

# Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*

Christina Prust<sup>1</sup>, Marc Hoffmeister<sup>2</sup>, Heiko Liesegang<sup>2</sup>, Arnim Wiezer<sup>2</sup>, Wolfgang Florian Fricke<sup>2</sup>, Armin Ehrenreich<sup>2</sup>, Gerhard Gottschalk<sup>2</sup> & Uwe Deppenmeier<sup>3</sup>

*Gluconobacter oxydans* is unsurpassed by other organisms in its ability to incompletely oxidize a great variety of carbohydrates, alcohols and related compounds. Furthermore, the organism is used for several biotechnological processes, such as vitamin C production. To further our understanding of its overall metabolism, we sequenced the complete genome of *G. oxydans* 621H. The chromosome consists of 2,702,173 base pairs and contains 2,432 open reading frames. In addition, five plasmids were identified that comprised 232 open reading frames. The sequence data can be used for metabolic reconstruction of the pathways leading to industrially important products derived from sugars and alcohols. Although the respiratory chain of *G. oxydans* was found to be rather simple, the organism contains many membrane-bound dehydrogenases that are critical for the incomplete oxidation of biotechnologically important substrates. Moreover, the genome project revealed the unique biochemistry of *G. oxydans* with respect to the process of incomplete oxidation.

Aerobic microorganisms usually oxidize their carbon sources to carbon dioxide and water. During this process, energy and intermediary metabolites required for biosynthesis are generated. In contrast, many acetic acid bacteria oxidize their substrates incompletely even under normal growth conditions. Among them is the  $\alpha$ -proteobacterium *G. oxydans*, a Gram-negative, obligately aerobic and rod-shaped acidophilic organism<sup>1</sup> that belongs to the family *Acetobacteriaceae*. *G. oxydans* is known for its incomplete oxidation of a wide range of carbohydrates and alcohols in a process that is referred to as oxidative fermentation. The corresponding products (aldehydes, ketones and organic acids) are secreted almost completely into the medium. The organism is able to grow in highly concentrated sugar solutions and at low pH values. High oxidation rates correlate with low biomass production, making *G. oxydans* strains suitable objects for industrial application<sup>2</sup>.

Because these organisms oxidize many substrates regioselectively, *G. oxydans* strains are used for the production of L-sorbose from D-sorbitol (vitamin C synthesis)<sup>3</sup> and 6-amino-L-sorbose from 1-amino-D-sorbitol for the synthesis of the antidiabetic drug Miglitol<sup>4</sup>. They also can be used industrially to produce D-gluconic acid, and ketogluconic acids from D-glucose and dihydroxyacetone from glycerol<sup>5</sup>. Strains of this genus also produce aliphatic, aromatic carboxylic and thiocarboxylic acids that find applications as flavoring ingredients<sup>6</sup>. Furthermore, *Gluconobacter* enzymes and whole cells are thought to be ideal for use as the biological element in sensor systems for the detection of alcohols, sugars and polyols<sup>7</sup>. Natural habitats of *G. oxydans* are flowers and fruits<sup>8</sup>. The organism is also found in

alcoholic beverages (wines and beers) and soft drinks where it causes off-flavors and spoilage<sup>9</sup>.

Here we report the complete genome sequence of *G. oxydans* 621H (DSM 2343). The genome data provide a rich source for metabolic reconstruction of the pathways leading to industrially important products derived from sugars and alcohols.

## RESULTS

### General features of the *G. oxydans* genome

The entire genome of *G. oxydans* 621H was sequenced to obtain detailed insights into the oxidative potential of the organism and to elucidate the mechanisms of incomplete oxidation. The circular chromosome consists of 2,702,173 base pairs (bp) with a G+C content of 60.8% (Table 1). In addition five plasmids were identified (163.1, 26.6, 14.6, 13.2 and 2.7 kb). Hence, the total size of the genome is 2,922,384 bp. There are a total of 2,664 predicted protein-encoding open reading frames (ORFs), four copies of rRNA operons and 55 genes encoding tRNAs. It is predicted that 89.9% of the DNA encodes proteins or stable RNAs. Biological roles were assigned to 1,877 ORFs (70.5%; see Supplementary Table 1 online). 446 ORFs showed similarities to hypothetical proteins from other organisms, and 341 ORFs were unique to *G. oxydans* (Table 1). A total of 1,650 proteins showed high homologies to proteins from other bacteria (threshold: blastp e-value < e<sup>-15</sup>). Of these proteins 76.4% (1,260) were most homologous to proteins from  $\alpha$ -proteobacteria (see Supplementary Table 2 online). Most prominent were *Rhodospirillum rubrum* and *Magnetospirillum magnetotacticum*. The GC skew analysis<sup>10</sup> of the

<sup>1</sup>Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August-University, Grisebachstr. 8, D-37077 Goettingen, Germany. <sup>2</sup>Goettingen Genomics Laboratory and Competence Center for Genome Research on Bacteria, Institute of Microbiology and Genetics, University of Goettingen, Grisebachstrasse 8, 37077 Goettingen, Germany. <sup>3</sup>Department of Biological Sciences, University of Wisconsin-Milwaukee, 3209 N. Maryland Ave., Milwaukee, Wisconsin 53211, USA. Correspondence should be addressed to U.D. (udeppen@uwm.edu).

**Table 1** General features of the *G. oxydans* genome

Size of the chromosome (bp)	2,702,173
Size of plasmids (bp)	163,186 (pGOX1) 26,568 (pGOX2) 14,547 (pGOX3) 13,223 (pGOX4) 2,687 (pGOX5)
Coding region (%)	89.9%
G+C (%)	60.8%
Ribosomal RNAs:	23S (4 copies) 16S (4 copies) 5S (4 copies)
tRNAs	55
Total number of ORFs	2,664
ORFs with assigned function	1,877
Conserved hypothetical ORFs	446
ORFs unique to <i>G. oxydans</i>	341
IS elements	18
Transposases	103

chromosome from *G. oxydans* indicated a bidirectional replication that starts at the proposed *oriC* sequence near the *dnaA* gene and ends near the calculated replication terminus at around 1.3 Mb (Supplementary Fig. 1 online).

Also present in the *G. oxydans* genome is a large number of repeated DNA elements, which are known to be involved in genomic rearrangements. We identified 82 insertion sequences (IS) and 103 transposase genes (Supplementary Table 1 online). Some of these copies are partially deleted, and therefore, presumed defective. Most of the IS, however, appear to have functional copies and may be responsible for the genetic instability leading to deficiencies in various physiological properties as observed in a variety of acetic acid bacteria<sup>11</sup>. Some IS elements can be grouped into the families IS12528 (ten copies) and IS1032 (eight copies)<sup>12</sup>.

*G. oxydans* has many two-component regulatory systems (15 sensor kinases and 16 response regulators, including 6 sensor-regulator pairs), which mediate the response to various environmental stimuli. The genome also codes for 75 regulatory proteins, many of which belong to previously described regulator families such as LysR (13), AraC (4), MarR (3), TetR (6), AsnC (2) and MerR (2). *G. oxydans* appears to have a simple chemosensory transducer protein system. The genome contains one complete set of chemotaxis genes that are organized in one gene cluster, and three copies of genes encoding methyl-accepting chemotaxis proteins (MCPs). The chemotactic response is accomplished by signal transmission between the receptor complexes and the flagellar motor complexes. There are two copies of CheY (GOX1551, GOX1556), which act as messenger proteins, transducing the signal from the MCPs to the flagella<sup>13</sup>. The genome contains 27 predicted flagellar genes, most of them organized in six gene clusters (see Supplementary Table 1 online). The flagellum is essentially composed of three substructures, a long helical filament that is connected via the flexible hook to the complex basal body, which is located in the cell wall. All genes encoding the structural proteins of the flagellum were found in the *G. oxydans* genome. These results are in accordance with morphological features of this organism that indicate the presence of peritrichous flagella<sup>14</sup>.

With respect to cell wall assembly, all essential genes were identified that are involved in biosynthesis of lipopolysaccharides and

peptidoglycan. The organism contains a gene cluster comprising 14 genes (GOX1477–1490) that is probably required for the production of a cell surface capsular polysaccharide. Furthermore, we identified all genes necessary for the biosynthesis of squalen and hopanoids (see Supplementary Table 1 online).

### Plasmids and phages

None of the five plasmids show homology to known plasmids of other *G. oxydans* strains (e.g., pAG5 (ref. 15) and pGO128 (M. Sievers, direct submission to GenBank; Gen ID: NC\_003374)). However, pGOX5 revealed substantial similarities over 1.5 kb with plasmid pJK21 from *Acetobacter europaeus*<sup>16</sup>. A BLAST search with the RepA proteins of plasmids pGOX1, pGOX2, pGOX3 and pGOX5 indicated high homologies to the corresponding proteins encoded by genes of plasmids from other  $\alpha$  proteobacteria.

Around 70% of the ORFs of the *G. oxydans* 621H plasmids encode hypothetical proteins. Genes encoding proteins with putative functions include a DNA helicase II (*umuD*), a restriction/modification system, a C<sub>4</sub>-dicarboxylate transporter, a heavy metal resistance system, proteins involved in plasmid replication and two alcohol dehydrogenases with unknown substrate spectra (GOX2594 and GOX2684; for additional information see Supplementary Table 3 online). Additionally, the megaplasmid contains the genetic information for DNA transfer via conjugation (Supplementary Table 1 online). We detected two regions (GOX2318–2357 and GOX1211–1226) with ORFs that show similarity to known phage genes. *G. oxydans* strains have been reported to contain dsDNA phages with contractile tails (Myoviridae, e.g., phage GW6210) and noncontractile phages with short tails (Podoviridae, e.g., JW2040)<sup>17</sup>. They were identified by electron microscopy and DNA sequences are not available.

### Respiratory chain

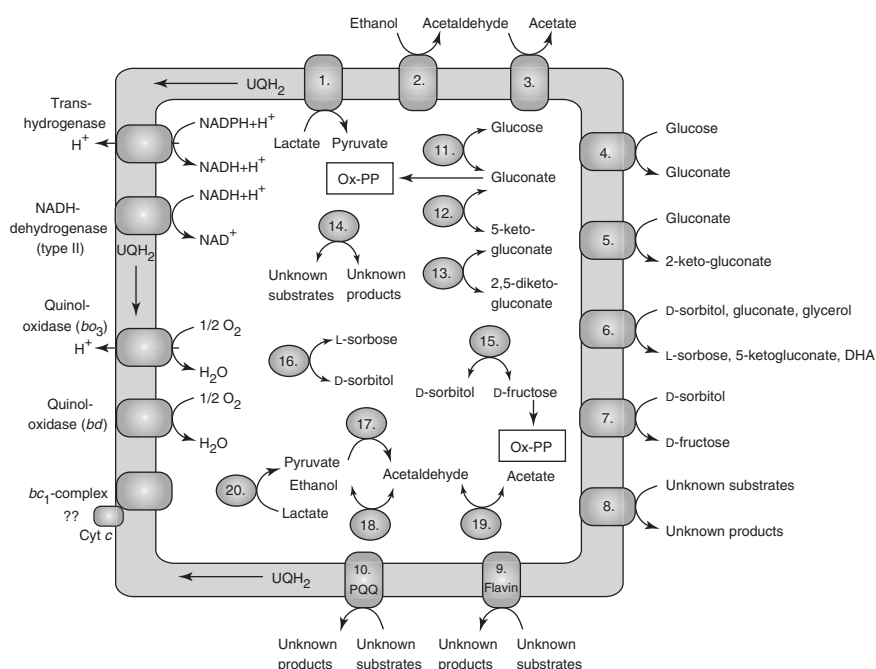
The knowledge of the DNA sequence forms the basis for the elucidation of the respiratory chain as a key element of the oxidative metabolism in *G. oxydans*. The core system is rather simple, consisting of a non-proton-translocating NADH:ubiquinone oxidoreductase and two quinol oxidases of the *bo*<sub>3</sub> and *bd* type, respectively (Fig. 1). The former protein has been characterized biochemically<sup>18</sup>. The identification of a *bd*-type quinol oxidase explains previous findings that *G. oxydans* contains a cyanide-insensitive terminal oxidase that is produced at low pH values<sup>19</sup>. The organism lacks a proton-translocating NADH: ubiquinone oxidoreductase (complex I) and a cytochrome *c* oxidase (complex IV). Therefore, the ability to translocate protons in the course of redox reactions is rather limited (Fig. 1). Genes encoding a ubiquinol:cytochrome *c* oxidoreductase (*bc*<sub>1</sub> complex) were identified. However, the function of the corresponding protein complex is unclear because reduced cytochrome *c* cannot be reoxidized by complex IV, which is missing in this organism. The electrochemical proton gradient is used to generate ATP via an F<sub>1</sub>F<sub>o</sub>-type ATP synthase. The corresponding genes are organized in two clusters, which are located at different positions on the chromosome. Cluster I (GOX1110–1113) encodes the hydrophobic membrane-bound subunits and cluster II (GOX1310–1314) contains the genes for the hydrophilic part of the enzyme (see Supplementary Table 1 online).

### Oxidoreductases

In addition to the standard respiratory complexes mentioned above, *G. oxydans* possesses a large and diverse set of genes encoding membrane-bound dehydrogenases that channel electrons into the respiratory chain (Fig. 1) as well as many intracellular oxidoreductases.

**Figure 1** Respiratory chain and sugar/alcohol metabolism in *G. oxydans* 621H. (1) Membrane-bound lactate dehydrogenase (GOX1253). (2) PQQ-dependent alcohol dehydrogenase (GOX1067–1068). (3) Membrane-bound acetaldehyde dehydrogenase (GOX0585–0587). (4) PQQ-dependent glucose dehydrogenase (GOX0265). (5) Membrane-bound gluconate-2-dehydrogenase (GOX1230–1232). (6) Glycerol/sorbitol dehydrogenase (SldAB; GOX854–855). (7) Sorbitol dehydrogenase (fructose forming; GOX2094–2097). (8) Uncharacterized membrane-bound oxidoreductases (see **Supplementary Table 3** online). (9) Uncharacterized flavin-containing oxidoreductases (see **Supplementary Table 3** online). (10) Uncharacterized PQQ-containing oxidoreductases (GOX0516, GOX1441, GOX1857, GOX1969). (11) Soluble glucose dehydrogenase (GOX2015). (12) Soluble gluconate dehydrogenase (GOX2187). (13) 2,5-diketogluconate reductase (GOX0644). (14) Uncharacterized cytoplasmic oxidoreductases (see **Supplementary Table 3**). (15) Soluble sorbitol dehydrogenase (fructose forming; GOX1432). (16) L-sorbose reductase (GOX0849). (17) Pyruvate decarboxylase (GOX1081). (18) Soluble acetaldehyde dehydrogenase (GOX2018). (19) Soluble alcohol dehydrogenase (GOX0313). OxPP indicates that the corresponding intermediates are further metabolized by the oxidative pentose-phosphate pathway.

All soluble oxidoreductases are NAD(P) dependent and catalyze reversible reactions. The membrane-bound dehydrogenases transfer electrons to ubiquinone (UQH<sub>2</sub>). The left part of the figure shows the components of the respiratory chain in *G. oxydans* 621H: Membrane-bound transhydrogenase (GOX0310–0312); NADH-dehydrogenase (type II), nonproton translocating NADH:ubiquinone oxidoreductase (GOX1675); quinol oxidase (*bo*<sub>3</sub>), cytochrome *bo*<sub>3</sub> ubiquinol oxidase (GOX1911–1914); quinol oxidase (*bd*), cytochrome *bd* ubiquinol oxidase (GOX0278–0279); *bc*<sub>1</sub> complex, ubiquinol:cytochrome *c* oxidoreductase (GOX0565–0567), *Cyt c*, cytochrome *c* (GOX0564). The question marks indicate that the function of the *bc*<sub>1</sub> complex is not clear yet. The function of the enzymes was verified by BLAST and motif searches of the corresponding ORFs against public databases.



The left part of the figure shows the components of the respiratory chain in *G. oxydans* 621H: Membrane-bound transhydrogenase (GOX0310–0312); NADH-dehydrogenase (type II), nonproton translocating NADH:ubiquinone oxidoreductase (GOX1675); quinol oxidase (*bo*<sub>3</sub>), cytochrome *bo*<sub>3</sub> ubiquinol oxidase (GOX1911–1914); quinol oxidase (*bd*), cytochrome *bd* ubiquinol oxidase (GOX0278–0279); *bc*<sub>1</sub> complex, ubiquinol:cytochrome *c* oxidoreductase (GOX0565–0567), *Cyt c*, cytochrome *c* (GOX0564). The question marks indicate that the function of the *bc*<sub>1</sub> complex is not clear yet. The function of the enzymes was verified by BLAST and motif searches of the corresponding ORFs against public databases.

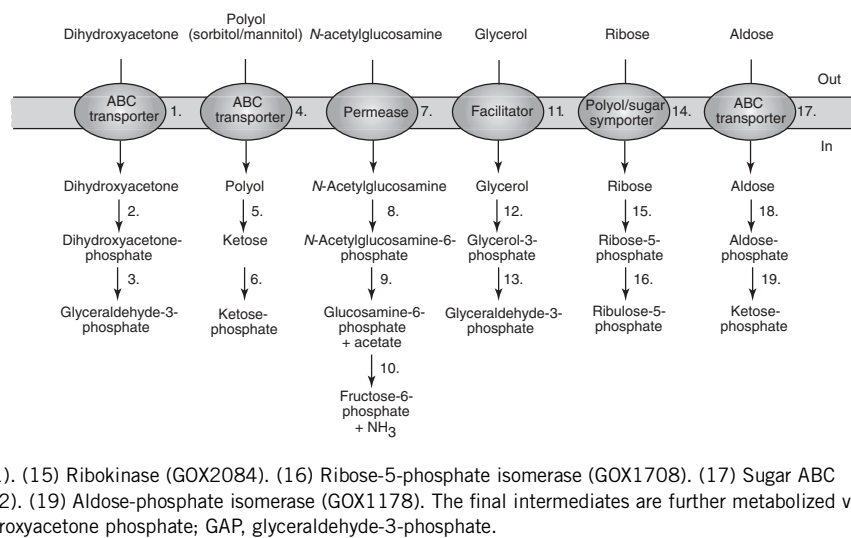
One striking feature of the *G. oxydans* genome is that 75 ORFs were identified that encode putative dehydrogenases/oxidoreductases of unknown functions (**Fig. 1**; **Supplementary Table 3** online). Among these ORFs, 23 were predicted to contain at least one transmembrane helix and four proteins that are probably located in the periplasm. The organism contains 15 putative oxidoreductases of the short chain dehydrogenase/reductase family that catalyze the reversible oxidation of alcohols to aldehydes with the concomitant reduction of NAD<sup>+</sup>. We identified four novel pyrroloquinoline quinone (PQQ)-dependent dehydrogenases of which three are membrane-bound with five transmembrane helices in the N-terminal domain. The C-terminal domain faces the outside and comprises the PQQ binding site and the catalytic center. The remaining protein represents a periplasmic PQQ enzyme. Five putative oxidoreductases belong to the zinc-binding dehydrogenase family. These proteins use NADPH<sup>+</sup> or NAD<sup>+</sup> and oxidize alcohols or reduce aldehydes. Six putative soluble oxidoreductases belong to the aldo-keto reductase family that comprise a number of related monomeric NADPH-dependent oxidoreductases. Furthermore, we identified four FAD-linked oxidases, three of which are referred to as putative D-lactate dehydrogenases. Two of these enzymes (GOX1170 and GOX2071) are predicted to be located in the cytoplasm of *G. oxydans* and are homologs of the corresponding mitochondrial protein from *Saccharomyces cerevisiae*<sup>20</sup>. The third protein (GOX1253) is related to the respiratory lactate dehydrogenase from *Escherichia coli*<sup>21</sup> and contains one transmembrane helix, indicating that it is membrane bound and located at the inner face of the cytoplasmic membrane. In addition, several other proteins were found that belong to different families of flavin-containing oxidoreductases (for a

detailed description of all putative oxidoreductases, see **Supplementary Table 3** online). These findings may lead to the development of new strategies for the employment of *G. oxydans* in a much greater variety of incomplete oxidations. Future studies will indicate the substrate spectra and the catalytic activities of the uncharacterized dehydrogenases in *G. oxydans*.

To prove that the genes encoding uncharacterized oxidoreductases are transcribed, we analyzed the expression of a selected subset by real time RT-PCR. This approach included genes encoding enzymes with known function as well as proteins that have been identified and annotated because of bioinformatic evidence. The experiments showed that all genes tested were expressed during growth on glucose, some of them in relatively high amounts. The results indicated that the corresponding oxidoreductases were produced and had a metabolic function in *G. oxydans* (**Supplementary Fig. 2** online).

A number of ORFs could be assigned to known oxidoreductases, for example, those containing the cofactor PQQ (**Fig. 1**). The alcohol dehydrogenase (involved in acetate formation), glucose dehydrogenase (involved in gluconate formation) and glycerol/sorbitol dehydrogenase (SldAB) belong to this group of quinoproteins<sup>22</sup>. The latter enzyme is the major polyol dehydrogenase in *G. oxydans* that exhibits a broad substrate specificity<sup>23</sup>. It catalyzes the oxidation of D-sorbitol, gluconate and glycerol, thereby producing L-sorbose, 5-ketogluconate and dihydroxyacetone, all of which are of great biotechnological interest<sup>24–26</sup>. Other proteins like gluconate-2-dehydrogenase and sorbitol dehydrogenase (SldSLC) contain flavins as prosthetic groups (**Fig. 1**). The latter protein catalyzes D-sorbitol oxidation to D-fructose and can be used for the biotechnological production of this sweetener<sup>5</sup>.

**Figure 2** Reconstruction of pathways for the utilization of polyols, sugars and sugar derivatives in *G. oxydans*. The enzymes catalyzing the reactions and the corresponding ORFs are indicated as follows. (1) Sugar ABC transporter (GOX2219–2221). (2) Dihydroxyacetone kinase (GOX2222). (3) Triosephosphate isomerase (GOX2217). (4) Polyol ABC transporter (GOX2182–2185). (5) Polyol dehydrogenase (GOX2181). (6) Ribulokinase (GOX2186). (7) Sugar permease (GOX1474). (8) N-acetylglucosamine kinase (GOX1475). (9) N-acetylglucosamine-6-phosphate deacetylase (GOX1473). (10) Glucosamine-6-phosphate deaminase (GOX1470). (11) Glycerol uptake facilitator protein (GOX2089). (12) Glycerol kinase (GOX2090). (13) Glycerol-3-phosphate dehydrogenase (GOX2088). (14) Polyol/sugar symporter (GOX0649, GOX0925, GOX1047, GOX2231). (15) Ribokinase (GOX2084). (16) Ribose-5-phosphate isomerase (GOX1708). (17) Sugar ABC transporter (GOX1179–1181). (18) Sugar kinase (1182). (19) Aldose-phosphate isomerase (GOX1178). The final intermediates are further metabolized via the oxidative pentose phosphate pathway. DHAP, dihydroxyacetone phosphate; GAP, glyceraldehyde-3-phosphate.



It has been shown that membrane-bound dehydrogenases transfer electrons to ubiquinone, which functions as electron donor for the quinol oxidases<sup>27</sup>. The active sites of the dehydrogenases are oriented towards the periplasm (Fig. 1). Thus, substances that are used as energy sources can be oxidized in the periplasmic space without the need to enter the cytoplasm<sup>2</sup>. The products of the oxidations are easily released into the medium via porins in the outer membrane of the Gram-negative bacteria. It has been shown that membrane-bound dehydrogenases are responsible for the rapid oxidation of biotechnologically important substrates<sup>27,28</sup>.

In addition to the membrane-bound dehydrogenases, there is an alternative route for the oxidation of sugars, alcohols and polyols in *G. oxydans*<sup>1,27</sup>. The second set of enzymes is located in the cytoplasm, indicating that the uptake of their substrates into the cell is required (Fig. 1). These proteins catalyze reversible reactions and are NAD(P)<sup>+</sup> dependent. The substrates are oxidized and the resulting intermediates are phosphorylated and further metabolized via the pentose phosphate pathway. The soluble, NAD(P)<sup>+</sup>-dependent enzymes are believed to participate in the synthesis of biosynthetic precursors and are obviously involved in the maintenance of cells in the stationary growth phase<sup>27</sup>. A close inspection of the genome sequence indicated that *G. oxydans* possesses several genes encoding enzymes of this class for the oxidation of glucose, gluconate, ketogluconate, ethanol and acetaldehyde (Fig. 1). In addition, three proteins were identified that may play a critical role in biotechnological applications for the synthesis of L-sorbose (NADPH-dependent sorbose reductase and fructose-forming sorbitol dehydrogenase) and xylitol (xylitol dehydrogenase)<sup>29–31</sup>.

### Intermediary metabolism

All genes encoding the enzymes of the oxidative pentose-phosphate and the Entner-Doudoroff pathways were identified. The oxidative pentose-phosphate pathway is thought to be the most important route for phosphorylative breakdown of sugars and polyols to CO<sub>2</sub> (ref. 32). It is noteworthy that the aforementioned soluble dehydrogenases as well as the glucose-6-phosphate and the 6-phosphogluconate dehydrogenase involved in these pathways produce NADPH + H<sup>+</sup> in the course of their reactions. It is still a matter of debate how this cofactor is oxidized. Interestingly, the genome data revealed the presence of a proton-translocating nicotinamide

nucleotide transhydrogenase in *G. oxydans* 621H (Fig. 1). The amino acid sequences of the three subunits are highly homologous to the membrane-bound transhydrogenase from *Rhodospirillum rubrum* (*pntAA*, *pntAB* and *pntB*)<sup>33</sup>. Transhydrogenase, in animal mitochondria and bacteria, couples hydride transfer between NADH + H<sup>+</sup> and NADP<sup>+</sup> to proton translocation across a membrane<sup>34</sup>. Hence, the enzyme might have two functions in *G. oxydans*: first, the oxidation of NADPH derived from intermediary metabolism, and second, the translocation of protons across the membrane thereby contributing to maintenance of the electrochemical proton gradient (Fig. 1).

Within the genome no ORF could be assigned that potentially encodes a phosphoenolpyruvate synthase or other phosphoenolpyruvate-synthesizing enzymes indicating that *G. oxydans* cannot produce C<sub>6</sub>-sugars via gluconeogenesis. Hence, the formation of glucose, for example, from pentoses or glycerol must depend on the oxidative pentose phosphate pathway. In accordance with the literature<sup>35</sup>, the citrate cycle is incomplete because genes encoding the succinate dehydrogenase are missing.

We predict that *G. oxydans* has the capability to take up and to channel several polyols, sugars and sugar derivatives into the oxidative pentose phosphate pathway (Fig. 2). In most cases the substrates are first phosphorylated by specific kinases and are further degraded by dehydrogenases and isomerases. Further genome analysis showed that *G. oxydans* contains metabolic pathways for the *de novo* synthesis of all nucleotides, amino acids, phospholipids and most vitamins. Proteins for the synthesis of the important cofactor PQQ are encoded by the *pqqA-E* operon (GOX983–987), which is highly homologous to the one found in *G. oxydans* IFO 3293 (ref. 36). The organism contains several cytochrome *c*-type proteins. A close inspection of the genome revealed the presence of all genes for the synthesis and transport of heme *c* and the maturation of cytochrome *c* (GOX1648–1652) (Supplementary Table 1 online)<sup>37</sup>. Ammonia is taken up by a specific transporter (GOX743) and is introduced into biosynthetic pathways by the catalytic activity of glutamate synthase (GOX1851–1852) and glutamine synthetase (GOX1121). We identified a gene encoding a pyruvate decarboxylase (GOX1081) indicating that *G. oxydans* can produce acetaldehyde from pyruvate. The former intermediate is then oxidized to acetate by a soluble acetaldehyde dehydrogenase (GOX2018)<sup>38</sup>.



## Transporters

*G. oxydans* possesses various transporters for the uptake of substrates and ions. We identified 29 ABC transporters, 5 symporters and 37 other permeases. Among them were those that can transport metal ions,  $\text{NH}_4^+$ , phosphate, sulfate, amino acids, purines, pyrimidines, di- and oligopeptides, sugars (e.g., ribose, galactose), polyols (e.g., mannitol and sorbitol) and sugar acids (gluconate and glucarate). Glycerol is taken up by a facilitator protein (Fig. 2). In addition, some components of a phosphotransferase system were determined. However, the sugar-specific proteins (EIIB, EIIC) were not identified (see Supplementary Table 1 online).

## DISCUSSION

The bacteria belonging to the genus *Gluconobacter* not only exhibit an extraordinary uniqueness in their biochemistry with respect to the process of incomplete oxidation but also in their growth behavior and response to extreme culture conditions<sup>2</sup>. On the basis of the genomic sequence of *G. oxydans* we can now more extensively describe the process of incomplete oxidation and the physiology of organisms carrying out this process. The organisms contain many membrane-bound dehydrogenases that are part of their strategy to thrive and to survive in nutrient-rich environments. The enzymes rapidly form sugars or sugar acids that are difficult to assimilate for most microorganisms, whereas *G. oxydans* can easily take advantage of these substrates.

The oxidized compounds are taken up and reduced in the cytoplasm, the reactions being catalyzed by a soluble set of oxidoreductases. Therefore, this deposit and withdrawal system of sugars and sugar alcohols is a clever strategy adopted by *G. oxydans* to survive in mixed microbial populations. Furthermore, incomplete oxidation of glucose and other aldoses, which are abundant in the natural habitats of *G. oxydans*, lead to the formation of sugar acids and to a decrease in the pH value, thereby preventing propagation of many other microorganisms.

The respiratory chain is designed to accelerate this process because proton-translocating abilities are limited. This prevents an increase in the electrochemical membrane potential that would otherwise lead to the inhibition of membrane-bound redox reactions. However, the low energy-transducing efficiency results in very low growth yields<sup>39</sup>. The inability to degrade glucose and other sugars via the Embden-Meyerhof pathway and the incomplete citrate cycle contribute to the inadequate utilization of the substrates. In summary, *G. oxydans* reveals an extreme adaptation to its nutrient rich habitats by out-competing other microorganisms. Furthermore, the unique metabolism makes it an ideal organism for microbial process development.

## METHODS

The *G. oxydans* 621H (DSMZ 2343) genome was determined by a whole-genome shotgun approach using plasmid and cosmid libraries as described previously<sup>40</sup>. The raw sequencing data were obtained from Integrated Genomics Inc. All generated sequences (53,000) were assembled into contigs, and misassembled regions caused by repetitive sequences were solved as described<sup>40</sup>. Closure of sequencing gaps was accomplished through primer-walking using plasmid, PCR products and cosmid templates. Sequences from cosmids served for verification of the orientation, order and linkage of the contigs. Initial ORF prediction was accomplished automatically using the program YACOP<sup>41</sup>. The ERGO software package<sup>42</sup> (Integrated Genomics, Inc., <http://www.integratedgenomics.com/>) was used for automatic annotation. All annotations were inspected manually through searches against PFAM, PROSITE, PRODOM and COGS databases, in addition to the BLASTP versus GenBank/EMBL and SWISSPROT databases. Putative frameshifts were checked and corrected manually.

The complete *G. oxydans* genome sequence has been deposited in GenBank (<http://www.ncbi.nlm.nih.gov>, accession nos. CP000009 (chromosome), CP000004 (plasmid pGOX1), CP000005 (plasmid pGOX2), CP000006 (plasmid pGOX3), CP000007 (plasmid pGOX4), CP000008 (plasmid5) and the Göttingen Genomics Laboratory database (<http://www.g2l.bio.uni-goettingen.de/>).

Note: Supplementary information is available on the Nature Biotechnology website.

## ACKNOWLEDGMENTS

This work was supported by grants from the Lower Saxony ministry of Science and Culture to the Goettingen Genomics Laboratory and by funds of the Competence Network Goettingen "Genome Research on Bacteria" financed by the German Federal Ministry of Education and Research (BMBF).

## COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 23 July; accepted 11 November 2004

Published online at <http://www.nature.com/naturebiotechnology/>

- De Ley, J., Swings, J. & Gossele, F. The genus *Gluconobacter*. in *Bergey's Manual of Systematic Bacteriology*, vol. 1 (eds. Krieg, N.R. & Holt, J.G.) 267–278 (Williams & Wilkins Co., Baltimore, 1984).
- Deppenmeier, U., Hoffmeister, M. & Prust, C. Biochemistry and biotechnological applications of *Gluconobacter* strains. *Appl. Microbiol. Biotechnol.* **60**, 233–242 (2002).
- Reichstein, T. & Grüssner, A. Eine ergiebige Synthese der 1-Ascorbinsäure (C-Vitamin). *Helv. Chim. Acta* **17**, 160–164 (1934).
- Schedel, M. Regioselective oxidation of aminosorbitol with *Gluconobacter oxydans*, a key reaction in the industrial synthesis of 1-deoxynojirimycin. in *Biotechnology*, vol. 8b *Biotransformations I*. (ed. Kelly, D.R.) 296–308 (Wiley-VCH, Weinheim, 2000).
- Adachi, O. *et al.* New developments in oxidative fermentation. *Appl. Microbiol. Biotechnol.* **60**, 643–653 (2003).
- Rabenhorst, J., Gatfield, I. & Hilmer, J.M. Natural, aliphatic and thiocarboxylic acids obtainable by fermentation and a microorganism therefore. Patent EP1078990 (2001).
- Macaulay, S., McNeil, B. & Harvey, L.M. The genus *Gluconobacter* and its applications in biotechnology. *Crit. Rev. Biotechnol.* **21**, 1–25 (2001).
- Gupta, A., Singh, V.K., Qazi, G.N. & Kumar, A. *Gluconobacter oxydans*: Its biotechnological applications. *J. Mol. Microbiol. Biotechnol.* **3**, 445–456 (2001).
- Batley, A.S. & Schaffner, D.W. Modelling bacterial spoilage in cold-filled ready to drink beverages by *Acinetobacter calcoaceticus* and *Gluconobacter oxydans*. *J. Appl. Microbiol.* **91**, 237–247 (2001).
- Grigoriev, A. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* **26**, 2286–2290 (1998).
- Kondo, K. & Horinouchi, S. Characterization of an insertion sequence, IS12528, from *Gluconobacter suboxydans*. *Appl. Environ. Microbiol.* **63**, 1139–1142 (1997).
- Mahillon, J. & Chandler, M. Insertion sequences. *Microbiol. Mol. Biol. Rev.* **62**, 725–774 (1998).
- Bren, A. & Eisenbach, M. How signals are heard during bacterial chemotaxis: protein-protein interactions in sensory signal propagation. *J. Bacteriol.* **182**, 6865–6873 (2000).
- Asai, T., Iizuka, H. & Komagata, K. The flagellation and taxonomy of genera *Gluconobacter* and *Acetobacter* with reference to the existence of intermediate strains. *J. Gen. Appl. Microbiol.* **10**, 95–125 (1964).
- Tonouchi, N., Sugiyama, M. & Yokozeke, K. Construction of a vector plasmid for use in *Gluconobacter oxydans*. *Biosci. Biotechnol. Biochem.* **67**, 211–213 (2003).
- Trcek, J., Raspor, P. & Teuber, M. Molecular identification of *Acetobacter* isolates from submerged vinegar production, sequence analysis of plasmid pJK2-1 and application in the development of a cloning vector. *Appl. Microbiol. Biotechnol.* **53**, 289–295 (2000).
- Robakis, N. *et al.* Isolation and characterization of two phages for *Gluconobacter oxydans*. *J. Gen. Microbiol.* **131**, 2467–2474 (1985).
- Matsushita, K., Shinagawa, E., Adachi, O. & Ameyama, M. Purification, characterization and reconstitution of cytochrome  $\alpha$ -type oxidase from *Gluconobacter suboxydans*. *Biochim. Biophys. Acta* **894**, 304–312 (1987).
- Matsushita, K., Nagatani, Y., Shinagawa, E., Adachi, O. & Ameyama, M. Effect of extracellular pH on the respiratory chain and energetics of *Gluconobacter suboxydans*. *Agric. Biol. Chem.* **53**, 2895–2902 (1989).
- Lodi, T. & Ferrero, I. Isolation of the DLD gene of *Saccharomyces cerevisiae* encoding the mitochondrial enzyme D-lactate ferricytochrome c oxidoreductase. *Mol. Gen. Genet.* **238**, 315–324 (1993).
- Rule, G.S., Pratt, E.A., Chin, C.C., Wold, F. & Ho, C. Overproduction and nucleotide sequence of the respiratory D-lactate dehydrogenase of *Escherichia coli*. *J. Bacteriol.* **161**, 1059–1068 (1985).
- Adachi, O. *et al.* New quinoproteins in oxidative fermentation. *Biochim. Biophys. Acta* **164**, 10–17 (2003).
- Matsushita, K. *et al.* 5-Keto-D-gluconate production is catalysed by a quinoprotein glycerol dehydrogenase, major polyol dehydrogenase, in *Gluconobacter* species. *Appl. Environ. Microbiol.* **69**, 1959–1966 (2003).

24. Hoshino, T., Sugisawa, T., Shinjoh, M., Tomiyama, N. & Miyazaki, T. Membrane-bound D-sorbitol dehydrogenase of *Gluconobacter suboxydans* IFO 3255 – enzymatic and genetic characterization. *Biochim. Biophys. Acta* **1647**, 278–288 (2003).
25. Herrmann, U., Merfort, M., Jeude, M., Bringer-Meyer, S. & Sahm, H. Biotransformation of glucose to 5-keto-D-gluconic acid by recombinant *Gluconobacter oxydans* DSM 2343. *Appl. Microbiol. Biotechnol.* **64**, 86–90 (2004).
26. Hekmat, D., Bauer, R. & Fricke, J. Optimization of the microbial synthesis of dihydroxyacetone from glycerol with *Gluconobacter oxydans*. *Bioprocess. Biosyst. Eng.* **26**, 109–116 (2003).
27. Matsushita, K., Toyama, H. & Adachi, O. Respiratory chains and bioenergetics of acetic acid bacteria. *Adv. Microb. Physiol.* **36**, 247–301 (1994).
28. Pronk, J.T., Levering, P.R., Olijve, W. & Van Dijken, J.P. Role of NADP-dependent and quinoprotein glucose dehydrogenases in gluconic acid production by *Gluconobacter oxydans*. *Enzyme Microb. Technol.* **11**, 160–164 (1989).
29. Shinjoh, M., Tazoe, M. & Hoshino, T. NADPH-dependent L-sorbose reductase is responsible for L-sorbose assimilation in *Gluconobacter suboxydans* IFO 3291. *J. Bacteriol.* **184**, 861–863 (2002).
30. Shibata, T. *et al.* Cloning of a gene for D-sorbitol dehydrogenase from *Gluconobacter oxydans* G624 and expression of the gene in *Pseudomonas putida* IFO3738. *J. Biosci. Bioeng.* **89**, 463–468 (2000).
31. Sugiyama, M., Suzuki, S., Tonouchi, N. & Yokozeki, K. Cloning of the xylitol dehydrogenase from *Gluconobacter oxydans* and improved production of xylitol from arabinol. *Biosci. Biotechnol. Biochem.* **67**, 584–591 (2003).
32. Asai, T. *Acetic Acid Bacteria: Classification and Biochemical Activities* (University of Tokyo Press, Tokyo, 1968).
33. Williams, R., Cotton, N.P., Thomas, C.M. & Jackson, J.B. Cloning and sequencing of the genes for the proton-translocating nicotinamide nucleotide transhydrogenase from *Rhodospirillum rubrum* and the implications for the domain structure of the enzyme. *Microbiology* **140**, 1595–1604 (1994).
34. Cotton, N.P.J., White, S.A., Peake, S.J., McSweeney, S. & Baz Jackson, J. The crystal structure of an asymmetric complex of the two nucleotide binding components of proton-translocating transhydrogenase. *Structure* **9**, 165–176 (2001).
35. Greenfield, S. & Claus, G.W. Non-functional tricarboxylic acid cycle and the mechanism of glutamate biosynthesis in *Acetobacter suboxydans*. *J. Bacteriol.* **112**, 1295–1301 (1972).
36. Felder, M. *et al.* The pyrroloquinoline quinone synthesis genes of *Gluconobacter oxydans*. *FEMS Microbiol. Lett.* **193**, 231–236 (2000).
37. Thöny-Meyer, L. A heme chaperone for cytochrome *c* biosynthesis. *Biochemistry* **42**, 13100–13105 (2003).
38. Chandra Raj, K., Ingram, L.O. & Maupin-Furlow, J.A. Pyruvate decarboxylase: a key enzyme for the oxidative metabolism of lactic acid by *Acetobacter pasteurianus*. *Arch. Microbiol.* **176**, 443–451 (2001).
39. Olijve, W. & Kok, J.J. An analysis of the growth of *Gluconobacter oxydans* in chemostat cultures. *Arch. Microbiol.* **121**, 291–297 (1979).
40. Deppenmeier, U. *et al.* The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between Bacteria and Archaea. *J. Mol. Microbiol. Biotechnol.* **4**, 453–461 (2002).
41. Tech, M. & Merkl, R. YACOP: enhanced gene prediction obtained by a combination of existing methods. *In Silico Biol.* **3**, 441–451 (2003).
42. Overbeek, R. *et al.* The ERGO™ genome analysis and discovery system. *Nucleic Acids Res.* **31**, 164–171 (2003).