrogenase has allowed for the study of the enzyme’s biochemistry.

\textit{Uptake Hydrogenase}

The historical study of nitrogen fixation-related enzymes also includes study of the associated hydrogen-oxidizing uptake hydrogenase system in \textit{Azotobacter}. This system consumes hydrogen gas, either nitrogenase-generated or exogenous\textsuperscript{146,147}. In the presence of a strong reducer at low pH, it can reduce protons to produce hydrogen\textsuperscript{146–148}, but these conditions are not found \textit{in vivo}. The system is membrane-bound\textsuperscript{147–149} and contains two peptide subunits, the large and small $\alpha$ and $\beta$ subunits\textsuperscript{147,150}. These are encoded by the \textit{hoxG} and \textit{hoxK} genes, respectively\textsuperscript{151}, though the product of \textit{hoxG} must be processed to remove a carboxyl-terminal stretch before it is functional\textsuperscript{150,152}. The enzyme contains a cofactor cluster containing iron and nickel, making it a NiFe hydrogenase\textsuperscript{147,152,153}. Other enzymes encoded by the \textit{hox} and \textit{hyp} operons are necessary for activity but not directly involved. HypE is
involved in processing the α subunit and insertion into the membrane\textsuperscript{149}; HypB functions in nickel processing and insertion into the enzyme\textsuperscript{153}; and HoxZ carries the hydrogen-derived electrons to the electron transport chain of the cells\textsuperscript{151,154–156}. The functions of the other gene products are not so well defined, but they are also important\textsuperscript{153,154}. Electrons from hydrogen travel to HoxZ and then through the electron transport chain, reducing cytochromes \textit{d} and \textit{b}, but especially cytochrome \textit{c}\textsuperscript{157,158}, and from there go on to reduce oxygen to water\textsuperscript{155}. As \textit{A. vinelandii}'s exposure to hydrogen mostly comes from nitrogenase, the activity of hydrogenase is highly linked to nitrogen fixation\textsuperscript{153,159–161}. This system will be explored more later in the chapter.

\textit{Alternative Nitrogenases}

The presence of alternative nitrogenase isoenzymes was hinted at in earlier studies, in the 60s and early 70s, by observation of the substitution of vanadium for molybdenum\textsuperscript{131,162–165}. In 1980, Bishop and colleagues observed nitrogen-fixing pseudorevertants derived from nitrogenase-negative mutants\textsuperscript{166}, which were able to grow despite the presence of tungsten\textsuperscript{4}, an element that poisons the Mo nitrogenase\textsuperscript{4,167–169}. None of these pseudorevertants had functional a Mo nitrogenase\textsuperscript{4,45}, suggesting the presence of an independent enzyme system. Later, they observed an alternative DNR directly, expressed in the absence of Mo, using two-dimensional gel electrophoresis\textsuperscript{27}, and nitrogen fixation in strains with the Mo nitrogenase genes deleted\textsuperscript{5}. The V nitrogenase was isolated soon after\textsuperscript{41,94,170}, followed by confirmation and isolation of a third nitrogenase containing neither Mo nor V but only Fe\textsuperscript{28,42,43}. These systems were then studied more extensively in \textit{A. vinelandii}\textsuperscript{47,49,90,97,104,106,136,171–178}. 8
Figure 1.2. Three versions of nitrogenase with different central cofactors in wild-type *Azotobacter vinelandii*, in the presence of different metals. A) When molybdenum (Mo) is present, regardless of the presence of vanadium (V), the Mo nitrogenase is active and the others are repressed. B) When Mo is absent and V is present, the V nitrogenase is active and the others are repressed. C) When only iron (Fe) is present, the Fe nitrogenase is active and the others are repressed. D) When tungsten (W) is present, the Mo nitrogenase is present but inactive, and the others are repressed. Source: World J Microbiol Biotechnol, “Protons and pleomorphs: aerobic hydrogen production in Azotobacters,” Vol 32, 2016, p2, J.D. Noar and J.M. Bruno-Bárcena, reprinted with permission of Springer
Nitrogenase Regulation

The regulation of these different versions, by fixed nitrogen and by the presence or absence of the metals required by different versions, has been studied as well (Figure 1.2). The presence of Mo represses activity of the V and Fe nitrogenases while activating Mo nitrogenase activity\textsuperscript{27,44,45,167}. Similarly, if Mo is absent, the presence of V represses the Fe nitrogenase\textsuperscript{43,44,48,167}. If neither is present or if the culture exhausts their supply, \textit{A. vinelandii} depends on the Fe nitrogenase to grow\textsuperscript{43,47}.

Regarding fixed nitrogen sources, \textit{A. vinelandii} is capable of using inorganic sources (ammonium, nitrate, nitrite) and several organic sources: aspartate, asparagine, glutamate, adenine, and urea\textsuperscript{179}. Ammonium salts are the most potent in repressing all forms of nitrogenase activity\textsuperscript{174,179–183}, at concentrations of 25 μM and above\textsuperscript{184}. The mechanism of this shut-off, which is almost immediate, seems to be a decrease in proton motive force, restricting the access of nitrogenase to reducing equivalents that it requires\textsuperscript{182}. Factors that modify the force of this repression include amount of oxygen and the cells’ potential respiratory rate (too little oxygen, or too much without adequate increase in respiration, enhance the repression of nitrogenase)\textsuperscript{6,185}; high pH also enhances repression, which provides supporting evidence for a mechanism involving proton motive force\textsuperscript{185}. Other fixed-nitrogen compounds can induce repression in \textit{A. vinelandii} also: nitrate salts, especially when cells are pre-adapted to them\textsuperscript{179,186}, and urea\textsuperscript{160} can fully repress nitrogen fixation; and organic compounds such as aspartate, adenine, yeast extract, or casamino acids can repress nitrogenase partially\textsuperscript{179,183,187}. \textit{A. vinelandii} does not appear to store or “park” nitrogenase protein inactivated by fixed nitrogen-induced regulation over long periods\textsuperscript{6,116,118}, in the way
that organisms such as *Rhodospirillum rubrum* use via ADP-ribosylation\(^{188}\); inactive protein is broken down in *A. vinelandii*.

**Genetics of Nitrogen Fixation**

In more recent decades (since the 1970s), the genetics and transcriptional regulation of nitrogen fixation have been examined. The genes encoding the structural proteins of the Mo nitrogenase, *nifHDK* (encoding DNR and dinitrogenase α and β subunits, respectively), have been well characterized\(^{113,189–192}\). Other genes in the cluster include *nifUSV*, which are responsible for aspects of enzyme and cofactor synthesis\(^{90,191,193}\); *nifEN*, also required for cofactor synthesis\(^{113,191,194}\); *nifTY*, *nifX*, and *nifF*, involved in enzyme synthesis and other things\(^{191}\); and *nifWZM*\(^{191,193}\). There are also several important *nif* genes in a separate cluster: *nifB*, also required for cofactor synthesis\(^{115,189,195,196}\); and *nifOQ*, of poorly understood function\(^{195}\).

The alternative nitrogenase genes have been characterized as well. The *vnf* operon encodes proteins of the V nitrogenase; many of its genes are homologous to those in the main *nif* cluster. The *vnfH* and *vnfDK* genes encode the DNR and V-containing dinitrogenase α and β subunits, respectively\(^{175,197}\), with the *vnfG* gene, encoding the δ subunit, between *vnfD* and *vnfK*\(^{98,175}\). The products of *vnfENX* are important for enzyme and cofactor synthesis\(^{198}\). Likewise, the Fe nitrogenase genes are found in the *anf* operon: *anfHDGK* encode the DNR and dinitrogenase α, δ, and β subunits\(^{42,98}\). Some genes annotated as *nif* (*nifUSV*) are also essential for activity of both alternative nitrogenases\(^{90}\); alternative homologs of these are not present. The V DNR gene, *vnfH*, is also required for activity of the Fe nitrogenase\(^{173,199}\).
Other genes in the *Azotobacter* genome are also important for nitrogen fixation. Iron is an essential element for all nitrogenases, and *rnf* genes in *A. vinelandii* help with iron clusters in nitrogenase\(^\text{200}\). For electron transfer to nitrogenase, a number of candidate genes have been identified (encoding flavodoxins and ferredoxins), but it’s unclear which are involved or essential, as deleting one or more them does not reduce nitrogenase activity\(^\text{201–203}\). Finally, sequencing the genome of *A. vinelandii* helped to identify other related genes\(^\text{13}\), and cloning the Fe nitrogenase into *E. coli* has helped identify which genes are essential for its function\(^\text{204}\).

The transcriptional regulation of these genes in *A. vinelandii* has also been studied. Studies of the global transcriptome or the nitrogenase operons specifically in different nitrogen-fixing conditions provide a helpful bird’s-eye perspective\(^\text{205}\). Specifically, each nitrogenase version has an activator protein, encoded by the *nifA*, *vnfA*, and *anfA* genes, which regulate expression depending on the presence or absence of necessary metals or fixed nitrogen\(^\text{91,172,206–209}\). Other genes, such as *nfrX*, *ntrC*, and *rpoN*, are involved in nitrogenase regulation also\(^\text{172,210,211}\), to coordinate optimal efficiency of this expensive process. V nitrogenase gene expression is not repressed by ammonium, but enzyme activity is absent, so this nitrogenase must be regulated posttranscriptionally\(^\text{174,206}\). Expression of *vnfH* in general shows some different patterns from those of the other structural *vnf* genes, so its product may have other roles than just dinitrogenase reduction\(^\text{174,178,199}\). The product of the gene involved in regulation of the *mod* Mo uptake operon, *modE*, is also involved in regulation of the alternative nitrogenases by Mo\(^\text{206,212}\). The product of the *nifL* gene is involved with the repression of nitrogenase by fixed nitrogen\(^\text{213–215}\); when it is nonfunctional, the cells excrete
ammonium\textsuperscript{213,215}. This NifL appears to be a sensor of redox changes in the cell\textsuperscript{214,216}, which corresponds with results showing reduced electron transport as the mechanism of nitrogenase shut-off by ammonium\textsuperscript{182,185}.

\textit{Nitrogen Fixation Under Oxygen}

As an obligate aerobe\textsuperscript{83,217,218}, \textit{A. vinelandii} has also been a useful model for the study of aerobic nitrogen fixation and how bacteria protect their oxygen-sensitive enzymes from oxidative damage: through enhanced respiration (“respiratory protection”), reversible conformational inactivation via the FeSII/Shethna protein, and other methods (such as superoxide dismutase). Oxygen and its reactive species are incompatible with nitrogenase enzymes, capable of reversibly inactivating them and preventing diazotrophic growth\textsuperscript{25,87,88,116,219}. Azotobacters employ a multifaceted approach to tolerate oxygen when fixing nitrogen; the most important facet is respiratory protection, an increase in respiration resulting in so-called “wasting” of oxygen\textsuperscript{7–9,11,32,84,87,143,220–224}. The cells redirect electron flux through the cytochrome \textit{bd} branch of their electron transport chain, starting from NADH dehydrogenase II\textsuperscript{225,226}, in which respiration is virtually decoupled from phosphorylation\textsuperscript{25,87,227–229}. This branch also produces much less superoxide\textsuperscript{225}. This redirection results in increased aldolase activity cycling carbon through the pentose phosphate cycle\textsuperscript{223}, increasing the production of carbon dioxide\textsuperscript{8,32,223} and the rate of respiration\textsuperscript{7,25,87,222}. Growth rate and yield are also affected, as cells devote carbon and energy to respiratory protection rather than to biosynthesis\textsuperscript{11,32,218,222,223}. Some question whether this increase in respiration truly creates an anoxic environment inside the cells, proposing instead
that it permits cells to maintain their oxygen-sensitive enzymes in more reduced, oxygen-tolerant states\textsuperscript{116,219,230}.

The other facets of azotobacters’ tolerance of oxygen include strategies to detoxify reactive oxygen species, reversible inactivation of nitrogenase to prevent irreversible inactivation, and possibly even detoxification of oxygen by components of the nitrogenase itself. \textit{A. vinelandii}’s iron-containing superoxide dismutase activity increases proportionally as oxygen exposure increases, especially in cells fixing nitrogen\textsuperscript{231}. The cells also increase catecholate siderophore production as oxygen increases, to chelate iron and prevent the generation of reactive oxygen species by Fenton chemistry\textsuperscript{53,232}. Some strains of \textit{Azotobacter} produce a layer of alginate polysaccharide in certain conditions, which may limit oxygen diffusion\textsuperscript{233}, though not all strains have this capability\textsuperscript{50}. If \textit{Azotobacter} experiences a sudden increase in oxygen exposure, or if it exhausts its carbon substrate, proliferation pauses (if substrate is available to increase respiration) or ceases\textsuperscript{9,25,87,118,219}. Rather than permitting oxygen to irreversibly inactivate its enzymes, \textit{Azotobacter} produces a small iron-sulfur protein called FeSII or Shethna\textsuperscript{119,234,235}, which reversibly inactivates nitrogenase by conformational change\textsuperscript{9,234}, to be reactivated when possible. Finally, it is possible that the dinitrogenase reductase component of nitrogenase itself contributes to oxygen tolerance if present in high enough proportions relative to oxygen, by reducing oxygen directly, an effect dubbed “autoprotection”\textsuperscript{236}.
Respiration

A. vinelandii has also been a model for the study of respiration in general, including aspects of the tricarboxylic acid cycle, the electron transport chain (ETC) and its branches, and the substrates involved in respiration. A. vinelandii’s ETC is interesting because it comprises multiple branches with different terminal oxidases in each\textsuperscript{225,228,229,237–239}, a setup that is essential for its adaptation to varying levels of oxygen. NADH and L-malate are the physiological substrates that donate electrons to begin the chain\textsuperscript{229,237,240–242}; NADPH may be used indirectly by passing electrons to NADH through a transhydrogenase\textsuperscript{243,244}. From there, all electrons go through ubiquinone (Q\textsubscript{8}) before passing to the different branches\textsuperscript{225,229,237,240,241}.

There are two main branches after ubiquinone. One is composed of \textit{b} and \textit{c} cytochromes and ends in a heme-copper oxidase; this branch has lower activity but higher affinity for oxygen, and has high phosphorylation activity, and is especially important in low-oxygen conditions\textsuperscript{25,225,239,245,246}. Some studies suggest that this branch may be further split into multiple branches with similar properties\textsuperscript{237,239,246,247}. The other main branch consists of cytochrome \textit{bd} oxidase, a quinol oxidase that is faster but has relatively lower affinity for oxygen, and has almost no ATP generation activity\textsuperscript{25,226,248–251}, though more recent studies have found that it is not completely uncoupled from oxidative phosphorylation\textsuperscript{252–254}. A. vinelandii up- or downregulates these branches depending on the redox conditions in its environment: the \textit{bd} branch is important in highly oxic conditions, and the other at lower oxygen levels\textsuperscript{225,226,246,248,251,255}. This organism’s setup represents an important strategy of adaptation to a complex redox habitat.
Polyhydroxyalkanoates

The production and biological role of polyhydroxyalkanoates is another area in which *A. vinelandii* has contributed as a model organism. This was discovered as early as 1937 by the observation of lipid-like granules in *Azotobacter* cells\(^{30,256}\), these granules were later identified as poly-β-hydroxybutyrate (PHB)\(^{257}\), and confirmed as present in *A. vinelandii* in 1958\(^{258}\). The granules have repeatedly been associated with cysts of *A. vinelandii*, staining like lipids using Sudan black B dye\(^{30,31}\). The amount of PHB in *A. vinelandii* cells peaks during the encystment process and then falls to a lower level as the cyst matures\(^{259–261}\). The granules of PHB stored in a cyst may prolong its survival in a semi-dormant state\(^{262–264}\), though the calculated rate of consumption of this polymer does not account for the more than decade-long survival of *Azotobacter* cysts in soil that has been observed\(^{265}\). Others have suggested that large amounts of PHB are not necessary for encystment\(^{266,267}\). Regardless, cells consume whatever is present when cysts germinate\(^{31}\). The location and movement of PHB granules in cells have also been studied: they originate at the periphery of the cell and migrate inward to the center as they form\(^{63,268,269}\). The presence of PHB around the cell’s membrane correlates with the cell’s transformation competence, though the latter is not dependent on the former\(^{63}\). Other studies relating to industrial applications of *A. vinelandii*’s polyhydroxyalkanoate production will be discussed later.
Agriculture

A. vinelandii’s roles in soils and agriculture have also been examined. Considering the problems that plague industrial nitrogen fertilizer use—depleted soil quality, environmental pollution, intensive energy use\textsuperscript{270-272}—biological nitrogen fixation in soils is of significant interest\textsuperscript{273}. Azotobacter is known to dwell in soils worldwide\textsuperscript{1,265,274-276}, though its abundance depends on the specific soil conditions\textsuperscript{275,276}. Cysts of A. vinelandii can survive at least 10 years in dry soil\textsuperscript{265}, and the vegetative cells can grow quite well on the phenolic acids and other nutrients in soil\textsuperscript{277}. A. vinelandii is also a plant growth-promoting rhizobacterium, living in the vicinity of plant roots and potentially increasing crop yields\textsuperscript{278,279}. The most obvious effect is A. vinelandii’s production of fixed nitrogen compounds. A closely related species, A. chroococcum, can fix about 40 milligrams of dinitrogen per gram of mannitol\textsuperscript{221}. Plants are capable, at least in certain conditions, of utilizing fixed nitrogen of bacterial origins\textsuperscript{280-282}. A. vinelandii itself can be modified not just to fix nitrogen to ammonium but to excrete excess into its surroundings, acting directly as a fertilizer\textsuperscript{19,213,215,283}, producing medium concentrations as high as 35 mM\textsuperscript{215}. However, A. vinelandii can also increase plant growth by other methods, such as increasing the nodulation activity of local rhizobia around legume roots\textsuperscript{284} and by producing plant growth-promoting hormones such as indole acetic acid (IAA)\textsuperscript{285-287}. Finally, A. vinelandii can enhance the nutritional value of crops, by increasing their vitamin content\textsuperscript{288} and protein quality and quantity\textsuperscript{287}.

Through these historical areas of research, Azotobacter vinelandii has contributed much valuable knowledge. Using modern systems and synthetic biology techniques and
approaches, study of this organism has great potential to contribute to multiple valuable areas of research in the future.

**Promising Areas for A. vinelandii Research**

*Azotobacter vinelandii* has also been studied as for various biotechnological potentials. These include the production of hydrogen from the nitrogenases, natural polymers of industrial interest, application of nitrogenase to alternative substrates, and more.

**Polyhydroxyalkanoates (PHAs)**

As mentioned, *A. vinelandii* produces hydrocarbon polymers in certain conditions. These polymers have certain advantages over petroleum-derived polymers: they are a form of renewable bioplastic\(^{289-295}\); they are biodegradable\(^{296-299}\); and they can make biocompatible medical devices\(^{300-306}\). However, compared to petroleum thermoplastics, bacterial PHAs have less versatility and are not always cost-competitive, so more research and development is needed to make them viable products.

*A. vinelandii* primarily produces poly-\(\beta\)-hydroxybutyrate (PHB). In a standard batch culture, most strains start producing PHB in late exponential phase, and then consume the polymer in stationary phase\(^{58,260,307}\). PHB can accumulate up to 70% of the cells’ dry weight\(^{26,57,58,308,309}\). As mentioned, there is a connection between PHB production and encystment; PHB may be a storage polymer for this process in strains that produce cysts\(^{260}\). In more controlled conditions, oxygen limitation induces its production in *Azotobacter* species\(^{26,32,307,308,310}\). When the limitation is lifted (or when the carbon substrate is exhausted),
cells consume the PHB$^{26,264,310}$. Formation of the polymer also acts as an electron sink when insufficient oxygen is present to accept extra reductants$^{310,311}$, or when NADH oxidation is impaired$^{58,312}$ (Figure 1.3). Iron limitation also may induce PHB production$^{307}$. The carbon substrate is also important: acetate and ethanol allow for little or no PHB production, compared with that produced from sugars or butanol$^{57,58,260}$.

Figure 1.3. Pathway from glucose to PHB in A. vinelandii, resulting in the loss of two molecules of CO$_2$ (one for each of the two acetyl-CoA residues that go into acetoacetyl-CoA) and conversion of one NADPH to NADP$^+$ (reducing acetoacetyl-CoA to D-$\beta$-hydroxybutyryl-CoA). ED = Entner-Doudoroff

As PHB production and degradation is an important metabolic pathway in *Azotobacter*, it is subject to regulation by a number of different factors. Most directly, the PhbR protein
upregulates the PHB-producing operon \((phbBAC)\) to which its gene \((phbR)\) is adjacent\(^{313,314}\). The stationary phase-associated sigma factor \((\sigma^{38}\) or RpoS) also plays a role in upregulating PHB production\(^{313,315}\), and the global two-component response regulator system, GacS/GacA, is also involved in this upregulation\(^{315–317}\). Other regulatory factors are involved as well. The \textit{arrF} gene encodes a small regulatory RNA, ArrF, expressed in conditions of iron limitation, which regulates the expression of iron-containing proteins such as superoxide dismutase and the FeSII/Shethna protein\(^{318}\). Knocking out the \textit{arrF} gene increases the production of PHB around 300-fold, even in well-aerated conditions; the ArrF sRNA represses genes related to PHB production\(^{307}\). Other small RNAs, regulated by the GacS/GacA system, include RsmA, which represses PHB production\(^{317}\). CydR (Fnr) is a transcriptional regulator protein that controls the low-affinity cytochrome \textit{bd} branch of the electron transport chain; knocking out the \textit{cydR} gene results in overproduction of PHB-producing proteins\(^{319}\). The rhodanese protein RhdA, which may regulate protection of \textit{A. vinelandii} from oxidants, represses PHB production\(^{320}\). PtsP, involved in regulating nitrogen fixation and respiratory protection in carbon-limited conditions, enhances the production of PHB\(^{321}\). It is likely that this summary is not an exhaustive list of all relevant factors, as some may remain to be discovered, but it is clear that polymer production is important to different aspects of \textit{A. vinelandii} metabolism.

For strains of \textit{A. vinelandii} that produce it, the production and regulation of alginate is also related to PHB metabolism in interesting ways. Alginate is a slimy polysaccharide of varying composition that many strains of \textit{A. vinelandii} produce, while other, “non-gummy” strains do not\(^{50}\). In some conditions, such as low aeration, strains that overproduce PHB also
produce more alginate than wild-type strains\textsuperscript{312}. The GacS/GacA system upregulates the production of both polymers\textsuperscript{316}, and in some conditions, mutations of PHB- or alginate-related genes both reduce the production of alginate\textsuperscript{322}. PHB and alginate are both involved in cyst structure and encystment and germination processes of \textit{A. vinelandii}\textsuperscript{260,267,323–326}, though strains unable to produce PHB can still encyst\textsuperscript{327}. However, PHB and alginate production may compete for the same carbon substrate, so in some conditions, knocking out the ability of \textit{A. vinelandii} to produce PHB or alginate increases yields of the other several-fold\textsuperscript{328,329}.

In order to solve the problems of cost-competitiveness and versatility of properties, a number of approaches have been attempted. For example, William Page and Olga Knosp developed a non-gummy strain of \textit{A. vinelandii}, strain UWD, that overproduces PHB constitutively; they transformed non-gummy strain UW (aka CA or OP\textsuperscript{50}) with genetic material from a gummy PHB-overproducing strain\textsuperscript{58}. Strain UWD consumes glucose more quickly than its parent (strain UW), and unlike UW in which low aeration induces PHB production, higher aeration correlates with higher PHB production in UWD (to a point)\textsuperscript{58,59}. NADH oxidation is impaired in strain UWD, so cells use PHB production as an electron sink to prevent reducing equivalents from building up\textsuperscript{58,312,330}. UWD produces yields of PHB almost 7-fold higher than its parent\textsuperscript{58}. Use and development of UWD and other over-producing strains (which could be engineered by modifying the regulatory factors mentioned previously\textsuperscript{331}) could improve the efficiency and competitiveness of bioplastic production.

William Page’s group and others have also studied the use of low-cost sources of carbon as substrates to replace costly, purified sugars and other compounds. For example, corn
syrup, cane or beet molasses, and malt extract allow PHB production comparable to glucose\textsuperscript{33,57,59,60,332}, and beet molasses actually stimulates the growth of \textit{A. vinelandii} UWD more than other substrates\textsuperscript{57}. One study estimated that, in 1992 at least, PHA production from UWD using beet molasses could be less expensive than polypropylene production\textsuperscript{33}. As a low-cost source of nitrogen, researchers tested fish peptone with strain UWD\textsuperscript{332–334}; added nitrogen greatly increases the yield of PHB\textsuperscript{58,59,333}, and adding fish peptone more than doubles the PHB yield from molasses\textsuperscript{333}. A two-stage process, with beet molasses in the first stage for biomass/protein production and then fish peptone and raw sugar in the second stage for PHB increases production even further\textsuperscript{334}. Peptone also increases pleomorphism in \textit{A. vinelandii} by weakening the structural integrity of the cell wall, resulting in weirdly shaped “fungoid” cells\textsuperscript{40,333}; this actually proves advantageous, as it facilitates extraction of the polymer from cells\textsuperscript{333}. Other low-cost substrates, such as swine wastewater, olive mill wastewater, and cheese whey, also show promise for production of PHAs\textsuperscript{335–338}.

To address the problem of versatility in the properties of PHA bioplastics, some have studied the possibility of producing copolymers in \textit{A. vinelandii}, instead of just poly-\textit{β}-hydroxybutyrate. When valerate (a 5-carbon fatty acid) is added to the growth medium, \textit{A. vinelandii} strain UWD produces poly(hydroxybutyrate-co-hydroxyvalerate) or P(HB-co-HV), a copolymer with two kinds of monomer and with improved material properties compared with PHB\textsuperscript{33,339}. PHA yield decreases in this condition though\textsuperscript{339}. The use of swine wastewater permits copolymer formation without added valerate, but supplementation improves yields\textsuperscript{335,336}. The development of other polymers and processing methods could increase the penetration of PHAs into the plastics market.
By creative use of what is known of PHA regulation, combined with improved design of production processes, *A. vinelandii* could act as a catalyst for competitive, efficient biodegradable biopolymer production.

**Alginates**

Alginates are slimy polysaccharide polymers mainly consisting of two sugar acids, (1-4)-\(\beta\)-D-mannuronic acid and \(\alpha\)-L-guluronic acid. Bacterial species in the family *Pseudomonadaceae*, including those in genera *Azotobacter*, *Pseudomonas*, and *Azomonas* produce this polymer\(^{340}\), as do brown algae. *A. vinelandii* converts hexoses such as glucose to trioses through the Entner-Doudoroff pathway\(^ {341-343}\), and then back to alginate via gluconeogenesis\(^ {343}\).

In *A. vinelandii*, depending on the condition, certain strains produce alginate and others do not. Wild-type strains produce it, but “non-gummy” strains, such as strain OP (aka UW or CA) can be isolated that generally do not\(^50\). Sequencing the genome of strain DJ, an offspring of strain CA, revealed a transposon insertion into the *algU* gene, which encodes an activator of alginate production, which explains its non-gummy phenotype\(^13\). However, PHB-overproducing strain UWD also produces much more alginate than its non-gummy parent (strain UW) when oxygen-limited in stationary phase\(^ {312}\), so environmental conditions and regulatory factors other than AlgU also influence production of the polymer.

Alginate in *A. vinelandii* is associated (and, in fact, required) for encystment\(^ {324,344-348}\); non-gummy strains do not form cysts\(^ {263,324,344,346,347}\), though they may form cyst-like structures\(^ {263}\). Alginate forms a protective part of the cyst coat, helping it resist
dessication. Its rate of production and molecular properties (molecular weight, ratio of mannuronic to guluronic acid, degree of acetylation) vary depending on the level of oxygen in the cells’ environment, possibly indicating a role for alginate in protecting the cells from excess oxygen. Non-gummy strains also show increased susceptibility to reactive oxygen species. The presence of alginate also affects the natural competence and transformation efficiency of *A. vinelandii*; gummy strains are less easily transformed, though protocol modifications can enhance their competence.

As with PHB, several independent factors regulate the production of alginate in *A. vinelandii* via transcriptional regulation of the *alg* operon starting with *algD*, which encodes GDP-mannose dehydrogenase, the rate-limiting enzyme in the process that converts GDP-mannose to GDP-mannuronic acid. The AlgU protein (aka RpoE), a putative sigma factor, upregulates the *alg* operon; when *algU* is knocked out, such as in strain OP (aka UW or CA), the cells usually do not produce alginate. Strain UWD, a derivative of strain OP, does produce alginate when oxygen-limited though; in this case, *algD* may be upregulated by RpoD (aka $\sigma^{70}$). The products of the *mucABCD* genes are negative regulators of the *alg* operon. Additionally, the GacS/GacA global two-component response regulator system and the stationary phase-associated sigma factor (RpoS) activate alginate production, the former via sRNAs that inhibit the repressor protein RsmA. Alginates are products of interest industrially; they can be used as stabilizing, gel-forming, or thickening agents in various processes in food, pharmaceutical, and cosmetic industries. Various approaches are being taken to develop processes in *A. vinelandii* to produce higher amounts and diverse forms of alginate for various potential uses.
example, regulating the oxygen transfer rate and oscillation of dissolved oxygen can affect the average molecular mass of the final alginate product, and to a lesser extent its yield and degree of acetylation (another important property). To regulate this, some have compared shake flask to bioreactor cultures and the power each consumes, as well as two-stage reactor processes with high aeration followed by low aeration. Optimizing the exposure of cells to CO₂ increases alginate yield and molecular mass, and adding 3-(N-morpholino)-propane-sulfonic acid (MOPS) to the culture medium can increase the degree of acetylation. Finally, genetic engineering can improve alginate production even further: knocking out PHB production increases alginate yield several-fold, though at the cost of some molecular mass, and reducing the amount of ubiquinone in cells also increases alginate yield. A double-knockout of phbR and muc26 produces less alginate but at a very high molecular mass, and knocking out alginate lyase may increase mass even further. Optimizing and combining these strategies with process engineering could result in a competitive system for production of alginates from A. vinelandii.

Synthesis Gas and Biological Fischer-Tropsch Chemistry

The developed world depends on the fossil fuel petroleum for a large number of high-value and high-demand products, including hydrocarbon fuels—such as diesel and gasoline—and a variety of important polymers, such as plastics and rubbers. However, oil reserves may meet global demands for only forty more years. Continued exploitation of fossil fuels will also exacerbate the problem of global climate change even sooner, leading to decreased biodiversity, sea-level rise, and expansion of pathogens’ geographic
Research into alternatives to petroleum will become increasingly necessary as we exhaust the planet’s supplies, and will be important for limiting the harmful effects of climate change and for improving the nation’s energy independence.

Reserves of the other fossil fuels—coal and natural gas—are larger, predicted to last about 60 and 130 years, respectively, but these products are not well-suited for transportation fuel, nor are they suitable substrates for polymer production. It is possible to convert coal, natural gas, and biomass into synthesis gas (syngas) by gasification of coal or biomass, or steam reforming of natural gas. Syngas consists mainly of carbon monoxide (CO) and hydrogen gas, and depending on the process, may contain significant quantities of nitrogen gas, methane, carbon dioxide, and minor impurities such as tars, dust, and nitric oxide. Steam reforming gives the cleanest product by the reaction \( \text{CH}_4 + \text{H}_2\text{O} \rightarrow \text{CO} + 3\text{H}_2 \), with the highest \( \text{H}_2:\text{CO} \) ratio (3:1) and with higher conversion efficiency, around 70\%.

Gasified coal and biomass have lower \( \text{H}_2:\text{CO} \) ratios, less than 1:1, with more impurities. \( \text{H}_2 \) is a high-value molecule itself, used in fuel cells and in other industrial reactions, but CO is an environmental pollutant in industrial and automotive exhaust gases; converting the CO from syngas and from these exhausts into more useful products would increase its value and clean up the gases before release.

Different processes exist for the conversion of CO to more valuable products. Several strains of acetogenic bacteria produce ethanol, acetate, and sometimes butanol from carbon and energy derived from CO and \( \text{H}_2 \) through the Wood-Ljungdahl pathway. These bacteria, mostly clostridia, are obligate anaerobes. They do not require \( \text{H}_2 \) for CO fixation, but it increases the process efficiency: for example, the reaction with CO alone is 6CO +
$3\text{H}_2\text{O} \rightarrow \text{C}_2\text{H}_2\text{OH} + 4\text{CO}_2$, but with the reductants derived from hydrogen instead of CO, the reaction is $2\text{CO} + 4\text{H}_2 \rightarrow \text{C}_2\text{H}_2\text{OH} + \text{H}_2\text{O}$; only one third the amount of CO is consumed per molecule of product formed$^{378}$.

A second process is the Fischer-Tropsch reaction, a chemical process using inorganic catalysts, especially the metals iron and cobalt$^{383}$. The catalyst consumes both CO and H$_2$ to produce hydrocarbons and oxygenated products of varied chain lengths, especially 1-alkenes, with some 2-methyl branching, as well as H$_2$O and CO$_2$$^{377,383,384}$. This reaction requires temperatures of several hundred degrees Celsius and often takes place at greater than one atmosphere of pressure$^{377,383,384}$; the range of effective temperatures is from 190 to 350ºC$^{375}$; in some cases 1 atm is sufficient$^{385,386}$.

Several years ago, Lee et al. discovered that A. vinelandii’s vanadium-containing nitrogenase converts CO into hydrocarbons such as ethylene ($\text{C}_2\text{H}_4$)$^{127}$. A. vinelandii’s molybdenum-containing nitrogenase catalyzes the same reactions but with much lower activity, about 0.1% that of the vanadium (V) version$^{126}$; its activity increases when D$_2$O replaces H$_2$O in the reaction mix$^{126}$ or when certain specific residues of the protein are substituted to enlarge the active site$^{128}$. The products of each enzyme in all conditions consist mainly of ethylene, with lower amounts of other 1- to 4-carbon hydrocarbons$^{126-128}$. The molybdenum (Mo) nitrogenase shows a tendency toward higher saturation and longer chains than the V nitrogenase, but in both, the longer the carbon chain, the more saturated it is$^{126}$.

The process appears to be possible because of the nitrogenase’s ability to bind two CO moieties simultaneously to the active site—its metal cofactor$^{103,387-390}$. Ian Dance used computer models to suggest that nitrogenase reduces the first bound CO to CH$_2$, releasing the
oxygen atom as \( \text{H}_2\text{O} \), then links the carbon of the second CO to the carbon of the first, then continues to reduce the product until releasing it as a hydrocarbon\(^{391}\). Alternatively, the nitrogenase could bind both CO moieties, link the carbons, and reduce the product to a hydrocarbon\(^{392}\). Three- or more-carbon chains may be formed by partial release of the chain followed by the linking of an additional CO. Considering that the main product by far is ethylene and that very little methane is formed\(^{126-128}\), it seems that the nitrogenase rapidly links two or more carbons before reducing and releasing them, but often releases the product before fully saturating it. The enzyme reduces its ligands by direct protonation (using \( \text{H}^+ \)) rather than by hydrogenation (using \( \text{H}_2 \)), so hydrogen gas is not consumed\(^{393}\).

Compared to the other known processes for converting CO into useful products, the nitrogenase catalysis has certain strengths and weaknesses. For example, acetogenic bacteria produce solvents while nitrogenase produces hydrocarbons, so the processes can be used simultaneously toward separate ends. However, acetogens consume hydrogen—a high-value component of syngas\(^{394}\)—as energy to increase the efficiency of solvent production\(^{378,381}\), while nitrogenase derives its energy and reductants from other sources (and actually produces \( \text{H}_2 \) as a separate product of its catalysis)\(^{393}\). The hydrogenase of the acetogens is responsible for taking up \( \text{H}_2 \) and generating reductants\(^{378,381}\), but possible impurities of biomass-derived syngas such as nitric oxide (NO) or even CO itself in high concentrations inhibit this hydrogenase, reducing the process efficiency\(^{378,381,382,395}\). Acetogens are obligate anaerobes, often requiring costly and difficult-to-maintain anoxic conditions\(^{382}\); their preferred medium also is complex and costly, containing multiple vitamins, minerals, and metals\(^{378,380,381}\), while nitrogenase-producing \textit{A. vinelandii} requires only inexpensive minimal medium with no
added nitrogen\textsuperscript{396,397}, which also reduces the risk of contamination by preventing growth of organisms with stricter growth requirements.

The Fischer-Tropsch (F-T) reaction produces compounds similar to those the nitrogenase produces\textsuperscript{377,383,384}. However, F-T produces large amounts of methane, a product very low in value\textsuperscript{394} because it is itself a precursor for the F-T reaction\textsuperscript{370,374}; nitrogenase produces little to no methane\textsuperscript{126–128}. F-T requires high temperatures\textsuperscript{375,377,383,384}, while nitrogenase catalyzes optimally at slightly above room temperature\textsuperscript{126–128}. F-T is unpredictable and produces a large variety of compounds\textsuperscript{377,383,384}; nitrogenase’s product range is narrow and specific\textsuperscript{126–128}. F-T consumes large amounts of syngas hydrogen to hydrogenate the carbon compounds it produces\textsuperscript{374,383,384,398,399} while nitrogenase does not\textsuperscript{393}, enabling it to make better use of syngas derived from coal and biomass, which has lower H\textsubscript{2}:CO ratios\textsuperscript{375–378}. F-T catalysts degrade, probably due to graphite buildup on the surface\textsuperscript{383}, while nitrogenase inactivation has not been observed\textsuperscript{128}. One notable advantage of F-T over nitrogenase, currently, is that its range of products includes longer-chain hydrocarbons useful for liquid fuels and waxes\textsuperscript{375,383}, but some groups are investigating the possibility of extending the nitrogenase’s product range using mutagenesis or different chemistries\textsuperscript{128,400}.

The main product of CO fixation by nitrogenase is ethylene (ethene, C\textsubscript{2}H\textsubscript{4}); other products are ethane (C\textsubscript{2}H\textsubscript{6}), propane (C\textsubscript{3}H\textsubscript{8}), and small amounts of propylene (propene, C\textsubscript{3}H\textsubscript{6}), methane, and four-carbon products\textsuperscript{126–128}. Longer-chain hydrocarbons are preferable as liquid fuels\textsuperscript{370}, but propane displays a number of potential advantages as a moderate-pressure automotive fuel\textsuperscript{401}. The main product, ethylene, is a high-value compound, produced industrially in larger quantities than any other organic compound\textsuperscript{402,403}, mostly from steam-
cracking of petroleum hydrocarbons. Its primary uses are incorporation into polymers such as polyethylene, polyvinyl chloride, and polystyrene, or conversion into antifreeze (ethylene glycol); secondary uses include anesthesia, forced ripening of fruit, and as welding gas.

Nitrogenase-catalyzed CO fixation shows potential as a process for converting CO into valuable products, and could offer an alternative to petroleum, especially for organic polymers. These aspects permit A. vinelandii’s nitrogenase enzymes to be considered as potentially useful biological Fischer-Tropsch catalysts. The catalytic abilities of A. vinelandii’s molybdenum- and vanadium-containing nitrogenases have been studied in vitro, but the CO-fixing ability of the iron-only nitrogenase has not been studied, nor has the in vivo CO-fixing ability of any of them. Considering that the three isoenzymes display a number of similarities but also notable differences, it is worth studying whether the iron (Fe) nitrogenase as a catalyst is better suited to CO fixation than either the V or Mo nitrogenases.

As mentioned previously, the three versions of nitrogenase share several features in common. The cofactors, though they each contain a different central metal, are all arranged in the same central trigonal prismatic arrangement of atoms, and the cofactor from each version can substitute for each of the others in all three apoproteins. All require the \textit{nifB} gene (involved in cofactor synthesis) and the \textit{nifV} gene (homocitrate synthesis) for activity. The protein residues important for cofactor-binding are conserved in all three.

When molybdenum is available in its environment, wild-type A. vinelandii preferentially uses the Mo nitrogenase for nitrogen fixation, especially around room temperature or
warmer; the two alternative isoenzymes are repressed\textsuperscript{27,44,47,206,416}. These two alternative nitrogenases share certain similarities with each other but not with the Mo version. Their predicted protein sequences are more homologous to each other than to the primary\textsuperscript{42,412–415}, and they both have a third subunit (δ) that the primary lacks\textsuperscript{42,46,407,414}, which is thought to play a role in cofactor incorporation\textsuperscript{417,418}. When fixing nitrogen gas, both V and Fe versions produce more molecules of H\textsubscript{2} per molecule of N\textsubscript{2} than the Mo version\textsuperscript{28,41,412,419,420}, and when reducing acetylene, both produce detectable amounts of ethane in addition to ethylene, while the primary produces only ethylene under most circumstances\textsuperscript{29,43,133,413,420}. In some situations, CO stimulates the alternatives to produce higher proportions of ethane from acetylene\textsuperscript{176,421}, this could be interpreted as evidence that the alternatives are reducing CO, not acetylene, into ethane. Finally, computer modeling predicts that the FeV and FeFe cofactors are both larger than the FeMo-cofactor\textsuperscript{422}.

In some ways, the Fe nitrogenase is as different from the V nitrogenase as the latter is different from the Mo nitrogenase. The Fe version produces more H\textsubscript{2} when reducing N\textsubscript{2} or C\textsubscript{2}H\textsubscript{2} than the others\textsuperscript{28,419,420}. Ethane makes up a larger proportion of the Fe nitrogenase’s acetylene reduction products than the V nitrogenase’s. The FeFe-cofactor is predicted to be larger than VFe-cofactor\textsuperscript{422}. So in almost every way that the V nitrogenase differs from the Mo nitrogenase, the Fe nitrogenase differs from the Mo version to a larger extent. It is possible that this pattern will hold for CO fixation, that the Fe nitrogenase will fix CO at a higher rate than the V nitrogenase, but this hypothesis has yet to be tested.
Other Minor Interests

Other avenues of research include using *Azotobacter vinelandii* to detoxify cyanide in wastewater from cassava processing\(^17\). *A. vinelandii*’s nitrogenases accept cyanide as a substrate, converting it into ammonia and methane\(^{132,135,136}\), or larger products such as methylamine in the case of the vanadium nitrogenase\(^{136}\). When incubated with wastewater containing 4 mM cyanide, *A. vinelandii* reduced the concentration by almost 70% within 66 hours. Additionally, the ammonia and methane produced from this may be harvested for fertilizer and energy. Other kinds of wastewater can be treated using *A. vinelandii* as well, to remove organic carbon without the addition of fixed nitrogen\(^{423,424}\) or to remove potentially toxic polyphenol compounds\(^{425}\).

Another potential use of *A. vinelandii* is providing fixed nitrogen for other industrially relevant microbes. Some algae and other microbes are capable of taking up bacterial siderophores and assimilating their metal contents\(^{426,427}\), but siderophores also contain fixed nitrogen, so those produced by *A. vinelandii* may act as a nitrogen source for other species. For example, algal species of industrial interest, such as *Neochloris oleoabundans* and *Scenedesmus* sp. BA032, are capable of growing with purified *A. vinelandii* siderophores as the sole nitrogen source, or even in co-culture with the bacteria themselves in the case of *Scenedesmus*\(^{18}\). The algae in this study preferred azotobactin rather than the catechol siderophores, but other algal species may have different capabilities\(^{18}\). Algae could also take up excreted ammonium directly from modified strains of *A. vinelandii*\(^{19,428}\).
Hydrogen

The interest most relevant to the following work is that of biohydrogen production in *Azotobacter vinelandii*. Interest in hydrogen as a clean alternative fuel is increasing because it is carbon-free and energy-dense\(^429-431\). Though most current research on biohydrogen focuses on fermentative or phototrophic processes\(^429,432,433\), aerobic, heterotrophic, nitrogenase-powered biohydrogen is increasingly being studied\(^20,434\). This approach has potential to serve as an alternative to the others, possibly as a value-adding strategy for biofertilization processes\(^18\).

As mentioned, all three nitrogenase complexes in *A. vinelandii* produce molecular hydrogen as a necessary byproduct of nitrogen fixation\(^28,46,420\); this evolution of hydrogen during nitrogen fixation occurs even under a high-pressure pure nitrogen atmosphere\(^123\). At its most efficient, the stoichiometric operation of the nitrogenase can be represented with the following equation:

\[
N_2 + 8H^+ + 8e^- \rightarrow 2NH_2 + H_2
\]

Various factors affect this equation, as discussed later. In structure and in function, “nitrogenase can be considered as a hydrogenase”\(^96\). *A. vinelandii* can be and has been studied as a model of this aerobic hydrogen production system.

This nitrogenase-generated hydrogen often is not directly observable, because *A. vinelandii* possesses a hydrogen-oxidizing membrane-bound uptake hydrogenase enzyme complex. Under an air atmosphere, this hydrogenase oxidizes nearly all the hydrogen produced by the molybdenum-containing nitrogenase\(^20,435,436\). See Figure 1.4 for an illustration of this system. Many microbial hydrogenases catalyze reversible reactions,
producing H$_2$ in most conditions$^{437}$, but $A.\ vinelandii$’s uptake hydrogenase and those of other nitrogen-fixing species are unidirectional under physiological conditions—they oxidize H$_2$ and do not reduce H$^+$—though a strong reducing agent and low pH can induce H$_2$ production$^{146,147}$. The enzyme contains a metal center made up of nickel and iron$^{147,438}$. The genes necessary to produce it are found in the $hox$ and $hyp$ operons; these are conserved in other organisms such as $Escherichia\ coli$, $Bradyrhizobium\ japonicum$, and $Rhodobacter\ capsulatus$$^{149,154,439}$. The genes $hoxKG$ encode structural subunits of the enzyme$^{151}$, while $hoxZMLOQRTV$ and $hypABFCDE$ encode necessary regulators, electron transport components, and proteins involved in processing and assembly of the hydrogenase$^{149,152,154,155,439}$. The regulation of these genes is not well-studied, but they are known to be upregulated in nitrogen-fixing conditions$^{205}$. 
In _Azotobacter_ species, nitrogenase produces hydrogen gas by reduction of protons; the amount depends on the nitrogenase version. A) When the uptake hydrogenase is active, it oxidizes most of the hydrogen to reduce oxygen to water. B) When hydrogenase is inactive, hydrogen escapes into the atmosphere. Source: World J Microbiol Biotechnol, “Protons and pleomorphs: aerobic hydrogen production in _Azotobacters_,” Vol 32, 2016, p3, J.D. Noar and J.M. Bruno-Bárcena, reprinted with permission of Springer

In addition to the membrane-bound uptake hydrogenase, _A. vinelandii_ and _Azotobacter chroococcum_ have genes homologous to those encoding a soluble hydrogenase enzyme\textsuperscript{13,15}. Its function has not been studied, nor has it been shown to be functional; studies with knockouts of the membrane-bound hydrogenase genes did not detect any hydrogen oxidation activity in nitrogen-fixing conditions\textsuperscript{154,155,440}.

Competition studies in _A. chroococcum_ have found that in carbon-limited, nitrogen-fixing conditions, the ability to oxidize hydrogen confers a growth advantage\textsuperscript{441,442}; one study in _A.
vinelandii failed to replicate this\textsuperscript{436} while another succeeded, in batch culture\textsuperscript{20}. Other studies have shown that energy from the oxidation of hydrogen—whether exogenous or produced by nitrogenase—can contribute to ATP production\textsuperscript{227,435} and nitrogenase activity\textsuperscript{443}. This energy also permits mixotrophic growth with mannose\textsuperscript{444,445}, and may play a role in the respiratory protection of nitrogenase from inactivation by oxygen\textsuperscript{435,443}. In other conditions, such as oxygen limitation, the situation may be reversed and the possession of an uptake hydrogenase is neutral\textsuperscript{154} or even disadvantageous\textsuperscript{441,442}. This may be due to the tighter coupling between ATP generation and the electron transport chain of hydrogenase compared to that dedicated to carbon substrates\textsuperscript{227}.

Hydrogen can serve as an electron donor in \textit{A. vinelandii}, increasing the total respiratory and oxygen consumption rates\textsuperscript{155,435,443}. However, environmental levels of oxygen can affect how much hydrogen \textit{A. vinelandii} produces. Hydrogen production correlates with nitrogenase activity, and above the point of oxygen limitation, increasing exposure to oxygen correlates with decreased nitrogenase activity\textsuperscript{88,435,446–448}. As mentioned, sudden increases in oxygen can shut down nitrogenase completely, if only transiently until the cell adapts\textsuperscript{446,447}. Similar results have been observed in more distantly related organisms: \textit{Klebsiella pneumoniae} represses nitrogenase synthesis when the oxygen concentration is too high\textsuperscript{449}; and higher oxygen concentration correlates with lower nitrogenase activity in \textit{Enterobacter aerogenes}, \textit{Xanthobacter flavus}, \textit{Azospirillum brasiliense}, and \textit{Burkholderia} and \textit{Rhizobium} species\textsuperscript{434,450–455}.

Various factors can affect yields of hydrogen from nitrogenase. \textit{A. vinelandii} prioritizes use of its molybdenum nitrogenase, which has consistently been shown to devote 25\% of its
electron flux—the flow of electrons through the complex that is constant even if no substrate is available—to hydrogen, producing one mole of H$_2$ per mole of N$_2$ fixed$^{28,123,456,457}$. The alternative nitrogenases display lower affinity for dinitrogen, so more of their electron flux goes to H$_2$: the vanadium-containing nitrogenase produces three to five moles of H$_2$ per mole N$_2$ fixed$^{420,457}$, and the iron-only nitrogenase ratio is between three and eleven to one$^{28,46,171}$. The total electron flux for each isoenzyme may not be equal, though; this and whether the above ratios can be compared are questions that are difficult to study, because the optimal conditions for purifying and testing each version have not been determined$^{28,419}$. However, it is apparent that a higher proportion of the electron flux through the alternative nitrogenases goes to hydrogen.

Excess oxygen (beyond the amount needed to relieve limitation) affects hydrogen yields in two ways: 1) by affecting the activity of nitrogenase and 2) by affecting the overall metabolic efficiency, though the line separating these effects is blurry. As mentioned, intensity of exposure to oxygen is inversely related to nitrogenase activity in Azotobacter, so relieving excess oxygen stress may increase the yields of nitrogenase-produced hydrogen. Excess oxygen also reduces overall growth efficiency in Azotobacter, especially in nitrogen-fixing cells: maintenance requirements increase one to two orders of magnitude when going from ammonia-grown to nitrogen-fixing conditions$^{458}$ or from oxygen-limited to carbon-limited growth$^{222}$. As oxygen increases, a higher proportion of substrate carbon consumed is oxidized to CO$_2$ instead of going to biomass$^{32,222,223,310}$; a higher proportion of electrons pass directly to oxygen without contributing to ATP production or nitrogen fixation$^{25,87,227}$. Studies of hydrogen production support these results. Linkerhögner and Oelze observed a
near six-fold decrease in yield of hydrogen upon increasing dissolved oxygen from 11.3 to 135 μM.

As *A. vinelandii* is naturally competent, the hydrogenase genes can be knocked out by homologous recombination. Also, spontaneous hydrogenase-negative mutants can be isolated by selection or mutagenesis. The hydrogenase can also be inactivated by chemical methods, thus affecting hydrogen yields. Acetylene in high concentrations, nitric oxide, carbon monoxide, or cyanide inhibit hydrogenase activity, and chelators such as ethylene diamine tetra-acetic acid (EDTA) or nitrilotriacetic acid (NTA) can prevent synthesis of active hydrogenase by binding necessary nickel and iron.

Some compounds can inhibit the nitrogenase, either by acting on it directly (as oxygen does) or by inducing the cell to repress its activity. Sources of fixed nitrogen, such as ammonia or nitrate, make nitrogen fixation redundant, so *Azobacter* shuts down its nitrogenases and ceases to produce hydrogen. However, mutants of the *nifL* gene can be generated that produce active nitrogenase even in the presence of fixed nitrogen.

Tungsten, not a functional component of any nitrogenase in nature, can be incorporated into the molybdenum nitrogenase as part of an iron-tungsten-cofactor (FeWco). Several studies showed a reduction in all activities—N$_2$ reduction, H$_2$ evolution, and ATP hydrolysis—with tungsten incorporation, but one in *Rhodobacter capsulatus* observed active but diminished hydrogen production under an argon atmosphere with tungsten-containing nitrogenase. Adding tungsten in nitrogen-fixing conditions inhibits the growth of wild-type strains of *Azotobacter*, in concentrations as low as 1 μM, though
increasing the concentration of molybdenum may relieve this inhibition competitively\textsuperscript{168}. However, in strains lacking repression of the alternative nitrogenases, added tungsten forces reliance on these alternatives\textsuperscript{4,167}; this can stimulate hydrogen yields, because a greater proportion of electron flux goes to hydrogen in these isoenzymes\textsuperscript{20}.

Other compounds can modify hydrogen yields by affecting the direction of nitrogenase electron flux. Removing external nitrogenase substrates, such as by replacing dinitrogen in the atmosphere with argon, induces nitrogenase to direct its entire flux solely to proton reduction, transiently producing from 1.2 to 11-fold more hydrogen (depending on the isoenzyme, substrate, and composition of the atmosphere) than when a substrate is present\textsuperscript{28,46,420}. Carbon monoxide inhibits nitrogen fixation or acetylene reduction by nitrogenase, but does not reduce the total electron flux—the balance goes toward proton reduction, and thus toward production of hydrogen\textsuperscript{131,144,465,466}.

It is possible to make point mutations to the nitrogenases that modify its substrate reduction activities and electron flux. For example, modifying NifD position 191 from glutamine to lysine redirects most electron flux toward hydrogen instead of acetylene or nitrogen; however, the total electron flux is reduced at least two-fold, so less product can be formed overall\textsuperscript{133,465,467}. However, this reduced activity might reflect reduced amount of active protein, rather than reduced activity in the active protein itself\textsuperscript{467}.

It is important to note that many of these studies were performed \textit{in vitro} on purified enzymes. Differences in the purification procedure or conditions may make comparisons between enzymes and between studies—as well as extrapolations to \textit{in vivo} applications—less valid.
To summarize, hydrogen production in *Azotobacter* and other diazotrophs is chiefly an interplay between two enzyme systems: hydrogenase and nitrogenase. Uptake hydrogenase oxidizes most of the hydrogen produced by the nitrogenase, but there are ways to remove it from the picture. Many factors affect the hydrogen-evolving activity of nitrogenase: the specific version of the enzyme and the central cofactor incorporated into that version, the level of oxygen to which the organism is exposed, the availability of fixed nitrogen, and other attributes of the atmosphere and the medium. Determining the optimal conditions for *Azotobacter* to produce hydrogen could lead to novel biotechnological processes, especially considering the genus’s other desirable features\textsuperscript{16,18,19}, and may at least serve as a model of a novel hydrogen production system.

**Conclusions**

Study of *Azotobacter vinelandii* and its cousin species has contributed much to bacteriological knowledge for the past century, and they continue to represent important species with great potential for research and industrial interest. The following work focuses on evaluating and optimizing the hydrogen-producing potential of a particular mutant strain of *A. vinelandii*, strain CA6.
CHAPTER 2. Genomic Comparison of *Azotobacter vinelandii* Mutant Strain CA6 with its Parent Strain CA

**Introduction**

*Azotobacter vinelandii* is a Gram-negative, soil-dwelling, obligately aerobic diazotroph, discovered more than a century ago\(^1\). *A. vinelandii* strain CA (or UW or OP) (ATCC 13705) was isolated as a non-gummy pigment-producing wild-type strain that was easier to study than “gummy” polysaccharide-producing strains\(^50\). A consortium of researchers has published the complete genome sequence of a variant of strain CA, a high-frequency transforming strain called DJ (ATCC BAA-1303)\(^13\).

*A. vinelandii* strain CA6 is a spontaneous mutant strain derived from strain CA. Tungstate prevents nitrogen fixation and growth of the strain CA, but it does not inhibit CA6\(^4\). Strain CA6 also displays impaired molybdate uptake\(^167\), and has been found to produce large quantities of hydrogen gas when fixing nitrogen. For this reason, we sequenced the genomes of both CA6 and its parent, CA, to identify the genetic bases of these mutant phenotypes.

**Materials and Methods**

*Strains and Growth Conditions*

*Azotobacter vinelandii* strains were grown in modified Burk broth\(^396\), sometimes containing 28 mM ammonium acetate, and incubated at 29-30°C shaking at 200 rpm.
DNA Sequencing

One hundred milliliter cultures of CA and CA6 were divided into 5 aliquots each and pelleted. The pellets were stored at -80ºC. Bacterial pellets were sent to the University of North Carolina-Chapel Hill’s Microbiome Core Facility (Chapel Hill, NC) for pyrosequencing on a Roche 454 GS FLX Titanium+ Sequencer (Roche, Switzerland). DNA from one pellet from each strain was extracted using a PowerSoil DNA Isolation Kit (MO-BIO, Carlsbad, CA) according to manufacturer’s instructions, and then prepared for sequencing according to manufacturer’s instructions. Each strain was sequenced using half of a 454 plate. A second pellet from each strain was used for re-sequencing at the same facility on the Ion Torrent PGM platform (Life Technologies).

Polymerase chain reactions (PCR) followed by Sanger dye-terminator sequencing were performed to confirm variations or provide extra coverage for low- or no-coverage regions of the genomes. Primers were designed using Clone Manager 9 (Sci-Ed Software, Cary, NC) and ordered from Integrated DNA Technologies (Coralville, IA). See Table 2.1 for primers used. PCR was performed with DNA as templates extracted with a DNeasy Tissue Kit (Qiagen, Valencia, CA) according to manufacturer’s instructions. Alternatively, cultures of bacteria were placed at -20ºC overnight, then thawed and used as template directly. Remaining reagents were ExTaq (TaKaRa, Japan), HiFi HotStart (KAPA, Woburn, MA), or Taq (Qiagen) PCR kits. Concentrations and thermal cycler protocols were according to manufacturer’s instructions. PCR products were run on 1% agarose gels in TAE buffer with ethidium bromide. When necessary due to multiple bands, the band at the predicted size was cut out of gel and extracted using QIAquick Gel Extraction Kit (Qiagen). PCR products were
purified using QIAquick PCR Purification Kit (Qiagen). Products were sent to Eton Bioscience (Research Triangle Park, NC) or ACGT, Inc. (Wheeling, IL) for dye-terminator sequencing. Sequences were analyzed using FinchTV (Geospiza, Seattle, WA) and Geneious v6 (Biomatters Ltd., New Zealand).

In Silico Sequence Analysis

454 reads were assembled using Geneious v6 software with the genome sequence of A. vinelandii strain DJ as a reference. Single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) were catalogued initially using Geneious with a 45% variant frequency cutoff. Ion Torrent data and Sanger sequencing reads were used primarily to confirm or reject variations detected in 454 data. De novo assembly was performed using Newbler (Roche, Switzerland) to confirm that there were no large rearrangements. Final annotated versions of the genomes of CA and CA6 were submitted to GenBank; accession numbers are CP005094 (CA) and CP005095 (CA6). FastQC was used to determine sequence quality.

In order to determine optimal settings for Geneious to detect variations between strains, strains were compared at different depth and variant frequency cutoffs, and the number of false variations detected at each setting was calculated. These calculations were done on 454 reads alone, Ion Torrent reads alone, and reads from both technologies combined, using Medium sensitivity assembly unless otherwise specified.
Table 2.1. Primers used

<table>
<thead>
<tr>
<th>Location of 5’ End in DJ</th>
<th>Direction</th>
<th>Primer Sequence</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1186882</td>
<td>Forward</td>
<td>5’-GTGCTGGCGAGATCGACGCTGTTC-3’</td>
<td>Re-sequence a low-coverage region</td>
</tr>
<tr>
<td>1187926</td>
<td>Reverse</td>
<td>5’-GCAGCGCGGAATGGGCTGACTA-3’</td>
<td>Re-sequence a low-coverage region</td>
</tr>
<tr>
<td>2866462</td>
<td>Forward</td>
<td>5’-GTAGGTATCGCTGGCCTTCCAA-3’</td>
<td>Confirm insertion in DJ</td>
</tr>
<tr>
<td>2866751</td>
<td>Reverse</td>
<td>5’-TCGCCCTGTATTAGCCAGCAAAGAC-3’</td>
<td>Confirm insertion in CA or CA6</td>
</tr>
<tr>
<td>3182590</td>
<td>Forward</td>
<td>5’-TGCGAACTACGCCAGTGACATT</td>
<td>Confirm insertion in DJ</td>
</tr>
<tr>
<td>3182937</td>
<td>Reverse</td>
<td>5’-CGCCCTGTATTAGCCAGCAAAGAC-3’</td>
<td>Confirm insertion in CA or CA6</td>
</tr>
<tr>
<td>2523005</td>
<td>Forward</td>
<td>5’-TGCGAACTACGCCAGTGACATT</td>
<td>Confirm insertion in DJ</td>
</tr>
<tr>
<td>2523341</td>
<td>Reverse</td>
<td>5’-CGCCCTGTATTAGCCAGCAAAGAC-3’</td>
<td>Confirm insertion in CA or CA6</td>
</tr>
<tr>
<td>53117895</td>
<td>Forward</td>
<td>5’-GCGGAGGCGGGCAATGGCGTGACTTA-3’</td>
<td>Re-sequence a low-coverage region</td>
</tr>
<tr>
<td>5318787</td>
<td>Reverse</td>
<td>5’-GCGGAGGCGGGCAATGGCGTGACTTA-3’</td>
<td>Re-sequence a low-coverage region</td>
</tr>
<tr>
<td>496737</td>
<td>Forward</td>
<td>5’-GCGCGGTGCTTCCTCCCGTCGTGGTGAC-3’</td>
<td>Confirm 42kbp deletion in CA6</td>
</tr>
<tr>
<td>497350</td>
<td>Reverse</td>
<td>5’-GCGCGGTGCTTCCTCCCGTCGTGGTGAC-3’</td>
<td>Confirm 42kbp deletion in CA6</td>
</tr>
<tr>
<td>2866243</td>
<td>Forward</td>
<td>5’-CCGTCCGGCGAGACGGCTGTCGTA-3’</td>
<td>Confirm deletion in CA6</td>
</tr>
<tr>
<td>2867001</td>
<td>Reverse</td>
<td>5’-CCGTCCGGCGAGACGGCTGTCGTA-3’</td>
<td>Confirm deletion in CA6</td>
</tr>
<tr>
<td>3726483</td>
<td>Forward</td>
<td>5’-GCTGGCTGAGCCGTCGTGGTGAC-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>3727301</td>
<td>Reverse</td>
<td>5’-GCTGGCTGAGCCGTCGTGGTGAC-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5102727</td>
<td>Forward</td>
<td>5’-GGCGAGGCGGGCAATGGCGTGACTTA-3’</td>
<td>Confirm 42kbp deletion in CA6</td>
</tr>
<tr>
<td>5103533</td>
<td>Reverse</td>
<td>5’-GGCGAGGCGGGCAATGGCGTGACTTA-3’</td>
<td>Confirm 42kbp deletion in CA6</td>
</tr>
<tr>
<td>1107618</td>
<td>Forward</td>
<td>5’-GACTGGCAGCGCGGCCGAACACAG-3’</td>
<td>Confirm deletion in CA</td>
</tr>
<tr>
<td>1107955</td>
<td>Reverse</td>
<td>5’-GACTGGCAGCGCGGCCGAACACAG-3’</td>
<td>Confirm deletion in CA</td>
</tr>
<tr>
<td>3868445</td>
<td>Forward</td>
<td>5’-CGAAGGGGTGTCGGCCTGGTGAC-3’</td>
<td>Confirm deletion in CA</td>
</tr>
<tr>
<td>3868718</td>
<td>Reverse</td>
<td>5’-CGAAGGGGTGTCGGCCTGGTGAC-3’</td>
<td>Confirm deletion in CA</td>
</tr>
<tr>
<td>5030255</td>
<td>Forward</td>
<td>5’-CCGTCTCCTCCCGCGAGCCGGAGACACGGATCA-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5034504</td>
<td>Reverse</td>
<td>5’-CCGTCTCCTCCCGCGAGCCGGAGACACGGATCA-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>2014002</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>2014929</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4415777</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4416690</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4424706</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4425701</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5124867</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5125741</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5125712</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>5125544</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4686350</td>
<td>Forward</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
<tr>
<td>4687563</td>
<td>Reverse</td>
<td>5’-ACCGAGATTTCCTCGGCGGCTGAG-3’</td>
<td>Confirm SNP in CA6</td>
</tr>
</tbody>
</table>

Note: The table lists primers used for various genomic locations and their corresponding purposes.
Figure 2.1. FastQC analysis of *A. vinelandii* CA (top) and CA6 (bottom) 454 reads
Results

Sequence Quality Analyses

In order to determine the genetic basis for the mutant phenotypes (slower growth, impaired molybdenum uptake, hydrogen production, and tungsten tolerance), we sequenced the genomes of *A. vinelandii* strains CA and CA6 and compared them with each other and with strain DJ.

Together, 454 pyrosequencing and Ion Torrent sequencing provided 63-65x coverage for each genome, mapping to a depth between 0 and 240 reads per position. More sequencing details can be found in Table 2.2. Gaps and areas of low depth were re-sequenced using PCR and dye-terminator sequencing. See Figure 2.1 for FastQC analysis of per-base quality for CA and CA6. We also calculated the percentages of each genome covered to a range of minimum depths (i.e. what proportion of positions is covered by at least x reads?) for each sequencing technology and the combination of the two (Figure 2.2). 454 provided much greater coverage: nearly 100% of each genome was covered by at least 20 reads (Fig 2.2a), whereas not even 90% of each genome was covered by at least 2 Ion Torrent reads (Fig 2.2b).
Figure 2.2. Proportion of genome covered to at least a minimum depth (reads per position) by A) 454, B) Ion Torrent, and C) both technologies combined

Table 2.2. Sequencing Statistics

<table>
<thead>
<tr>
<th></th>
<th>454 CA</th>
<th>454 CA6</th>
<th>Ion Torrent CA</th>
<th>Ion Torrent CA6</th>
<th>Combined CA</th>
<th>Combined CA6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Reads</td>
<td>607,789</td>
<td>568,103</td>
<td>463,121</td>
<td>452,254</td>
<td>1,070,910</td>
<td>1,020,357</td>
</tr>
<tr>
<td>Used Reads (%)</td>
<td>606,450 (99.8%)</td>
<td>566,774 (99.8%)</td>
<td>424,539 (91.7%)</td>
<td>424,875 (93.9%)</td>
<td>1,030,989 (96.3%)</td>
<td>991,649 (97.2%)</td>
</tr>
<tr>
<td>Mean Read Length</td>
<td>412.2 ±126.3</td>
<td>429.5 ±113.3</td>
<td>214.3 ±37.3</td>
<td>214.2 ±41</td>
<td>330.7 ±139.4</td>
<td>337.2 ±139.3</td>
</tr>
<tr>
<td>Read Length Range</td>
<td>40-597</td>
<td>40-595</td>
<td>14-394</td>
<td>14-393</td>
<td>14-597</td>
<td>14-595</td>
</tr>
<tr>
<td>Mean Depth</td>
<td>46.8±13</td>
<td>45.6±14.4</td>
<td>18.9±14.1</td>
<td>19.2±15.6</td>
<td>66.1±22.7</td>
<td>65.4±25.8</td>
</tr>
<tr>
<td>Depth Range</td>
<td>0-160</td>
<td>0-139</td>
<td>0-106</td>
<td>0-108</td>
<td>0-190</td>
<td>0-228</td>
</tr>
<tr>
<td>Coverage</td>
<td>47</td>
<td>46</td>
<td>17</td>
<td>17</td>
<td>64</td>
<td>63</td>
</tr>
</tbody>
</table>

Four regions in DJ compared to CA and CA6 were large (300-800 bp) stretches of no coverage. We considered the possibility that these were insertions in DJ, but the surrounding reads did not align to each side of the insertions as would be expected for real polymorphisms; the depth simply declined until it reached zero and then increased again on
the other side of the insertion. PCR and Sanger sequencing confirmed that DJ, CA, and CA6 have the same sequences at these four sites, but there seem to be secondary structures such as hairpins that interfere with conventional sequencing at these loci.

Small Variation Detection

We calculated the number of true variations detected and the number of false variations discovered with various cutoff settings, to determine the optimal settings in Geneious. A higher depth cutoff setting was advantageous because this directed the software to search only in high-depth areas, and the greater the depth, the less likely a false variation will appear in the majority of reads. For example, it is more likely a spurious SNP will occur in the majority of 10 reads than in the majority of 50 reads. Similarly for higher frequency cutoffs; a spurious variation is more likely to occur in more than 10% of reads than in more than 50% of reads. Figure 2.3a shows total false variations detected in CA with different settings. For Figure 2.3b, we extrapolated from the data in 2.3a to calculate how many false variations would be expected throughout the whole genome, if the proportion found in the region of adequate depth remained constant. For example, if a genome of 1000 bases is searched with settings of at least 50 depth and at least 81% frequency, and two false variations are found in the 5% of the genome covered by at least 50 reads, the extrapolated result would be 40 false variations (20 times 2).
For 454 sequencing, when coverage was greater than 100 reads per base, a frequency cutoff of 41% or more was sufficient to eliminate all but one false variation in CA’s sequence; the cutoff could be set as low as 31% in CA6’s. However, less than 0.5% of each genome had coverage of at least 100 (see Figure 2.2). At the same frequency cutoffs, when nearly 100% of each genome was covered (cutoff of 20 reads), Geneious detected 550 false variations in CA and 211 in CA6. At a frequency cutoff of 91% and depth cutoff of 20 reads, Geneious found 9 false variations in CA and 6 in CA6.

For Ion Torrent sequencing, the coverage was too low to be very useful in detecting variations, but it helped to confirm or reject variations detected by 454. Combining Ion Torrent and 454 reads in one assembly helped increase genome coverage and reduce the
number of false variations found, especially in mid- and high-range frequency cutoff settings. With this combination, nearly 100% of the genome was covered by at least 30 reads, and with this depth cutoff, only 334 false variations in CA had a frequency of at least 41%, but 409 in CA6 at 31%.

For finding true variations, 454 was again more effective than Ion Torrent: 454 detected 100% of true variations when searching areas covered by at least 20 reads and 95% with at least 50 reads, while Ion Torrent detected only 77% with coverage of only 5 reads or more. This correlates well with the overall percentages of genome covered to these depths. Ninety to 95% of true small variations had variant frequencies above 90% in 454 reads; this was only true of 53 to 61% of true variants discovered by Ion Torrent. Even when reads from both technologies were combined, this was true of only 68 to 77% of true variations, but all could be detected with coverage to a depth of at least 30 reads.

All of the above analyses were done with Geneious assembly algorithms set to medium sensitivity. This setting seemed to capture all of the high-quality reads in a reasonable time. However, use of medium-high sensitivity allowed for the assembling of additional, previously unused reads: about 1300 more from the 454 dataset and about 18500 more from Ion Torrent. The latter set was sufficient to increase each genome’s coverage by about one. However, extra reads from 454 either increased or caused no change to false variation detection rates at most detection settings. Extra reads from Ion Torrent increased the false variations detected by up to almost 100,000 at lower frequency and depth cutoff settings, but decreased them by up to almost 1000 at mid-range settings (higher settings still showed decrease, but to a lesser extent).
We examined more closely the 454 reads included in a medium-high sensitivity assembly but rejected in medium sensitivity assembly. Many aligned very well along half their length in one section of the reference genome, while the other half aligned very well in another random locus, sometimes very distant from the first. These reads seemed to be sequencing errors in which parts of different reads had been combined—chimeras. This phenomenon was absent in the same category of Ion Torrent reads, which seemed simply to be of low quality.

**Patterns of False Variations**

Of the false variations detected using cutoffs of 20 depth and 30% frequency when assembling 454 reads, more than 97% were homopolymer confusion: when sequencing a string of one type of nucleotide, 454 often mistakes a longer string as a shorter string, creating the appearance of an insertion or deletion (depending on whether the reference genome is an ancestor or descendant).

**True Variations**

Mapping of reads from *A. vinelandii* CA to the genome sequence of *A. vinelandii* strain DJ, a high-frequency transforming variant of CA whose genome had previously been published\textsuperscript{13}, followed by some follow-up confirmation, revealed a number of small variations, both single nucleotide polymorphisms (SNPs) and insertions/deletions (indels) between the two strains (Table 2.3), but overall the genomes displayed greater than 99.9%
pairwise identity. Most notable, though not a small variation, was a 1-kb transposon present in CA but not DJ, inserted between two open reading frames.

Table 2.3. Variations Discovered in Genome Comparison. Position column contains nucleotide position in DJ. * - stop codon. Variant Frequency refers to percentage of sequencing reads showing each variation.

<table>
<thead>
<tr>
<th>Locus Tag</th>
<th>Gene Function</th>
<th>Position</th>
<th>CDS Position</th>
<th>Base/AA Change</th>
<th>Variant Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>AvCA6_14900</td>
<td>Glycerol-3-phosphate acyltransferase</td>
<td>1470474</td>
<td>854/987</td>
<td>G:A/G:D</td>
<td>100%</td>
</tr>
<tr>
<td>AvCA6_27910</td>
<td>Flagellar biosynthesis</td>
<td>2871708</td>
<td>1568/2085</td>
<td>G:A/E:*</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>many</td>
<td>3868591</td>
<td></td>
<td>C:T</td>
<td>56%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5102868-5144962</td>
<td>-42095bp/many genes</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>AvCA_11230</td>
<td>Cation transporter</td>
<td>1072498</td>
<td>765/1071</td>
<td>+GACCAT/+DH</td>
<td>96%</td>
</tr>
<tr>
<td>AvCA_12720</td>
<td>Peptidase, modulator of DNA gyrase</td>
<td>1240550</td>
<td>896/1443</td>
<td>C:T/Q:*</td>
<td>100%</td>
</tr>
<tr>
<td>AvCA_18210</td>
<td>hypothetical protein</td>
<td>1805893</td>
<td>375/762</td>
<td>-GC/Frameshift</td>
<td>94%</td>
</tr>
<tr>
<td>AvCA_24220</td>
<td>Flagellar motor switch</td>
<td>2417604</td>
<td>144/996</td>
<td>G:A/E:*</td>
<td>98%</td>
</tr>
<tr>
<td>AvCA_36670</td>
<td>TolQ proton channel</td>
<td>3728269</td>
<td>481/696</td>
<td>G:A</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4686485</td>
<td></td>
<td>+1053bp</td>
<td>100%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locus Tag</th>
<th>⨳</th>
<th>Position</th>
<th>CDS Position</th>
<th>Base/AA Change</th>
<th>Variant Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avin_02860</td>
<td>Hypothetical</td>
<td>270373</td>
<td>163/369</td>
<td>C:T/P:L</td>
<td>98%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>278677</td>
<td></td>
<td>C:T</td>
<td>100%</td>
</tr>
<tr>
<td>Avin_06870</td>
<td>Hybrid histidine kinase</td>
<td>650766</td>
<td>2715/2802</td>
<td>G:C/H:D</td>
<td>100%</td>
</tr>
<tr>
<td>Avin_06870</td>
<td>Hybrid histidine kinase</td>
<td>652277</td>
<td>807/2802</td>
<td>-AGGGCG/-AL</td>
<td>100%</td>
</tr>
<tr>
<td>Avin_07570</td>
<td>Membrane bound protease regulator</td>
<td>716487</td>
<td>564/870</td>
<td>-CGCGAG/-RE</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1504215</td>
<td></td>
<td>-A</td>
<td>81%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2199316</td>
<td></td>
<td>+G</td>
<td>98%</td>
</tr>
<tr>
<td>Avin_22710</td>
<td>Heavy metal translocating P-type ATPase</td>
<td>2265546</td>
<td>357/2160</td>
<td>G:A/A:T</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3182849</td>
<td></td>
<td>+TGGCGGATGGCGGA</td>
<td>100%</td>
</tr>
</tbody>
</table>
The comparison of reads from strain CA6 with DJ’s genome revealed a conspicuous large deletion in the CA6 genome, encompassing 42 whole or partial open reading frames. Because of this deletion, identity between CA6 and DJ is only 99.2%. Table 2.4 contains a complete list of deleted open reading frames, their direction, and any open reading frames elsewhere in the genome that may complement their function; Figure 2.4 shows the locations and orientations of the genes contained.
Table 2.4. Genes and putative genes deleted in CA6, and putative homologs elsewhere in the genome. * - partial deletion.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
<th>Direction</th>
<th>Possible Homologs</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>fhuC</em></td>
<td>Iron(III)-siderophore ABC transport ATP-binding</td>
<td>Forward</td>
<td><em>hmuV, Avin_26450, 47160</em></td>
</tr>
<tr>
<td><em>fhuD</em></td>
<td>Iron(III)-siderophore-binding periplasmic protein</td>
<td>Forward</td>
<td>Avin_30870</td>
</tr>
<tr>
<td>fhuB</td>
<td>Iron-hydroxamate transporter permease subunit</td>
<td>Forward</td>
<td>Avin_26440, 47150, 42600</td>
</tr>
<tr>
<td></td>
<td>Aspartate racemase</td>
<td>Reverse</td>
<td>Avin_25560, 25580, 25570, 25650</td>
</tr>
<tr>
<td></td>
<td>MbtH-like protein; antibiotic biosynthesis?</td>
<td>Reverse</td>
<td>Avin_25630</td>
</tr>
<tr>
<td></td>
<td>DUF81 family membrane protein; sulfite exporter?</td>
<td>Forward</td>
<td>Avin_01050</td>
</tr>
<tr>
<td></td>
<td>HupE/UreJ protein; hydrogenase/urease accessory protein, nickel binding?</td>
<td>Forward</td>
<td>cooJ</td>
</tr>
<tr>
<td></td>
<td>hypothetical membrane protein</td>
<td>Forward</td>
<td>~10 elsewhere in genome</td>
</tr>
<tr>
<td></td>
<td>hypothetical membrane protein</td>
<td>Forward</td>
<td>~10 elsewhere in genome</td>
</tr>
<tr>
<td></td>
<td>hypothetical protein (mammalian gene homolog?)</td>
<td>Reverse</td>
<td>Avin_52080</td>
</tr>
<tr>
<td>hypE</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hypD</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hypC</td>
<td>Hydrogenase assembly chaperone (HypC/HupF)</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hypF</td>
<td>Hydrogenase maturation carbamoyltransferase</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hypB</td>
<td>Hydrogenase nickel incorporation</td>
<td>Reverse</td>
<td>Avin_00470</td>
</tr>
<tr>
<td>hypA</td>
<td>Hydrogenase nickel incorporation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxV</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxT</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxR</td>
<td>Rubredoxin-type Fe(Cys)4 protein</td>
<td>Reverse</td>
<td>Avin_48490</td>
</tr>
<tr>
<td>hoxQ</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxO</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxL</td>
<td>Hydrogenase assembly chaperone (HypC/HupF family)</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxM</td>
<td>Hydrogenase expression/formation</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td>hoxZ</td>
<td>Mn/Fe-hydrogenase, b-type cytochrome subunit</td>
<td>Reverse</td>
<td>Avin_49100, 10120, 03240</td>
</tr>
<tr>
<td>hoxG</td>
<td>Membrane bound nickel-dependent hydrogenase, large subunit</td>
<td>Reverse</td>
<td>Avin_04360-04380, hoxYHW</td>
</tr>
<tr>
<td><strong>modC1</strong></td>
<td>Mo transporter, inner membrane ATP-binding component</td>
<td>Reverse</td>
<td><strong>modC2</strong> (Avin_01280), 01580</td>
</tr>
<tr>
<td><strong>modB1</strong></td>
<td>Mo transporter membrane protein</td>
<td>Reverse</td>
<td><strong>modB2</strong> (Avin_01290), 50720</td>
</tr>
<tr>
<td><strong>modA1</strong></td>
<td>molybdenum transporter, periplasmic molybdate-binding</td>
<td>Reverse</td>
<td><strong>modA2</strong></td>
</tr>
<tr>
<td><strong>modE</strong></td>
<td>Mo regulation, Mo processing homeostasis</td>
<td>Reverse</td>
<td>Avin_33430, 22590</td>
</tr>
<tr>
<td><strong>modG</strong></td>
<td>Mo processing, homeostasis</td>
<td>Forward</td>
<td>Avin_22590, 33430</td>
</tr>
<tr>
<td></td>
<td>ABC transporter ATP binding component (Mo?)</td>
<td>Forward</td>
<td>mtlK, cysA, Avin_21300, 23060, 14020</td>
</tr>
<tr>
<td></td>
<td>ABC type tungstate transporter, permease component</td>
<td>Reverse</td>
<td></td>
</tr>
<tr>
<td><strong>modA3</strong></td>
<td>molybdate ABC transporter, periplasmic molybdate-binding</td>
<td>Reverse</td>
<td>Avin_26680</td>
</tr>
</tbody>
</table>

---

**Table 2.4 Continued**

| **hoxK** | Uptake hydrogenase small subunit (Precursor) | Reverse |
| Transposase | Forward | Several elsewhere |
| General substrate transporter | Forward | Avin_18040, 24940 |
| Hypothetical protein | Forward |
| Periplasmic hybrid histidine protein kinase, two-component | Forward | gacS, Avin_34990, 29290, retS, 25170 |
| Response regulator with metal dependent phosphohydrolase activity (two-component) | Forward | Avin_25160, 29300 |
| **modC1** | Mo transporter, inner membrane ATP-binding component | Reverse | **modC2** (Avin_01280), 01580 |
| **modB1** | Mo transporter membrane protein | Reverse | **modB2** (Avin_01290), 50720 |
| **modA1** | molybdenum transporter, periplasmic molybdate-binding | Reverse | **modA2** |
| **modE** | Mo regulation, Mo processing homeostasis | Reverse | Avin_33430, 22590 |
| **modG** | Mo processing, homeostasis | Forward | Avin_22590, 33430 |
| | ABC transporter ATP binding component (Mo?) | Forward | mtlK, cysA, Avin_21300, 23060, 14020 |
| | ABC type tungstate transporter, permease component | Reverse |
| **modA3** | molybdate ABC transporter, periplasmic molybdate-binding | Reverse | Avin_26680 |
| | ImpA-related (T4SS?) | Reverse | Avin_26680 |
| | Hypothetical protein (ImpL-related; T4SS?)* | Reverse | Avin_26680 |
Figure 2.4. Schematic representation of a section of CA’s genome, moving from the top left corner to the bottom right, with annotations marking the CA6 42kbp deletion (white) and all genes included (black). Arrows indicate strand orientation of each gene. Brackets indicate genes involved in uptake hydrogenase. Image produced using Geneious (http://www.geneious.com)

This large deletion in CA6 eliminates several operons from the strain, including the *hox* operon encoding the membrane-bound uptake hydrogenase, with its accessory protein-encoding genes in the *hyp* operon\cite{149,154,439}; most of the iron transporter-encoding *flu* operon; and the well-characterized molybdenum transporter-encoding *mod1* operon and putative molybdenum-transporter-encoding *mod3* operon (Figure 2.4)\cite{212,468,469}. 

56
Figure 2.5. A) Variations between the genomic sequences of *A. vinelandii* CA and CA6, mapped onto the genome of CA6. Black annotations are mutations present in CA6 (compared to strains CA and DJ); grey are mutations in CA (compared to CA6 and DJ). B) Mutations in the genomic sequence of *A. vinelandii* DJ, compared to parent CA and sibling CA6, mapped onto the genome of CA. Numbers on the outside show base numbers of genome locations. Image produced using Geneious (http://www.geneious.com)
Other Small True Variations

Figure 2.5a shows an atlas with the locations of all variations between CA and CA6, Table 2.3 lists other minor variations between strains, and Figure 2.5b shows an atlas of variations between DJ and the other strains.

SNPs, insertions, and deletions were identified from 454 reads in silico using Geneious software and confirmed with Sanger or Ion Torrent sequencing when necessary. Because strain DJ was derived from CA\textsuperscript{13}, we concluded that bases present in CA but not DJ were deletions in DJ, and that bases present in DJ but not CA were insertions in DJ. For CA6, no such conclusions could be made by comparing CA6 to DJ in silico, because both CA6 and DJ are derived from CA. We concluded that bases present in CA6 and CA but not DJ were deletions in DJ; bases present in CA6 but not CA or DJ were insertions in CA6; bases present in DJ but not CA or CA6 were insertions in DJ; and bases present in DJ and CA but not CA6 were deletions in CA6.

When comparing each strain to the other two, CA6 was found to contain 3 SNPs, including a nonsense mutation in the flagellar biosynthesis gene \textit{flhA}; CA contains 3 SNPs, including a nonsense mutation in the flagellar motor switch \textit{fliG}; 3 small insertions (2-8 bases in regions of repeats), and one new transposon insertion (1053bp); and DJ contained 6 SNPs, 6 small insertions (1-14 bases), and 4 small deletions (1-6 bases). Notably, two of DJ’s small insertions introduce a frameshift into their open reading frame, which in CA encode a TonB-dependent vitamin B12 receptor and a proline iminopeptidase. All these variations are listed in Table 2.3, along with their effects on proteins.
Discussion

The genes lost in *A. vinelandii* CA6 with its large deletion have the potential to explain many, if not all, of the mutant’s unique phenotypes (tungsten tolerance, impaired growth, impaired molybdenum uptake, and hydrogen gas production). Knocking out the *mod1* operon, especially the regulatory gene *modE*, results in lower molybdate accumulation at high concentrations, as previously observed in CA6\(^{167,468}\). ModE is also involved in molybdate-induced repression of the alternative nitrogenases\(^{206,212}\); its loss in CA6 may allow the expression of alternative nitrogenases that are not poisoned by tungstate. Previous studies have suggested that the *hox* operon contained within the deleted region in CA6 is responsible for the oxidation of hydrogen gas\(^ {153,154,439}\); without this operon, hydrogen that *A. vinelandii*’s nitrogenases produce as a byproduct evolves into the headspace, reducing its energy-efficiency and impairing its growth.

Most of *A. vinelandii* CA6’s few other variations do not seem significant in effect; however, both CA6 and its parent CA seem to have nonsense mutations in important motility-related genes. This may explain why we have not observed motility in either strain in any condition, despite claims about *A. vinelandii*’s alleged motility\(^1\).

There is a possibility that the strain we used as the wild-type, CA, has mutated and is no longer identical to the parent strain of CA6, HS2, and DJ. In particular, the transposon insertion present in the genome of our CA but not in the other strains seems to disrupt the promoter of a cyclic-AMP regulatory protein that is thought to regulate several catabolic operons; this disruption may affect CA’s substrate utilization phenotype. However, physiological studies in later chapters do not support this.
Allegedly, *A. vinelandii* DJ is a high-frequency transforming variant of strain CA\textsuperscript{13}, so we hoped to elucidate the genetic basis for this phenotype by comparing the two strains’ genomes. Unfortunately, none of the variations in DJ is conspicuous in its effect, but they may offer clues for further study. Several proteins may contain single amino acid changes or small insertions or deletions (one or two amino acids), but most notable are the frameshift mutations—a single-base deletion in the gene encoding a TonB-dependent vitamin B12 receptor, and a two-base insertion in the proline iminopeptidase gene. Future studies should focus on these variations.

Considering the coverage of each genome achieved with 454 pyrosequencing, this technology seems preferable to Ion Torrent, though this depends somewhat on the procedure used for each. Coverage aside, 454 reads seem to be much higher quality than Ion Torrent data (see proportion of usable reads for each in Table 2.2). However, even taking only high-quality reads into account, Ion Torrent is several times less expensive than 454.

In order to recommend optimal settings for variation detection based on our data, we must establish priorities: first, to discover all true variations; second, to minimize false variations. For our 454 data, a depth of 23 was the highest minimum depth that would discover all true variations (because overall genome coverage declined quickly beyond this depth, Figure 2.2a). Variant frequencies for CA’s true variations were all above 81%, and the same was true for all but one of CA6’s (the last had a frequency above 51%; it may be anomalous); so a minimum variant frequency of 51% is most conservative, but a minimum of 81% suffices for more than 97% of true variations.
After establishing the most conservative settings, it is usually necessary to suffer a number of false variations that will also be detected. As Figure 2.3b shows, the rate of false variations is minimized at the point of minimum depth where the genome coverage percentage begins to drop off, at the highest possible variant frequency cutoff. Extrapolating the rates of false variations loses accuracy as the sample size decreases; for example, one false variation found in CA at a depth of at least 70 reads (which covered almost 5% of the genome) was extrapolated as 101 false variations if the coverage were extended to 100% of the genome, while the same single variation was extrapolated to 644 variations when found at a depth of at least 100 reads (which covered less than 0.2% of the genome). To more accurately measure the optimum depth for minimizing false variations, it would be necessary to gather more data. However, given the data we have, using 454 reads, we can recommend a minimum depth of 20 and minimum variant frequency of 70%, which would limit CA’s false variations to 86 and CA6’s to 11, a fairly manageable amount.

That the majority of variations were homopolymer confusions is an artifact of the sequencing technologies. Both 454 and Ion Torrent measure the lengths of homopolymers by proportionally comparing the measured quantities to that of a single base (quantities of light in 454 and of pH change in Ion Torrent); there is some variation in quantities, and this is exacerbated with long homopolymer strings, such that an arbitrary decision must sometimes be made when giving an exact length. This may be avoided by checking with a different kind of sequencing technology, such as Illumina, which is not prone to this type of artifact.

Another type of artifact we encountered combined homopolymer confusion with the presence of copies of long regions in various places in the genome—in this case, transposase
coding sequences. Since these sequences, repeated in multiple genomic loci, are identical or nearly so, and are too long for any single read to span them, often it is impossible to know which particular locus matches a given read. If there is a homopolymer present, many such reads may show a false variation. Therefore, assembly software may assign all those reads with the correct homopolymer length to one copy locus, and all those with the incorrect length to another locus, spuriously giving the appearance of a high-frequency true variant. We observed this in CA’s genome, which seemed to show that more than 98% of reads in a region covered by a depth of more than 150 reads displayed a homopolymer length variation. It is important to be aware of this phenomenon.

The reads from *A. vinelandii* CA seemed to be lower quality than from CA6, in both 454 and Ion Torrent. The number of false variations in CA was higher in 454 reads, and the number of usable reads from Ion Torrent was lower. We have no reasonable speculations about why this might be.

*A. vinelandii* CA6’s mutant phenotypes, especially tungsten tolerance, are intriguing, and it has been used before when studying *A. vinelandii*’s alternative nitrogenases. We uncovered the likely genetic source of CA6’s phenotypes—deletion of a large multi-gene region—but convincingly linking gene losses with specific phenotypes will require mutational physiological analyses. However, the specific genes lost in the deletion event suggest other phenotypes that have not yet been observed in CA6, such as iron limitation due to the loss of the iron transport-related *fhu* operon. Future studies should examine these gene losses in more depth, and determine whether *A. vinelandii* CA6 has potential as a hydrogen-producing strain.
Despite the large deletion in strain CA6, different strains of *A. vinelandii* retain high inter-strain genomic similarity over time, remaining more than 99% identity. The mechanism that generated CA6’s deletion is a curiosity, but the strains are, for the most part, genetically stable.
CHAPTER 3. Physiological Evaluation of *Azotobacter vinelandii* CA6 Mutant

Phenotypes, Especially Hydrogen Production

**Introduction**

The genomic comparison in Chapter 2 revealed a large deletion in *Azotobacter vinelandii* strain CA6; many genes present in the parent strain CA are absent in CA6. This deletion explains the known mutant phenotypes of CA6: diazotrophic tolerance of tungsten\(^4,167\), caused by loss of a protein (ModE) that represses alternative, tungsten-resistant nitrogenase systems\(^206,212\); impaired uptake of molybdenum\(^167,468\), caused by loss of a molybdate uptake system operon\(^212,468,469\); and reduced growth compared to the parent strain when fixing nitrogen\(^20,167\). The genes lost in CA6’s deletion suggest other phenotypes—including iron utilization impairment from loss of the iron uptake-related *fhu* operon, and diazotrophic hydrogen evolution due to loss of the membrane-bound uptake hydrogenase *hox* operon—for which some evidence has been produced\(^20\).

As discussed in Chapter 1, hydrogen is an energy-dense compound with great industrial value, and while most approaches to biohydrogen employ photosynthetic or anaerobic fermentative microbes, an aerobic heterotrophic hydrogen producer could add value to some processes. To evaluate the potential of *A. vinelandii* to fill this role, it was worthwhile to investigate the following questions: does iron utilization impairment affect strain CA6’s metabolism and yield parameters? Though CA6’s nitrogen fixation is insensitive to tungsten, is it sensitive to regulation in the presence of fixed nitrogen? What yields can be expected from CA6, and how to these compare to the parent strain or other strains missing fewer
genes? Can forcing the use of alternative nitrogenases increase the yield of hydrogen from CA6? We used chemostat culture techniques to address these questions.

Methods

Strains and Media

Figure 3.1 displays the strains utilized in this study and their parent-offspring relationships. CA and DJ are available from the American Type Culture Collection under numbers 13705 and BAA-1303, respectively; R.L. Robson generously shared a culture of HS2, and CA6 is part of NCSU’s Department of Microbiology Culture Collection\textsuperscript{4,50,154}.

**Azotobacter vinelandii** strains

![Diagram of Azotobacter vinelandii strains]

- **CA**
  - Parent
  - Bush/Wilson, 1959

- **Spontaneous mutation**

- **Marker inserted in hoxK**

- **CA6**
  - Tungsten-tolerant
  - Bishop et al. 1980

- **DJ**
  - Reference genome
  - Setubal et al. 2009
  - High-frequency transforming variant

- **HS2**
  - *hoxK* knockout
  - Menon et al. 1992

Figure 3.1. Strains of *A. vinelandii* used in this study. Arrows indicate direction of parent-offspring relationship.
Azotobacter vinelandii strains were grown in modified nitrogen-free Burk broth unless otherwise noted, it contained, per liter, 0.2g MgSO$_4$•7H$_2$O, 90mg CaCl$_2$•2H$_2$O, 0.25mg NaMoO$_4$•2H$_2$O, 0.2g KH$_2$PO$_4$, 0.8g K$_2$HPO$_4$, 2.5mg FeSO$_4$•7H$_2$O, 2.5mg ferric citrate, and 20g sucrose. Sucrose and iron solutions were filter-sterilized with 0.2 μm filters; other solutions were autoclaved and the various components were combined as needed. Inocula were cultured by incubating at 30°C, shaking at 200 rpm. For solid medium, agar was added to 1.5% w/v.

**Bioreactor Setup**

Chemostat experiments were performed in a Biostat Bplus 2L bioreactor (Sartorius, Germany) equipped with sensors for pH (Mettler-Toledo, US), temperature, and dissolved oxygen (Mettler-Toledo or Hamilton, Nevada, US), mass flow control (airflow), and agitation.

Chemostat experiments were performed in a culture volume of 700mL; temperature was held at 30°C and pH was allowed to vary freely. Compressed air was pumped in through a 0.2 μm filter at a rate of 0.315±0.05 liters per minute (lpm); the system was open to gas exchange with the atmosphere. Exhaust gases (O$_2$, N$_2$, CO$_2$, H$_2$, and Ar) were monitored redundantly in real-time using in-line O$_2$/CO$_2$ EasyLine Continuous Gas Analyzers, Model EL3020 (ABB, Germany) and a Pfeiffer OmniStar quadrupole mass spectrometer. Culture turbidity was monitored in real-time using an in-house device that continuously measures light transmittance through a section of glass tubing through which culture is pumped.
(United States Patent 6975403). Arbitrary Unit (AU) measurements from this device can be converted to OD_{600} units with the following formula: OD_{600} \text{ units} = 0.0177 \times \text{AU} – 1.9583.

Media were prepared in 20-L glass carboys as three separate solutions—5 L salts in a 10-L bottle, autoclaved 45-60 minutes; 12 L phosphates in the final 20-L carboy, autoclaved 60-75 minutes; and 3 L remaining components in 3-L bottle, filter-sterilized—then combined by transferring the salts and sucrose solutions into the final carboy using filtered air pressure through silicone tubing. The final volume was placed on a scale and the weight loss was monitored and recorded continuously throughout the experiments.

For each experiment, the reactor, set to the appropriate temperature (30°C) and dissolved oxygen concentration (30%), was inoculated by injection of pre-grown cells in a flask batch culture. When the culture reached exponential phase, the inflow of fresh medium and removal of excess culture volume was initiated. Steady state was achieved before starting any sampling.

System parameters were defined as follows: dilution rate is the value obtained after reaching equal feed and harvest flows that allows for obtaining a constant volume of culture in the reactor (D = F/V). Yield is defined as the ratio of product generated to limiting substrate (sucrose) consumed. Productivity is the value of product generated per limiting substrate consumed per unit of time.

*Transient Responses to Additions in Chemostat*

Medium used was nitrogen-free Burk as described above, containing 50 mg/L FeSO_{4} \cdot 7H_{2}O, autoclaved along with the 5-L salt solution. When the three component
solutions were combined, a precipitate formed at the bottom of the carboy where it was inaccessible to the feed tube; this served to enhance any observed effects of iron limitation. This medium was used for all three strains tested. Sucrose was the sole source of carbon and energy. Samples of each batch of medium were saved at -80°C for later analysis.

The dilution rate for these chemostat experiments was set to 0.038 h⁻¹. Once steady state was achieved, a pulse of one of the following was injected into the reactor: A) 7 mL containing 1.4 g sucrose; B) sucrose as in A, plus 0.7 mL containing 1.75 mg each FeSO₄·7H₂O and ferric citrate; C) 7 mL containing 1.54 g ammonium acetate; or a combination of C with A or B. Adding citrate is an established method for maintaining iron in a soluble form over the length of a chemostat experiment⁴⁷¹,⁴⁷². No pulses were made until after at least four retention times had passed and steady state was reestablished. Retention time is the time required to replace one reactor-volume of culture with fresh medium, 1/D; for example, at D = 0.038 h⁻¹, 700 mL pass through the system in approximately 26 h—one retention period.

*Determinaton of Yields by Stepwise Changes*

Medium used was nitrogen-free Burk as described above, except it contained 5 mg FeSO₄·7H₂O and 0.5 g sodium citrate per liter, filter-sterilized along with the sucrose in the 3-L bottle. The citrate prevented the other components from precipitating, so everything remained in solution, as described previously⁴⁷¹; citrate does not serve as a carbon source in these conditions⁴⁷³. For tungstate-containing chemostats with CA6, sodium tungstate was added to 1 mM final concentration. To increase sucrose concentrations, 3 L complete
nitrogen-free Burk medium was assembled with enough extra sucrose to increase the concentration in the medium remaining in the 20-L carboy to the desired values. This 3 L was filter-sterilized and added to the 20-L carboy using filtered air pressure. A sample of each batch of medium at each concentration was saved at -80°C for later analysis.

The dilution rate for these experiments was set to 0.066 h\(^{-1}\). Once steady state was achieved, at least three samples were taken and analyzed before the sucrose concentration was changed. At least one retention period was allowed to pass between each sample; at D = 0.066 h\(^{-1}\), 700 mL pass through the system in approximately 15 h.

**Sample Analysis**

Each sample was spun down for 5 minutes at 10,000 rpm in an Eppendorf Centrifuge 5417R (Eppendorf, US), then the supernatant and pellet were saved in separate tubes at -80°C for later analysis. Optical density (OD\(_{600nm}\)) was determined using an Ultrospec 1100 pro spectrophotometer. Dry weight was obtained by filtering a portion of sample using vacuum suction through a 0.2 μm filter of known mass (mixed cellulose esters, EMD Millipore, Germany); the filter was then dried at 60-70°C for 1-7 days and re-weighed until weight was constant to determine the dry weight of biomass per sample volume.

The initial and residual concentrations of sucrose were determined from samples of initial media and supernatants of reactor samples analyzed by high-performance liquid chromatograph (Shimadzu, Japan) run under isocratic conditions at 65°C. The mobile phase was water at 0.5 mL/min, and the column was Supelcogel™ Ca (300 mm x 7.8 mm, Supelco™ Analytical, Bellefonte, PA, USA). The column was coupled to a refractive index
detector. Absolute values were obtained by measuring known quantities of sucrose to generate a standard curve.

**Results**

*Transient Response Experiments Confirm Iron Uptake Limitation in CA6*

The phenotypes in *A. vinelandii* CA6 of slower growth, tolerance to tungstate, and impaired molybdate uptake had previously been observed\(^{167}\), and batch cultures had also demonstrated iron uptake impairment and hydrogen evolution capabilities\(^{20}\). Therefore, to determine whether the iron-impaired phenotype could influence hydrogen production, chemostat experiments were performed comparing *A. vinelandii* strains CA, CA6, and HS2 (a strain in which the hydrogenase gene *hoxK* was knocked out\(^{154}\)). The strain HS2 was included in this study since it retains the *fhu* operon and other metal transporters missing in CA6, so HS2 cultures should not encounter limiting levels of iron as readily as CA6.

After achieving steady-state conditions with each strain in a low-iron, carbon-limiting, nitrogen-free Burk medium, pulses of sucrose and iron or sucrose alone were tested. This strategy allows for identification of limiting growth factors and has been utilized for medium optimization\(^{474,475}\). If the tested element acts as growth-limiting factor, the products of the culture should increase transiently and then recover to steady-state levels after the washout. As expected, the addition of iron with sucrose did not affect the amount of biomass, hydrogen, or CO\(_2\) generated by CA or HS2 from the additional carbon, whereas the increase in these products approximately doubled when iron was added with sucrose to steady-state CA6 cultures (Figure 3.2).
In addition to regulation of the alternative nitrogenases in the presence of heterometals, wild-type *A. vinelandii* represses nitrogen fixation in the presence of fixed nitrogen\textsuperscript{6,22,396}. *A. vinelandii* CA6 lacks Mo-induced repression of alternative nitrogenases; to discover if it also lacks regulation of nitrogenase activity in the presence of fixed nitrogen, steady-state cultures of CA6 and HS2 were tested by pulsing nitrogen in the form of ammonium acetate, by itself or in combination with sucrose or sucrose and iron. All pulses produced an immediate
cessation of hydrogen production in both strains, and hydrogen production gradually resumed as the ammonium washed out or was consumed (Figure 3.3).

Figure 3.3. Transient response to nitrogen pulses applied to aerobic, carbon-limited, steady-state chemostat cultures (D = 0.033 h⁻¹) of A. vinelandii CA6 growing at 30°C in nitrogen-free Burk medium. Left-most vertical dashed line indicates the time of pulse; to the left of it the culture is in steady state, and to the right a transient phase occurs as the cells respond to the addition of the limiting component (1.54 g ammonium acetate, 1.4 g sucrose, and 1.75 mg FeSO₄·7H₂O). To the right of the right-most vertical dashed line, steady state resumes after pulse contents are consumed or washed out.
Figure 3.4. Generation rates of biomass (dry weight), CO$_2$, and hydrogen in response to increases in sucrose feeding rates of *A. vinelandii* strains growing in chemostat culture (D=0.066 h$^{-1}$) at 30°C in iron-sufficient, nitrogen-free Burk medium. Triangles (Δ) give values for strain CA; circles (○) for HS2, empty squares (□) for CA6, and filled squares (■) for CA6 plus tungsten. Error bars (present for all points) indicate standard deviations. Yield values given in Table 3.1 were obtained for each strain by calculating the slope of the best-fit line for these points for each strain. A) Dry weight values, in carbon-moles (standardized by moles of carbon). B) Values for hydrogen, given as hydrogen produced, i.e. total hydrogen minus environmental hydrogen. C) Values for CO$_2$, given as CO$_2$ produced, i.e. total CO$_2$ minus environmental CO$_2$. 

![Graph A: Dry weight production rate (C·mmol/h)]

![Graph B: H$_2$ production rate (mmol/h)]

![Graph C: CO$_2$ production rate (mmol/h)]
Determination of Yields by Stepwise Increases in Concentration of the Growth-Limiting Substrate Using Chemostats

We conducted a series of tightly controlled aerobic, carbon-limiting chemostat culture experiments in iron-sufficient and nitrogen-free medium. By varying the concentration of sucrose in the medium, we established stepwise steady states and compared the biomass (dry weight), CO$_2$, and hydrogen production capabilities of A. vinelandii strains CA, CA6, and HS2. Steady-state values of dry weight, hydrogen, and carbon dioxide over sucrose consumed were plotted and the yields were calculated from the slope of the resulting best-fit lines. Figure 3.4 shows graphs of product over substrate for biomass, hydrogen, and CO$_2$, and Table 3.1 provides calculated yields and productivity values for the three tested strains. Notably, under the conditions tested, CA6 and HS2 are nearly identical in yields for all three products, and each produces almost 100-fold more hydrogen per gram of substrate than CA, and slightly less (not statistically significant) biomass and CO$_2$ (Table 3.1).

Table 3.1. Kinetic and yield parameters from stepwise chemostat experiments, calculated from slopes of lines in Figure 3.4

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dilution rate (h$^{-1}$)</td>
<td>CA</td>
</tr>
<tr>
<td>Biomass yield (mg dry weight g$^{-1}$ substrate)</td>
<td>23.58</td>
</tr>
<tr>
<td>Hydrogen yield (μg H$_2$ g$^{-1}$ substrate)</td>
<td>1.93</td>
</tr>
<tr>
<td>CO$_2$ yield (g CO$_2$ g$^{-1}$ substrate)</td>
<td>1.43</td>
</tr>
<tr>
<td>H$_2$ productivity (μg H$_2$ g$^{-1}$ substrate h$^{-1}$)</td>
<td>0.13</td>
</tr>
<tr>
<td>CO$_2$ productivity (g CO$_2$ g$^{-1}$ substrate h$^{-1}$)</td>
<td>0.09</td>
</tr>
<tr>
<td>Biomass productivity (mg dry weight g$^{-1}$ substrate h$^{-1}$)</td>
<td>1.56</td>
</tr>
</tbody>
</table>
Enhancement of CA6 Hydrogen Production by Addition of Tungstate

When molybdate is present in nitrogen-free medium, CA preferentially uses its primary nitrogenase, but CA6 produces both the primary and iron-only nitrogenases in these conditions, resulting in its tungstate tolerance phenotype\textsuperscript{167}. Considering that the iron-only nitrogenase is less efficient—a higher proportion of its electron flux goes to reducing protons instead of N\textsubscript{2}—we expected that tungstate would poison CA6’s primary nitrogenase, forcing it to rely on the alternative nitrogenase and increasing hydrogen yield as a result. We repeated the stepwise chemostat experiment with CA6 with 1 mM tungstate and no added molybdate in the medium (Figure 3.4 and Table 3.1). This experiment is possible only with CA6, not the other strains, because of tungsten’s nitrogenase-poisoning effect\textsuperscript{4,168}. Compared to CA6 in the absence of tungstate, tungsten-grown CA6’s hydrogen yield increased about 4.5-fold, while biomass yield decreased about 30\% and CO\textsubscript{2} about 10\%.

Discussion

The primary genetic change that generates phenotypic differences between \textit{A. vinelandii} CA and its offspring CA6 is a 42-kbp deletion, which removes a number of genes, including genes encoding metal transporters and the uptake hydrogenase, and allows for co-expression of the primary and alternative nitrogenases. \textit{A. vinelandii} CA6’s known mutant phenotypes are impaired growth in some conditions, impaired molybdate uptake, and tolerance to tungstate via derepression of alternative nitrogenases\textsuperscript{4,167}. The discovery of the large deletion in CA6 unveiled two other changes that led to the observation of new phenotypes. First, an impairment of iron transport due to the deletion of the \textit{fhu} operon was observed\textsuperscript{20}: we
explored this phenotype functionally and determined the conditions required to avoid iron limitation when cultivating CA6. The second phenotype was hydrogen evolution by CA6: it corresponded to the lack of hydrogenase genes (especially *hox*, as confirmed by *hoxK*-knockout strain HS2) within the deleted region in CA6. The products of these operons are responsible for the oxidation of hydrogen gas, and without these enzymes, the hydrogen that *A. vinelandii*’s nitrogenases generate as a byproduct escapes the cell. Additions of ammonium and iron also confirmed the nitrogenases’ role in *Azotobacter* hydrogen generation since the presence of ammonium inhibited both nitrogenase activity and hydrogen production.

Eliminating our mutant’s iron limitation using an iron-sufficient, nitrogen-free medium allowed for a comparison of the hydrogen-producing capabilities of CA, CA6, and HS2. The findings reveal that both mutants (CA6 and HS2) produce similar amounts of hydrogen in these conditions—the deficiency of a single gene (*hoxK*) was sufficient to replicate the hydrogen-production and impaired-growth phenotypes of CA6 in the conditions tested (Table 3.1). Surprisingly, despite CA6 co-expressing both the primary and less efficient iron-only nitrogenases, we observed no difference in hydrogen produced from CA6 compared with strain HS2, which produces only the primary nitrogenase in the presence of molybdate. Additionally, the unique capability of CA6 to grow in the presence of tungstate allowed us to test the role of the iron nitrogenase in hydrogen evolution and demonstrate that increased hydrogen yields can be obtained using CA6. In these conditions, when of the three nitrogenase isoenzymes only the iron-only nitrogenase could fix nitrogen, the cells produced larger amounts of hydrogen, generating about 4.5-fold more hydrogen than with tungstate.
absent; this is in line with previous observations in several nitrogen-fixing species\textsuperscript{28,46,171,419,476–478}. Another interesting possibility is that tungsten replaces Mo in the central dinitrogenase cofactor, rendering the enzyme capable only of proton reduction to dihydrogen\textsuperscript{464}. In such a scenario, the product of \textit{nif} genes would not fix nitrogen, only produce hydrogen.

\textit{A. vinelandii} has long been recognized as a robust nitrogen-fixing species, and strain CA6 represents an unusual type of hydrogen-producing microbe. It is an obligate aerobe, unlike strains that produce hydrogen through fermentation; and also a chemotroph, deriving its energy from organic substrates rather than light. Such a biohydrogen strategy has the potential to fill a special niche in the range of hydrogen production technologies in which these characteristics are desirable. CA6’s hydrogen yield from sucrose is about 5- to 60-fold lower than other candidate biohydrogen systems (mostly fermentative), though only up to 16-fold lower when tungstate was present; it may be better suited as a model for aerobic hydrogen than for industrial production, but modifications to its genome and culture conditions may lead to dramatic improvements. Several such potential improvements—e.g. the optimization of \textit{A. vinelandii}’s exposure to oxygen, and removal of a metabolic pathway that may draw electrons away from nitrogenases—will be discussed in the following chapter.

To conclude this chapter, we have characterized CA6, a mutant strain of \textit{A. vinelandii} lacking a number of important genes, which produces significant amounts of hydrogen.
CHAPTER 4. Evaluation of Azotobacter vinelandii Biomass and Hydrogen Yields
Growing Under Increasing Dissolved Oxygen Exposure

Introduction

In Chapter 3, Azotobacter vinelandii CA6 and other strains were evaluated in highly aerated environments to ensure that carbon remained limiting and utilized through the same pathway over a wide range of substrate concentrations. However, as described in Chapter 1, exposure of A. vinelandii to excess oxygen in nitrogen-fixing conditions induces a change in metabolism called “respiratory protection,” characterized by an increase in the rate of respiration without a corresponding increase in yields. Nevertheless, A. vinelandii is an obligate aerobe, and oxygen-limiting conditions induce other metabolic changes including an alternative pathway of carbon utilization, also discussed in Chapter 1: the production of polyhydroxyalkanoates—carbon storage polymers that also act as an electron sink—the most common being poly-β-hydroxybutyrate (PHB). These metabolic changes at different levels of oxygen exposure are part of A. vinelandii’s physiology that affects hydrogen yields and complicates the interpretation of metabolic models. One study on a hydrogenase-knockout strain of A. vinelandii found a near six-fold decrease in hydrogen yields when dissolved oxygen levels were increased twelve-fold.

We investigated the possibility that changing culture aeration could improve yields of hydrogen in A. vinelandii CA6 and other strains described herein. Using chemostats with varied levels of dissolved oxygen, we compared double mutants—lacking both hydrogenase and the ability to produce polyhydroxyalkanoates—to single mutants, lacking only
hydrogenase (or, in the case of CA6, the hydrogenase and adjacent operons). We predicted that double mutants would show a more linear relationship between biomass and dissolved oxygen, that excess aeration would show an inverse relationship to apparent yields of hydrogen and biomass, and that CA6’s use of an alternative nitrogenase would affect yields to a greater extent at lower dissolved oxygen than at higher.

**Methods**

*Strains and Media*

Strains of *Escherichia coli* and *A. vinelandii* utilized in this study are listed in Table 4.1, along with plasmids and primers used. *A. vinelandii* CA6 and DJ are part of NCSU’s Department of Plant and Microbial Biology Culture Collection.

*E. coli* strains were grown in lysogeny broth (LB) (with 1.5% w/v agar added for solid medium). Unless otherwise noted, *A. vinelandii* strains were grown in modified nitrogen-free Burk broth: unless otherwise noted, it contained, per liter, 0.2 g MgSO₄·7H₂O, 90 mg CaCl₂·2H₂O, 0.25 mg NaMoO₄·2H₂O, 0.2 g KH₂PO₄, 0.8 g K₂HPO₄, 2.5 mg FeSO₄·7H₂O, 2.5 mg ferric citrate, and 20 g sucrose. Sucrose and iron solutions were filter-sterilized with 0.2 μm filters; other solutions were autoclaved and the various components were combined as needed. Inocula were cultured by incubating at 30°C, shaking at 200 rpm.

Antibiotics were added to media as needed: 100 μg/mL ampicillin for *E. coli* or 50 μg/mL for *A. vinelandii*; 1 μg/mL kanamycin for *A. vinelandii*. 
Table 4.1. Strains, plasmids, and primers. Bold sequence indicates added restriction site; underlined sequence indicates homology to target.

<table>
<thead>
<tr>
<th>Name</th>
<th>Relevant characteristics</th>
<th>Reference or source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. vinelandii</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA6</td>
<td>Tungsten-tolerant, spontaneous mutant of CA</td>
<td></td>
</tr>
<tr>
<td>CA6 ΔphbC</td>
<td>CA6 with phbC-knockout, amp&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>HS2</td>
<td>CA with hoxK-knockout, kan&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>DJ</td>
<td>High-frequency transforming variant of CA</td>
<td>This study</td>
</tr>
<tr>
<td>DJ ΔhoxK</td>
<td>DJ with hoxK-knockout, kan&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td>DJ ΔphbC ΔhoxK</td>
<td>DJ with phbC- and hoxK-knockouts, kan&lt;sup&gt;f&lt;/sup&gt; and amp&lt;sup&gt;f&lt;/sup&gt;</td>
<td>This study</td>
</tr>
<tr>
<td><strong>E. coli</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOP10</td>
<td>F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(araleu)7697 galU galK rpsL (StrR) endA1 nupG</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Thunderbolt GC10</td>
<td>F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80lacZΔM15 ΔlacX74 endA1 recA1 Δ(ara, leu)7697 araD139 galU galK nupG rpsL λ– T1R</td>
<td>GeneChoice</td>
</tr>
<tr>
<td>DH5α</td>
<td>supE44 ΔlacU169 hsdR17 recA1 endA1 gyrA96 hi-1 relA1</td>
<td>Invitrogen</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUC19</td>
<td>amp&lt;sup&gt;f&lt;/sup&gt;, lacZ α-subunit, pMB1 origin</td>
<td>New England Biolabs</td>
</tr>
<tr>
<td>pUC-phb</td>
<td>pUC19 + A. vinelandii genomic inserted at multi-cloning site</td>
<td>This study</td>
</tr>
<tr>
<td>pUC-phb+amp</td>
<td>pUC-phb with amp&lt;sup&gt;f&lt;/sup&gt; cassette inserted in A. vinelandii phbC region</td>
<td>This study</td>
</tr>
<tr>
<td><strong>Primers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pUCphbCF</td>
<td>AGATGT-<strong>A</strong>AGCTT**-**</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>(HindIII site) TATGGTCCCTTGCTGCTTGAG</td>
<td></td>
</tr>
<tr>
<td>pUCphbCR</td>
<td>TGAACA-<strong>G</strong>AATTC**-**</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>(EcoRI site) GCCTTTCGCCCTATGTTCCTTC</td>
<td></td>
</tr>
<tr>
<td>pUCampF</td>
<td>AGTAGT-<strong>G</strong>TGCA-<strong>G</strong></td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>(PstI site) TGGTCTGACAGTTACCAATG</td>
<td></td>
</tr>
<tr>
<td>pUCampR</td>
<td>AGTACA-<strong>G</strong>AGCTC**-**</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td>(SacI site) GCCTCGTGATACGCCTATT</td>
<td></td>
</tr>
<tr>
<td>ampTestF</td>
<td>GTCAGCTCAGACGCGCCCTTCGTAG</td>
<td>This study</td>
</tr>
<tr>
<td>ampTestR</td>
<td>TATGCAGCAGCGCAGCGCAGCGCAGCGAAG</td>
<td>This study</td>
</tr>
<tr>
<td>phbIntLF</td>
<td>CGGCTGGTGCTGTTGATG</td>
<td>This study</td>
</tr>
</tbody>
</table>
Table 4.1 Continued

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>phbIntLR</td>
<td>ACGCTCGTCGTTTGGTGATGG</td>
<td>This study</td>
</tr>
<tr>
<td>phbIntDF</td>
<td>TCGCCGCATACACTATTCTC</td>
<td>This study</td>
</tr>
<tr>
<td>phbIntDR</td>
<td>AGGGTCCTTGGCAAGTATTCC</td>
<td>This study</td>
</tr>
<tr>
<td>hoxK2F</td>
<td>CGGCTTGTCGCTGATGACCTT</td>
<td>This study</td>
</tr>
<tr>
<td>hoxK2R</td>
<td>CGGTATGGAGCGGGGATTTTC</td>
<td>This study</td>
</tr>
<tr>
<td>kanR5</td>
<td>GGGATCGACGTGGTGAGTAAC</td>
<td>This study</td>
</tr>
</tbody>
</table>

Polymerase Chain Reactions and Restriction Digestions

Primers were designed using Clone Manager 9 software (Sci-Ed Software, Cary, NC) and ordered from Integrated DNA Technologies (IDT, Coralville, IA); see Table 4.1 for oligo sequences. Polymerase chain reactions (PCR) were performed using Qiagen Taq polymerase (Qiagen, Valencia, CA) or Phusion High-Fidelity polymerase (Bio-Rad, Hercules, CA) according to manufacturer’s instructions.

Restriction digests were performed as specified later with HindIII, EcoRI, PstI, SacI (Promega, Madison, WI), Sau3AI (Fisher, Waltham, MA), or NdeI (New England Biolabs, Ipswich, MA) according to manufacturers’ instructions.

Preparation of Competent E. coli Cells

_E. coli_ Thunderbolt electrocompetent cells were prepared by growing shaking in LB at 37°C up to an optical density at λ = 600nm (OD_{600}) of 0.5 to 1, then chilled on ice for 30 minutes and pelleted in chilled rotor at 4000 x g for 20 minutes, resuspended in cold water, pelleted and resuspended again, pelleted again and resuspended in sterile 10% (v/v) glycerol, then pelleted and resuspended in 10% glycerol to approximately 3 x 10^{10} cells per mL. Fifty-
microliter aliquots were frozen on dry ice and stored at -80°C. *E. coli* DH5α competent cells were prepared as described previously481.

**Plasmid Construction**

*E. coli* TOP10 competent cells (Invitrogen, Carlsbad, CA) were transformed with plasmid pUC19 (Invitrogen) by the heat-shock method according to manufacturer’s instructions, plated on LB with ampicillin to select for transformants. Plasmid was extracted from transformant cultures by the alkaline lysis method482 using a Qiagen Miniprep kit according to manufacturer’s instructions, and plasmid size was determined by 1% agarose gel electrophoresis with ethidium bromide staining. Plasmid identity was further confirmed by band pattern analysis by electrophoresis after cutting plasmid with restriction enzyme Sau3AI (Fisher) according to manufacturer’s instructions.

A region of *A. vinelandii* CA6 genome spanning from within phbC to within phbA was amplified by PCR with HindIII and EcoRI sites incorporated into forward and reverse primers, respectively (primers pUCphbCF and pUCphbCR; see Table 4.1). PCR and pUC19 plasmid were digested with HindIII and EcoRI according to manufacturer’s instructions and purified by gel extraction with Zymo Research Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA) according to manufacturer’s instructions. Products were ligated with Promega T4 DNA Ligase according to manufacturer’s instructions. Electrocompetent *E. coli* Thunderbolt cells were transformed with ligation using electroporation: 1 μL ligation was mixed with 50 μL cells, then a ECM399 electroporator (BTX, San Diego, CA) applied 2.5 kV at 25 μF to an electroporation cuvette containing the mixture. One milliliter LB was
added, the mixture was transferred to a centrifuge tube, and the tube was incubated at 37°C for 20 minutes. Cells were plated onto LB agar with ampicillin and 800 μg 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and incubated at 37°C for 24 hours, then white colonies were screened by PCR with pUCphbCF and pUCampF primers on extracted plasmids.

The resulting pUC-phb plasmid, as well as an ampicillin resistance cassette amplicon generated by PCR with pUCamp primers with pUC19 as template, were cut with SacI and PstI; these enzymes remove a 201-bp section of gene from phbC. After purification by gel-extraction, restricted plasmid was treated with Antarctic Phosphatase (New England Biolabs) according to manufacturer’s instructions for a 3’ overhang and ligated together with the ampicillin resistance cassette amplicon. Competent *E. coli* DH5α cells were transformed with this ligation by the heat-shock method except that the heat-shock interval was 1 min and 300 μL LB was added; transformants were plated on LB with ampicillin and incubated at 37°C for 24 hours. Colonies were screened by colony PCR using pUCphbC and phbInt primers, and a positive was saved as pUC-phb+amp.

*Transformation of A. vinelandii.*

*A. vinelandii* competence was induced and cells were transformed using previously described methods. In brief, strains CA6 and DJ were grown on transformation (TF) medium (Burk agar with 10-fold increased magnesium, 28mM ammonium acetate, and no added iron or molybdenum) for 3-4 days until green pigment appeared. Growth was washed off with 1 mL sterile Burk buffer (0.2 g/L KH₂PO₄ and 0.8 g/L K₂HPO₄), mixed with 50 μL
pUC-phb+amp plasmid linearized with NdeI and/or HS2 PCR product (comprising hoxK flanking regions surrounding a kanR cassette amplified from A. vinelandii HS2 using primers hoxK2F and hoxK2R), incubated at 30°C for 1 hour, and then plated onto selective Azotobacter growth (AG) agar (Burk with 0.5 g/L yeast extract added) with ampicillin and/or kanamycin and incubated at 30°C. Colonies were screened by PCR using phbInt and ampTest and/or hoxK2F+kanR5 primers.

Bioreactor Setup and Oxygen Requirement Chemostats

Chemostat experiments were performed as described in Chapter 3. Temperature was held at 30°C; pH was allowed to vary; filtered compressed air was sparged at a rate of 0.315±0.05 liters per minute (lpm); the system was open to gas exchange with the atmosphere. Culture volume was maintained constant across different agitation rates by raising a metal dip tube to the appropriate height. Exhaust gases and culture turbidity were monitored and recorded redundantly in real-time. Medium was prepared in 20-L glass carboys and consisted of modified Burk medium, with 10 g/L sucrose, 5 mg/L FeSO₄·7H₂O, and 0.5 g/L sodium citrate (to prevent precipitation). Sodium tungstate dehydrate was added to a concentration of 1 mM as noted. A sample of each batch of medium was saved at -80°C for later analysis.

To begin an experiment, the reactor, set to maintain the appropriate temperature (30°C) and dissolved oxygen concentration (30%), was inoculated by injection of pre-grown batch cells. When the reactor culture reached exponential phase, continuous culture (the inflow of fresh medium and removal of excess culture volume) was initiated. Steady state was achieved before starting any sampling.
System parameters were defined as follows: dilution rate is the value obtained when feed and harvest flows are equal, which allows for obtaining a constant volume of culture in the reactor \((D = F/V)\), with \(D\) as dilution rate, \(F\) as flow rate, and \(V\) as culture volume). Yield is defined as the ratio of product generated to limiting substrate (sucrose) consumed. Productivity is the value of product generated per limiting substrate consumed per unit of time. The dilution rate for these experiments was set to 0.066 h\(^{-1}\). Once steady state was achieved, at least three samples were taken and analyzed before the agitation rate was changed. At least one retention period was allowed to pass between each sample; at \(D = 0.066\) h\(^{-1}\), 700 mL pass through the system in approximately 15 h.

Sample Analysis

Each sample was processed as described in Chapter 3. A small volume was spun down for 5 min at 10,000 rpm in a Centrifuge 5415C (Eppendorf, US), then the supernatant and pellet were saved in separate tubes at -80°C for later analysis. \(OD_{600}\) was determined using an Ultrospec 1100 pro spectrophotometer. Dry weight was obtained by filtering a portion of sample through a 0.2 μm filter of known mass (mixed cellulose esters, EMD Millipore, Germany) using vacuum suction; the filter was then dried at 60-70°C for 7 days and weighed to determine the dry weight of biomass per sample volume.

Samples from \(phbC\)- and \(hoxK\)-knockout mutants were tested for antibiotic resistance by pipetting 10-ul spots from \(10^{-2}\), \(10^{-3}\), \(10^{-4}\), and \(10^{-5}\) dilutions onto selective and non-selective agar, 3 spots per plate per dilution. These were incubated at 30°C, and then colonies were counted and colony-forming units per ml values were calculated.
The initial and residual concentrations of sucrose were determined from samples of each batch of initial medium and from supernatants of reactor samples analyzed by high-performance liquid chromatograph (Shimadzu, Japan) run under isocratic conditions at 65°C (mobile phase was water at 0.5 mL/min, Supelcogel™ Ca column (300 mm x 7.8 mm, Supelco™ Analytical, Bellefonte, PA, USA). The column was coupled to a refractive index detector. Absolute values were obtained by measuring known quantities of sucrose to generate a standard curve.

*Nile Blue Staining of Poly-β-hydroxybutyrate*

Nile Blue A (Chem-Impex Int’l Inc, Wood Dale, IL, USA) was dissolved in water to 1% w/v. This solution was applied to *A. vinelandii* cells air-dried onto a glass slide, itself air-dried for 10 min at 30°C, then viewed under Axioskop2 oil immersion microscope with phase contrast and epifluorescence at 450nm for excitation of emission at approximately 600nm.

*Statistics*

Significance reporting is the result of Gosset’s *t* tests, with significance indicating *p* < 0.05.

*Results*

*Amount of Aeration Affects Hydrogen Production in Azotobacter vinelandii*

Chapter 3 examined hydrogen yields in various strains of *A. vinelandii* in highly aerated conditions. In these conditions, *A. vinelandii* strain CA6, lacking 42 genes due to a large
deletion (Chapter 2), produced equal amounts of hydrogen to strain HS2 which lacks only its uptake hydrogenase\textsuperscript{154}; the yields from both were around 30 mmoles H\textsubscript{2} per mole sucrose consumed. However, oxygen exposure affects the activity of nitrogenase, and thus the production of hydrogen in \textit{A. vinelandii}\textsuperscript{86,88,224,448,483}.

To explore this relationship, we conducted a series of chemostat culture experiments in iron-sufficient and nitrogen-free medium, as previously\textsuperscript{20}. The strains grown in these experiments were both hydrogenase-deficient derivatives of \textit{A. vinelandii} prototrophic strain CA: strain CA6 and a \textit{hoxK} knockout strain, DJ Δ\textit{hoxK}. Aeration of the cultures was varied by controlling the agitation speed (between 100 and 750 rpm), while airflow was kept constant to prevent dilution of gaseous products. Measurements and samples were taken when cultures reached steady state at each agitation setting.

Upon each increase in agitation speed, the culture ceased growth and hydrogen production (and thus nitrogen fixation) while carbon dioxide production transiently increased (Figure 4.1). This behavior is consistent with previous observations of conformational and respiratory protection strategies in \textit{A. vinelandii}\textsuperscript{9,25,87,118,219}. 

87
Figure 4.1. Example of cessation of hydrogen production (□) and concurrent increase in CO₂ production (Δ) upon increase in agitation (marked by dashed line at time 0). Data taken from A. vinelandii DJ ΔphbC ΔhoxK culture upon increase from 250 to 500rpm

The observed rate of oxygen transfer from gas to liquid phase did not increase proportional to the change in agitation speed across the whole range of agitation settings; the largest change in rate was observed between 250 and 750rpm (Figure 4.2a). Lack of a linear relationship is expected because the formation of a vortex (which occurred at 500rpm but not below) greatly increases oxygen transfer rates⁴⁸⁴. Oxygen remained the limiting substrate at agitation speeds below 500rpm, so dissolved oxygen remained undetectable and significant residual sucrose was measured in the supernatant of samples (Figure 4.2a). At 500rpm and above, carbon became limiting, so measured dissolved oxygen levels increased with stepwise
increases of agitation. Specific oxygen consumption increased most dramatically between 250 and 750rpm (Figure 4.2b).

![Graph of specific oxygen consumption and residual sucrose](image)

**Figure 4.2.** Squares (□) represent strain CA6 and triangles (Δ) represent strain DJ ΔhoxK. Error bars represent standard deviations. a) Percent saturation of dissolved oxygen (empty symbols) and percent residual sucrose in supernatant (full symbols) in steady state at each agitation speed. b) Specific oxygen consumption (mmol O\textsubscript{2} per hour per g dry weight) in steady state at each agitation speed; consumption was calculated as the difference in oxygen partial pressures between inflow and exhaust gases.

As predicted, the greater the agitation/dissolved oxygen, the lower the yield of hydrogen in each strain (Figure 4.3a). Compared to a yield of about 70 mmol H\textsubscript{2} per mol sucrose at 750rpm, CA6 produced about 163 mmol mol\textsuperscript{-1} g\textsuperscript{-1} at 100rpm. Also as predicted, as dissolved
oxygen decreased, the difference in specific hydrogen yields between CA6 and the single-gene-knockout strain DJ ΔhoxK increased, indicating that the lack of regulation of alternative nitrogenases in CA6 plays a larger role in hydrogen production at lower aerations.

![Graph showing the relationship between hydrogen yield and agitation speed.](image)

Figure 4.3. Empty symbols represent PHB-producing strains and full symbols PHB knockout strains. Squares (□) represent strain CA6 and triangles (Δ) represent strain DJ ΔhoxK. Error bars represent standard deviations. a) Hydrogen yields (mmol H₂ per mol sucrose) in steady state at each agitation speed. b) Dry weight yields (C-mol dry weight per mol sucrose) in steady state at each agitation speed. c) Carbon dioxide yields (mol CO₂ per mol sucrose) in steady state at each agitation speed.

Yields of biomass (measured as dry weight) also decreased as dissolved oxygen increased, especially between 250 and 750rpm (Figure 4.3b). The rate of this decrease was
measured as about 0.2 mg dry weight loss per gram of sucrose for each mg increase in dissolved oxygen per hour. Carbon dioxide yields increased as agitation speed increased, especially between 250 and 750rpm (Figure 4.3c); the pattern correlated well with substrate-corrected oxygen consumption, with coefficients of determination ($R^2$) between 0.748 and 0.999.

Not surprisingly, at lower aerations, poly-β-hydroxybutyrate (PHB) granules were observed in strains CA6 and DJ ΔhoxK. Granule frequency in CA6 was highest at lower agitation speeds (100 and 250rpm) and decreased at 500rpm and above. In DJ ΔhoxK, frequency increased up to a peak at 500rpm and then decreased (Table 4.2).

Table 4.2. Poly-β-hydroxybutyrate (PHB) observed in cells at different agitation speeds

<table>
<thead>
<tr>
<th>A. vinelandii strain:</th>
<th>Agitation speed (rpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
</tr>
<tr>
<td>CA6</td>
<td>+++</td>
</tr>
<tr>
<td>CA6 ΔphbC</td>
<td>—</td>
</tr>
<tr>
<td>DJ ΔhoxK</td>
<td>+</td>
</tr>
<tr>
<td>DJ ΔphbC ΔhoxK</td>
<td>—</td>
</tr>
</tbody>
</table>

Key: +++ to + indicates amount of PHB observed, — means none, ND means not determined.

Production and Confirmation of phbC Knockout Strains

We feared that accumulation of PHB storage polymer might skew the observed relationships between biomass and product yields—PHB production leads to increased cell size/mass without concurrent increase in catalytic capacity, resulting in lower apparent specific yields. To evaluate this relationship, we needed mutants of A. vinelandii deficient in PHB production but otherwise isogenic with strains CA6 and DJ ΔhoxK. Review of the
literature suggested that knocking out \textit{phbC}, rather than other genes in the \textit{phb} operon, would abolish PHB production completely with the fewest other effects\textsuperscript{34,313,314,327,485}. We constructed two new strains of \textit{A. vinelandii}—CA6 \textit{ΔphbC} and DJ \textit{ΔhoxK ΔphbC}—lacking a functional PHB synthase (PhbC) by inserting an ampicillin resistance cassette near the 5’ end of \textit{phbC} by double recombination. This PHB-negative phenotype was confirmed by growing cells in low-oxygen, high-carbon medium and then staining PHB granules with Nile Blue A (Figure 4.4). We attempted to create a knockout strain from HS2 also, but were unable to isolate any transformants of this strain; additionally, PHB granules were never observed in this strain under any condition.

Figure 4.4. The presence or absence of PHB granules in \textit{A. vinelandii} strains stained with Nile Blue A, grown in low-aeration, high-sugar conditions, observed under phase contrast and epifluorescence microscopy
Tangentially, we observed motile cells in cultures of both DJ Δ*hoxK* and DJ Δ*phbC* Δ*hoxK*, but have never observed motility in CA, HS2, or CA6 strains. This gives phenotypic support for the motility-reducing mutations observed in these strains’ genomes (stop codons introduced in two different flagellum-related genes in CA and CA6) (Chapter 2).

*Evaluation and Comparison of PHB-Negative Strains in Chemostats*

As before with *A. vinelandii* strains CA6, HS2, and DJ Δ*hoxK*, we conducted chemostat experiments with varied agitation speeds with the new knockout strains. Over the course of these experiments, we observed no decrease in the proportion of cells resistant to the antibiotics used to select for transformants; integration of the constructs into their genomes appeared to be stable. Additionally, we observed no PHB granules at any point in CA6 Δ*phbC* or DJ Δ*phbC* Δ*hoxK*.

We expected that PHB production would affect biomass measurements at lower aerations, and indeed observed significantly lower dry weight yields in strain CA6 Δ*phbC* compared to CA6 at agitation speeds 100, 250, and 500rpm, but not at 750rpm, and DJ Δ*hoxK* yields were significantly higher than yields of DJ Δ*phbC* Δ*hoxK* at 500rpm (peak PHB production) but not at other agitation speeds (Figure 4.3b). In terms of linearity of relationship between biomass yield and aeration, CA6 and CA6 Δ*phbC* were about equal (R² ~ 0.91 from 100 to 750rpm) while DJ Δ*phbC* Δ*hoxK* strain’s relationship was much more linear than DJ Δ*hoxK* (R² = 0.86 vs. 0.26, respectively). Relationships between CO₂ yield and oxygen consumption were also more linear in PHB-negative strains (average R² 0.99 vs. 0.87 for PHB-positive strains).
Other parameters, such as CO₂ yields and productivities, hydrogen yields and productivities, and oxygen consumption rates showed similar patterns and values over the range of agitation speeds (Figure 4.3).

**Tungsten Increases Hydrogen Yields in Hypoxic Conditions**

In our previous study, adding sodium tungstate to a highly aerated culture of *A. vinelandii* CA6 increased its yield of hydrogen produced over four-fold²⁰. This effect is presumably due to tungstate poisoning of the molybdenum nitrogenase, forcing CA6 to rely on its alternative nitrogenases that produce more moles of hydrogen per mole of dinitrogen fixed²⁸,⁴⁶,¹⁷¹. To examine this effect at lower aerations, we added sodium tungstate to the inflowing medium at a concentration of 1 mM in cultures of CA6 and CA6 ΔphbC in 250 and 500 rpm agitation conditions.

In strain CA6, addition of tungsten at 500rpm reduced the yield of dry weight by half, while CO₂ yield and oxygen consumption remained about the same. Hydrogen yield was 30% higher, increased from about 109 mmol/mol sucrose to about 145 mmol/mol. In CA6 ΔphbC, dry weight yield, CO₂ yield, and oxygen consumption did not change, but hydrogen yield increased about 86%. At 250rpm, the increase in hydrogen yield was about 50%, up to the highest yield observed in this species: 265.4 mmol/mol (Figure 4.5).
Figure 4.5. Hydrogen yields (mmol H₂ per mol sucrose) in strain CA6 in steady state at each agitation speed, with or without tungsten added. Empty bars represent tungsten-free conditions and full bars represent the presence of 1 mM sodium tungstate. Error bars represent standard deviations.

Discussion

In this study of *A. vinelandii* strains, changes in culture dissolved oxygen levels and the presence of tungsten in the medium allowed for a hydrogen yield over 9-fold higher than in previously tested high-aeration, tungsten-free conditions (from 28.1 to 265.4 mmol H₂/mol sucrose). Generally, cell exposure to lower dissolved oxygen resulted in higher yields of dry
weight and hydrogen, as expected because hydrogen tends to be growth-linked in diazotrophs; carbon dioxide generation decreased simultaneously as more of the available carbon went to biomass.

The effect of oxygen on *A. vinelandii’s* growth is well-studied. Increasing oxygen beyond limiting concentrations can greatly increase the organism’s maintenance coefficient, reducing the efficiency of growth by requiring the allotment of more and more available energy to respiratory protection. The concurrent increase in CO₂ production with oxygen consumption (Figures 4.2b and 4.3c) demonstrates this “wasting” of both substrate and oxygen, resulting from an increase in aldolase activity that removes carbon as CO₂ from the tricarboxylic acid cycle. However, one limitation of this interpretation is that almost all studies of oxygen effects in *Azotobacter* have focused on the molybdenum nitrogenase, so how much these results can be extrapolated to the alternative isoenzymes remains a mystery. We observed one behavior separating CA6 strains (which produce an alternative nitrogenase in all nitrogen-fixing conditions) from their more highly regulated siblings: that hydrogen yields in CA6 backgrounds increased more drastically at lower aerations than in other backgrounds (Figure 4.3a). This may be meaningful, and justifies the comparison of the CA6 strains to their single- and double-knockout cousins. It is possible that the iron-only nitrogenase is more oxygen-sensitive than the Mo version, and thus is more active at lower aerations than at higher.

The measurement of hydrogen can be used as a proxy for nitrogenase activity. The Mo nitrogenase consistently produces 1 mole hydrogen for every two moles of ammonia generated, so activity of the nitrogenase system and the yield and productivity of
ammonia can easily be calculated. Thus, using a single-gene uptake hydrogenase knockout strain such as our DJ ΔhoxK and a continuous measurement of hydrogen as in our system, it is possible to monitor nitrogenase activity in many different conditions in real time, much more simply than the typical acetylene reduction assay. However, the precision this approach allows is limited to applications measuring only Mo nitrogenase activity. Due to the diversity of H₂-to-N₂ ratios observed for the V and Fe nitrogenases²⁸,⁴⁶,¹³¹,¹⁷¹,⁴¹⁰, the relationship of hydrogen production to nitrogenase activity for these enzymes is uncertain; therefore, cells expressing these enzymes alone, or in combination with the Mo version as in the case of CA6¹⁶⁷, cannot be precisely monitored using this approach.

Intuitively, it seems like hydrogen evolution due to lack of uptake hydrogenase should decrease the overall fitness and growth yield of the cells. However, in our previous work²⁰ and in that of others in A. vinelandii and Azotobacter chroococcum⁴³⁶,⁴⁴¹,⁴⁴², lack of hydrogenase appears to have a neutral or even positive effect on fitness, especially in hypoxic conditions (such as in the present study). This may be because the electron transport chain of hydrogenase is more tightly coupled to phosphorylation than that of carbon substrates²²⁷. However, it is fortunate for the production of hydrogen in A. vinelandii, since it means yields need not be compromised by reduction in fitness.

We expected the production of poly-β-hydroxybutyrate (PHB) in hypoxic conditions to complicate our comparison of A. vinelandii metabolism at different aerations, by increasing the observed biomass without concurrent increase in the culture’s catalytic capacity. This would result in lower observed specific yields (product per unit of biomass). By creating and comparing PHB-negative strains with their isogenic PHB-positive parents, we teased out this
effect. CA6 ΔphbC and DJ ΔphbC ΔhoxK produced significantly less biomass at aerations in which significant amounts of PHB were observed in their parent strains (Figure 4.3b). With CO₂ as a product, specific yields (corrected for biomass) were higher in PHB-negative strains, as expected since total biomass was higher but the catalytic fraction was not. This relationship was not observed with hydrogen as a product; PHB-positive strains seemed to produce higher yields when more PHB was observed in their cells. Further study is needed to verify the cause of this phenomenon. However, PHB serves as an electron sink when oxygen is limiting⁴¹⁰–⁴¹²; the reduction of acetoacetyl-coenzyme A (-CoA) to D-β-hydroxybutyryl-CoA requires the conversion of NADPH to NADP⁺ (Figure 1.3). This relieves the regulatory burden of a buildup of NADPH, allowing carbon utilization, and thus nitrogen fixation and metabolism in general, to continue⁴¹⁰,⁴¹¹. Therefore it is possible that PHB-producing strains are able to fix nitrogen (and thus produce hydrogen) to a greater extent than non-producing strains at lower aerations.

*A. vinelandii* is capable of adapting to a wide range of oxygen concentrations even when fixing nitrogen, which serves to enhance its fitness in nature, but makes study of its carbon metabolism more complex. When fixing nitrogen at high aeration, respiratory protection skews carbon flow toward CO₂, while at low aeration, a portion of the carbon assimilated goes to PHB production. When optimizing cellular yields in this organism, it is best to treat oxygen like wine: best in moderation. Likewise, moderation is advisable when optimizing hydrogen production in *A. vinelandii*; we observed up to 6-fold increases in hydrogen yields (in tungsten-free conditions) in low-aeration conditions compared with high-aeration conditions. However, the productivity of hydrogen was reduced when oxygen provided was
insufficient to allow complete consumption of the carbon source: the highest yield in CA6 was at 250rpm, but the highest productivity was at 500rpm. The optimal condition may require a compromise and fall in between these values, controlling agitation to achieve the minimum detectable dissolved oxygen levels.

Although the hydrogen yields in this study were lower than yields commonly observed in dark fermenters such as *Clostridium*, *Thermotoga*, and *Enterobacter* species\(^\text{20}\), *Azotobacter vinelandii* has other characteristics that make it valuable for various applications, as discussed in Chapter 1. As one example, hydrogen production is inextricably linked to nitrogen fixation, which could be harnessed to provide fixed nitrogen fertilizer to plants or other industrially relevant microorganisms\(^\text{18,280–282}\); assuming a one-to-one ratio of H\(_2\) produced per N\(_2\) fixed in DJ strains, these strains produce up to 8.5 mg ammonia per hour (Figure 4.6). The relationship between hydrogen and ammonia in CA6 is less straightforward because of its use of alternative nitrogenases. A process combining these products with hydrogen production is more likely to be economically competitive.
Figure 4.6. Ammonia productivity (mg per hour) in steady state at each agitation speed in DJ strains. Empty symbols represent PHB-producing strains and full symbols PHB knockout strains. Error bars represent standard deviations.

To conclude this chapter, we have further improved the production of hydrogen from *A. vinelandii* and determined the effects of modifying oxygen exposure and PHB generation pathways on this production.
CHAPTER 5. Hydrogen and Azotobacter vinelandii: Future Potential Directions

This work evaluated Azotobacter vinelandii as a model for aerobic biohydrogen production and generated proof-of-concept data demonstrating prospectives for process development. Azotobacter vinelandii strain CA6 produced hydrogen through an unknown genetic mechanism and contained other phenotypic limitations. This study identified those mutations and allowed for adjustment of growth conditions to accommodate them. From that point, knowledge of the organism’s biology permitted an increase in hydrogen yield over 9-fold through modification of environmental conditions and medium composition (Figure 5.1).

![Figure 5.1](image)

Figure 5.1. Hydrogen yields (mmol H$_2$ per mol sucrose) in A. vinelandii CA6 growing in chemostat culture at D = 0.066 h$^{-1}$ with 10 g/L initial sucrose in different conditions: low aeration/dissolved oxygen (250 rpm agitation), high aeration/dissolved oxygen (1000 rpm), and the presence or absence of 1 mM tungsten (W). See chapters 3-4
Aerobic biohydrogen production is an unusual approach and poorly studied. Most microbial biohydrogen research focuses on dark fermenters that shed excess reductant as hydrogen; phototrophs that use light energy to split water; or microbial fuel cells that generate enough current to split water\textsuperscript{429,432,433,487}. Aerobic, heterotrophic biohydrogen approaches are advantageous in that they do not require the exclusion of oxygen or the large surface area necessary for light capture. While it does have its own limitations, such as requirements for aeration and organic carbon substrates, combining it with other competitive industrial processes can increase the value of both.

\textit{A. vinelandii} offers a promising foundation for such a combination. As discussed in Chapter 1, it is a candidate for numerous biotechnological processes of interest. It produces lipoid polymers, polyhydroxyalkanoates, that can accumulate up to 70\% of the cells’ dry weight and can be purified and made into biodegradable thermoplastics\textsuperscript{289–293}. It fixes nitrogen, potentially providing this nutrient to plants or other industrially relevant organisms\textsuperscript{18,19,428}. And it produces another polymer, alginate, a polysaccharide used as a stabilizing, gel-forming, or thickening agent in food, pharmaceutical, and cosmetic industries\textsuperscript{354,355}. And this is only a partial list of its potential uses.

There are many approaches that may still improve \textit{A. vinelandii}’s hydrogen production even further. In a continuous system, the dilution rate could be increased up to near the maximum specific growth rate to optimize productivity (the quantity of hydrogen produced per unit of time). Increasing the concentration of carbon substrate to its maximum limiting value may also improve productivity and possibly even yield, as the carbon-to-oxygen ratio
may be a more important factor than oxygen alone. Alternatively or additionally, controlling agitation and maybe airflow to maintain dissolved oxygen at a minimum detectable level (just enough to be neither limited nor in excess) could increase process efficiency. A non-continuous system, such as fed-batch culture, may also be valuable.

On the catalyst side of optimization, the highly competent *A. vinelandii* strain DJ may be preferable to CA6. Knocking out the *modE* gene could reproduce the CA6 phenotypes of tungsten tolerance and alternative nitrogenases unregulated by molybdenum$^{206,212}$, allowing exploitation of the alternatives’ higher hydrogen production. Knocking out the *nifL* gene could allow hydrogen production from nitrogenase even in the presence of fixed nitrogen, such as from low-cost crude carbon substrates$^{213,215,216,463}$. This list is also partial. And all of these modifications can be done without antibiotic resistance cassettes remaining in the genome$^{64}$.

At the least, this work provides a proof-of-concept of the processes involved, which could provide benefits for other applications. The system’s full potential remains to be explored.
REFERENCES


52. Palanché, T., Blanc, S., Hennard, C., Abdallah, M. A. & Albrecht-Gary, A.-M. 
Bacterial Iron Transport: Coordination Properties of Azotobactin, the Highly 

53. Cornish, A. S. & Page, W. J. The catecholate siderophores of *Azotobacter vinelandii*: their 
affinity for iron and role in oxygen stress management. *Microbiology* **144**, 1747–


55. Harris, J. O. A Study of the Effect of Growth Substrate on the Respiration of 


57. Page, W. J. Production of poly-β-hydroxybutyrate by *Azotobacter vinelandii* strain 
UWD during growth on molasses and other complex carbon sources. *Appl Microbiol 

58. Page, W. J. & Knosp, O. Hyperproduction of Poly-β-Hydroxybutyrate During 

59. Page, W. J. Suitability of commercial beet molasses fractions as substrates for 
polyhydroxyalkanoate production by *Azotobacter vinelandii* UWD. *Biotechnol Lett* **14**, 

60. Page, W. J. Production of poly-β-hydroxybutyrate by *Azotobacter vinelandii* UWD in 
media containing sugars and complex nitrogen sources. *Appl Microbiol Biotechnol* **38**, 

61. Page, W. J. Genetic transformation of molybdenum-starved *Azotobacter vinelandii*: 


89. Boiardi, J. L. Metabolic cost of nitrogen incorporation by N\textsubscript{2}-fixing \textit{Azotobacter vinelandii} is affected by the culture pH. \textit{Biotechnol Lett} \textbf{16}, 1195–1198 (1994).


403. OECD. Ethylene (Screening Information Data Set). (2005).


