

**Entschlüsselung des Genoms von
Gluconobacter oxydans 621H – einem Bakterium
von industriellem Interesse**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität zu Göttingen

vorgelegt von
Christina Prust
aus Göttingen

Göttingen, 2004

D7

Referent: Prof. Dr. G. Gottschalk

Korreferent: Prof. Dr. U. Deppenmeier

Tag der mündlichen Prüfung: 29. Juni 2004

für meinen Mann Jörg

INHALTSVERZEICHNIS

ABKÜRZUNGSVERZEICHNIS

1	EINLEITUNG	1
2	MATERIAL UND METHODEN	5
2.1	Organismen und Plasmide	5
2.2	Nährmedien, Puffer und Stammlösungen	6
2.2.1	Komplexmedien	6
2.2.2	Antibiotika und Medienzusätze	7
2.3	Kultivierung und Stammhaltung von Mikroorganismen	7
2.3.1	Zellanzucht	7
2.3.2	Stammhaltung	7
2.3.3	Reinheitskontrolle	8
2.4	Messung der optischen Dichte	8
2.5	Standardtechniken für das Arbeiten mit Nukleinsäuren	8
2.5.1	Behandlung von Geräten und Lösungen	8
2.5.2	Reinigung und Konzentrierung von Nukleinsäuren	9
2.5.2.1	Phenol/Chloroform-Extraktion	9
2.5.2.2	Fällung von Nukleinsäuren	9
2.5.3	Konzentrationsbestimmung und Reinheitskontrolle von DNA	9
2.5.4	Standard-Agarosegelelektrophorese	10
2.5.5	Größenbestimmung von Nukleinsäuren	11
2.6	Isolierung von Nukleinsäuren	12
2.6.1	Isolierung genomischer DNA mit dem AquaPure GenomicDNA Kit	12

2.6.2	Isolierung genomischer DNA mit CTAB	13
2.6.3	Isolierung von Plasmid-DNA	14
2.6.4	Isolierung von DNA-Fragmenten aus Agarosegelen	14
2.6.5	Aufreinigung von PCR-Produkten	15
2.6.6	Phenol-Chloroform-Extraktion	15
2.7	Enzymatische Modifikation von DNA	16
2.7.1	Spaltung von DNA durch Restriktionsendonukleasen	16
2.7.2	Dephosphorylierung von DNA-Fragmenten	16
2.7.3	Herstellung von stumpfen Enden	17
2.7.4	Ligation von DNA Fragmenten	17
2.8	Übertragung von DNA in <i>E. coli</i> und Selektion rekombinanter Klone	17
2.8.1	Herstellung von transformationskompetenten <i>E. coli</i> -Zellen (INOUE et al., 1990)	17
2.8.2	Transformation von <i>E. coli</i> DH5 α durch Hitzeschock	18
2.8.3	Transformation von <i>E. coli</i> durch Elektroporation	18
2.8.4	Selektion rekombinanter Klone durch den Blau-Weiß-Test	19
2.9	<i>In vitro</i>-Amplifikation von DNA durch Polymerasekettenreaktion (PCR)	20
2.10	Sequenzierung von DNA	21
2.11	Die Prozessierung von Rohsequenzen	22
2.11.1	Einführung in das Staden Package	22
2.11.2.	Das Pregap4 Interface	22
2.11.3	Gap4	23
2.11.3.1	Der <i>Contig Editor</i>	24
2.11.3.2	Das <i>Template Display</i>	24
2.12	Methoden zur Überwindung von Sequenzlücken in einem Genomsequenzierungsprojekt	25

2.12.1.	Identifikation von Contigüberlappungen mittels Gap4 oder BLASTN	26
2.12.2	Lückenschluss durch <i>Walkings</i> auf Plasmiden	26
2.12.3	Lückenschluss mit Hilfe von PCR	27
2.12.3.1	Auf der Cosmid-Genbank basierende PCR	27
2.12.3.2	Kombinatorische PCR	27
2.13	Auffinden repetiver Sequenzen	28
2.13.1	Identifikation repetiver Sequenzen mit Hilfe von und Gap4	28
2.13.2	Auffinden repetitiver Sequenzen mit Hilfe von RepVis und BLASTN	29
2.14	Sequenzanalyse	29
2.14.1	ORF-Vorhersage	30
2.14.2	Genomweite Annotation mit ERGO	31
2.15	Bezugsquellen für Biochemikalien und Enzyme	33
3	EXPERIMENTE UND ERGEBNISSE	35
3.1	Vorarbeiten und Überblick über die Strategie des Sequenzierungsprojekts	35
3.2	Editieren der Rohsequenzen	37
3.3	Lückenschluss	41
3.3.1	Lückenschluss durch Walkings und BLASTN	42
3.3.2	Erstellen eines „Supercontigs“	43
3.3.3	Kombinatorische Ansätze	46
3.3.4	Abschluss der Sequenzierung	47
3.4	Identifizierung von Plasmiden	50
3.4.1	Auffinden der Plasmide mit Hilfe von Gap4	50
3.4.2	Isolierung der identifizierten Plasmide	51

3.4.3	Analyse der Plasmide	54
3.4.3.1	Analyse von pGOX1	54
3.4.3.2	Analyse von pGOX2	57
3.4.3.3	Analyse von pGOX3	58
3.4.3.4	Analyse von pGOX4	59
3.4.3.5	Analyse von pGOX5	59
3.5.	Analyse des Chromosoms von <i>G. oxydans</i>	60
3.5.1	Allgemeine Merkmale des Chromosoms	60
3.5.1.1	ORF-Vorhersage und Übersicht über die Annotation	62
3.5.2	Identifizierung des Replikationsursprungs	64
3.6	Rekonstruktion des Stoffwechsels	66
3.6.1	Transportsysteme in <i>G. oxydans</i>	67
3.6.1.1	Erleichterte Diffusion	68
3.6.1.2	ABC-Transporter	68
3.6.1.3	Symporter und Antiporter	69
3.6.1.4	Permeasen	70
3.6.2	Substratverwertung	71
3.6.2.1	Glykolyse und Gluconeogenese	72
3.6.2.2	Pentosephosphat-Weg	74
3.6.2.3	Entner-Doudoroff-Weg (KDPG-Weg)	77
3.6.2.4	Zitratzyklus	79
3.6.3	Biosynthese der Aminosäuren	81
3.6.3.1	Biosynthese von Histidin	81
3.6.3.2	Biosynthese der aromatischen Aminosäuren Phenylalanin, Tyrosin und Tryptophan	84
3.6.3.3	Biosynthese von Serin, Glycin und Cystein	86
3.6.3.4	Biosynthese von Alanin, Leucin, Isoleucin und Valin	87
3.6.3.5	Biosynthese von Aspartat, Asparagin, Lysin, Threonin und Methionin in <i>G. oxydans</i>	89
3.6.3.6	Biosynthese von Glutamin, Glutamat, Prolin und Arginin in <i>G.</i> <i>oxydans</i>	91
3.6.4	Transhydrogenasen	92

3.6.4.1	Analyse der Transhydrogenase aus <i>G. oxydans</i>	93
3.6.5	Membrangebundene Prozesse	97
3.6.5.1	Rekonstruktion der Atmungskette	97
3.6.5.2	Die ATP-Synthase	101
3.6.6	Dehydrogenasen	102
3.6.6.1	Pyrrlochinolin Chinon (PQQ)	107
3.7	Abschließende Bemerkungen	108
4	DISKUSSION	110
4.1	Mobile Elemente	110
4.2	Betrachtungen zur Energiekonservierung	114
4.3	Chinoproteine	119
4.3.1	Allgemeine Merkmale von Chinoproteinen	120
4.3.2	PQQ-abhängige Proteine aus <i>G. oxydans</i>	123
4.3.2.1	Die membrangebundene Alkohol-Dehydrogenase aus <i>G. oxydans</i>	124
4.3.2.2	Chinat-Dehydrogenase	126
4.3.3	Der Cofaktor PQQ	128
4.4	<i>G. oxydans</i> im Dienst der Biotechnologie	129
4.4.1	Essigsäureproduktion	130
4.4.2	Produktion von Vitamin C	131
4.4.3	Miglitol und 1-Deoxynojirimycin	136
4.4.4	Produktion von Dihydroxyaceton	138
4.4.5	Produktion von Xylitol	139
4.4.6	Produktion von Gluconat und Ketogluconaten durch <i>G. oxydans</i>	142
4.4.6.1	Membrangebundene Produktion von Ketogluconaten	142
4.4.6.2	Cytoplasmatische Produktion von Ketogluconaten	144
4.5	Ausblick	148

5	ZUSAMMENFASSUNG	149
6	LITERATURVERZEICHNIS	152

Abkürzungsverzeichnis

<i>A.</i>	<i>Acinetobacter</i>
°C	Grad Celsius
A	Adenin
Abb.	Abbildung
ABC	ATP-bindende Kasette, ATP-binding cassette
AICAR	5'-Phospho-Ribosyl-5-Amino-4-Imidazol-Carboxamid
Amp	Ampicillin
ATP	Adenosintriphosphat
BLAST	Basic Local Alignment Search Tool
BLASTN	BLAST auf Nukleotidebene
BLASTP	BLAST auf Proteinebene
bp	Basenpaar
bzw.	beziehungsweise
C	Cytosin
C-	Carboxy
ca.	circa
cm	Zentimeter
CoA	Coenzym A
Da	Dalton
DH	Dehydrogenase
DMSO	Dimethylsulfoxid
DNA	Desoxyribonukleinsäure
dNTP	Desoxyribonukleotid-5'-triphosphat
<i>E.</i>	<i>Escherichia</i>
ed(s).	Herausgeber
EDTA	Etylendiamintetraessigsäure
einschl.	einschließlich
et al.	et alteri (und andere)
Fa.	Firma
FAD	Flavinadenindinukleotid
G	Guanin
g	Gramm

<i>G.</i>	<i>Gluconobacter</i>
Gap	Genomasmblierungsprogramm, genome assembly program
H ₂ O _{dest.}	destilliertes Wasser
Hrsg.	Herausgeber
IPTG	Isopropyl-β-D-thiogalactopyranosid
IS	Insertionssequenz(en)
k	Kilo
kb	Kilobasenpaare
l	Liter
μ	Mikro (10 ⁶)
M	Molar (Mol pro Liter)
Mb	Megabasenpaare
min	Minute
mod.	modifiziert
N-	Amino-
NAD(P)	Nicotinadenindinukleotid(phosphat)
NCBI	National Center for Biotechnology Information
OD	Optische Dichte
ORF	“open reading frame”, offener Leserahmen
<i>P.</i>	<i>Pseudomonas</i>
PCR	Polymerasekettenreaktion
pH	negativer dekadischer Logarithmus der Protonenkonzentration
PQQ	Pyrrlochinolinchinon
PRPP	5-Phosphoribosyl-1-Pyrophosphat
<i>R.</i>	<i>Rhizobium</i>
RNA	Ribonukleinsäure
rRNA	Ribosomale Ribonukleinsäure
RT	Raumtemperatur
s.	siehe
<i>S.</i>	<i>Sinorhizobium</i>
SDS	Natriumdodecylsulfat
sp.	Spezies
TAE	Tris-Acetat-EDTA
TE	Tris-EDTA

T _m	Schmelztemperatur
U	Unit (Einheit der Enzymaktivität)
u.a.	unter anderem
u.s.w.	und so weiter
v/v	Volumen pro Volumen
vgl.	vergleiche
w/v	Masse pro Volumen
X-Gal	5-Brom-4-chlor-3-indolyl- β -galactosid
z. B.	zum Beispiel

Aminosäuren

A	Ala	Alanin	M	Met	Methionin
C	Cys	Cystein	N	Asn	Asparagin
D	Asp	Aspartat	P	Pro	Prolin
E	Glu	Glutamat	Q	Gln	Glutamin
F	Phe	Phenylalanin	R	Arg	Arginin
G	Gly	Glycin	S	Ser	Serin
H	His	Histidin	T	Thr	Threonin
I	Ile	Isoleucin	V	Val	Valin
K	Lys	Lysin	W	Trp	Tryptophan
L	Leu	Leucin	Y	Tyr	Tyrosin

1 EINLEITUNG

Gluconobacter-Arten gehören, zusammen mit Vertretern der Gattungen *Acetobacter* und *Gluconacetobacter*, zu der Familie der *Acetobacteriaceae*. Sie unterscheiden sich von denen der Gattung *Acetobacter* dadurch, dass sie Acetat und Lactat nicht zu Kohlendioxid und Wasser oxidieren können (DE LEY et al., 1984). Während die Produktion von Essigsäure aus Ethanol vor allem durch Organismen der Gattung *Acetobacter* besonders effizient ist, dominieren *Gluconobacter* Spezies bei der Oxidation verschiedener Zucker. Essigsäurebakterien sind strikt aerobe, Gram-negative, stäbchenförmige Bakterien, welche der α -Subklasse der Proteobakterien zugeordnet werden. *Gluconobacter*-Arten sind durch polar angeordnete Flagellen schwach beweglich. Im Gegensatz dazu sind Vertreter der Gattung *Acetobacter* peritrich begeißelt (LEISINGER, 1965). In der Natur findet man *Gluconobacter* Spezies vor allem auf Früchten und in alkoholischen Getränken wie Bier und Wein. Dabei tragen sie durch die Säureproduktion auch zum Verderben der Getränke und Früchte bei. Die optimale Wachstumstemperatur beträgt 25-30 °C bei einem pH-Wert von 5,5-6. Viele *Gluconobacter*-Stämme sind jedoch auch in der Lage, in einem Komplexmedium bei einem konstanten pH-Wert von 2,5 zu wachsen (DEPPENMEIER et al., 2002; MACAULEY et al., 2001; GUPTA et al., 2001).

In der Regel oxidieren aerobe Mikroorganismen ihre Kohlenstoffquelle vollständig zu Kohlendioxid und Wasser. Bei diesem Abbauprozess werden sowohl Energie als auch Intermediärprodukte für die Biosynthese von Zellbestandteilen generiert. Nur unter ungünstigen Wachstumsbedingungen (Substratüberangebot, Spurenelementlimitierung oder in Anwesenheit von toxischen Verbindungen bzw. von Inhibitoren) werden die Substrate von einigen Bakteriengruppen unvollständig oxidiert. Im Gegensatz dazu oxidieren Essigsäurebakterien die entsprechenden Substrate auch unter „normalen“ Bedingungen unvollständig, und es werden energiereiche Produkte ausgeschieden. Neben Alkoholen, Aldehyden und Ketonen können Essigsäurebakterien auch verschiedene Zucker und Zucker-Alkohole wie D-Glucose, Glycerin und D-Sorbitol oxidieren. Derartige Reaktionen werden als oxidative Fermentation bezeichnet, da sie die unvollständige Oxidation von Zuckern oder Alkoholen beinhalten, einhergehend mit der Akkumulierung der korrespondierenden Oxidationsprodukte im Medium.

Für die Oxidation der verschiedenen Zucker, Polyole und Alkohole ist *G. oxydans* mit zwei verschiedenen Enzymsystemen ausgestattet (DE LEY et al., 1984). Die Systeme unterscheiden sich sowohl in der Funktion als auch in der Lokalisierung in der Zelle. Das erste, in der Cytoplasmamembran verankerte Enzymsystem, ist für die direkte Oxidation der Substrate verantwortlich. Die zugehörigen Enzyme katalysieren die Dehydrogenierung verschiedener nicht-phosphorylierter Zucker, Alkohole und Polyole. Die aktiven Zentren der Enzyme ragen dabei in den periplasmatischen Raum, so dass ein Transport der Substrate in die Zelle nicht notwendig ist. Die unvollständig oxidierten Produkte werden direkt und nahezu quantitativ ins Medium ausgeschieden (DEPPENMEIER et al., 2002). Die membrangebundenen unvollständigen Oxidationsreaktionen finden dabei stereo- und regioselektiv statt (KULHANEK, 1989). D-Sorbitol wird z. B. spezifisch zu L-Sorbose umgesetzt. Die Elektronen aus den Oxidationsreaktionen werden über Ubichinon auf eine Chinol-Oxidase von Typ *bo₃* übertragen (MATSUSHITA et al., 1994). Dieses Enzym nutzt molekularen Sauerstoff als terminalen Elektronenakzeptor in einer Reaktion, die mit einer Protonentranslokation einhergeht. Die Kopplungseffizienz ist allerdings sehr gering, so dass der Prozess der unvollständigen Oxidation mit einer limitierten Biomasse-Produktion einhergeht. Für die zweite Gruppe von Oxidoreduktasen ist die Aufnahme der Substrate in die Zelle erforderlich. Es handelt sich meist um NAD(P)-abhängige Enzyme, die die Substrate erst nach Phosphorylierung oxidieren. Die entstehenden Produkte werden in der Regel über den Pentosephosphat-Weg metabolisiert. Die Verwertung der Substrate durch die NAD(P)-abhängigen Enzyme dient nicht der Energiegewinnung, sondern dem intrazellulären Metabolismus zur Synthese von Biomasse (MATSUSHITA et al., 1994).

Die charakteristischen Eigenschaften von *G. oxydans*, zu denen sowohl die stereo- und regioselektive Oxidation verschiedener Substrate als auch die Ausscheidung der unvollständig oxidierten Produkte ins Medium zählen, machen diesen Organismus auch für die Biotechnologie sehr interessant. Schon seit Jahrhunderten werden *Gluconobacter*- und *Acetobacter*-Stämme erfolgreich für die Produktion von Essig eingesetzt. Dieser Prozess wird durch zwei verschiedene Enzyme katalysiert. Im ersten Schritt wird Ethanol zu Acetaldehyd umgesetzt. Dieses erfolgt durch die Enzymaktivität der membrangebundenen Alkohol-Dehydrogenase (MATSUSHITA et al., 1999). In

einem zweiten Schritt erfolgt die Oxidation von Acetaldehyd zu Essigsäure. In diese Katalyse ist die ebenfalls membrangebundene Aldehyd-Dehydrogenase involviert (MATSUSHITA et al., 1994). Weitere Produkte, die unter Einsatz von *Gluconobacter oxydans* hergestellt werden, sind z. B. Vitamin C (L-Ascorbinsäure), Dihydroxyaceton als Selbstbräunungsmittel, und Miglitol, ein Medikament für die Behandlung von nicht-Insulin-abhängiger Diabetes mellitus. Dihydroxyaceton entsteht durch die Aktivität der membrangebundenen Glycerin-Dehydrogenase (ADACHI et al., 2001). Bei der Synthese von Vitamin C katalysiert *G. oxydans* die stereo- und regioselektive Umsetzung von D-Sorbitol zu L-Sorbose. Durch weitere chemische Schritte entsteht letztendlich Ascorbinsäure (REICHSTEIN & GRÜSSNER, 1934). Diese Verbindung weist mehrere Isomere auf, jedoch ist nur die L-Form biologisch aktiv. Daher ist die Herstellung enantiomerenreiner Substanzen besonders wichtig. Chemisch ließe sich derartiges nur mit einer aufwendigen und kostenintensiven Schutzgruppenchemie verwirklichen.

In der Vergangenheit wurden bereits einige Dehydrogenasen identifiziert und charakterisiert, die an der Oxidation verschiedener Substrate beteiligt sind. Als Beispiel seien die Alkohol- (MATSUSHITA et al., 1999), Aldehyd- (MATSUSHITA et al., 1994), Sorbitol- (CHOI et al., 1995) und Glycerin-Dehydrogenase (ADACHI et al., 2001) genannt. Diese Enzyme enthalten als prosthetische Gruppe Pyrrolochinolinchinon (PQQ). Zusätzlich weisen die Alkohol-, Aldehyd- und Sorbitol-Dehydrogenase Untereinheiten auf, die Häm c enthalten. Darüber hinaus wurden auch Flavin-abhängige Dehydrogenasen beschrieben, welche Gluconat, Ketogluconat, Sorbitol und Sorbose oxidieren (DEPPENMEIER et al., 2002). Es wird jedoch vermutet, dass das oxidative Potential weitaus größer ist, als bislang bekannt.

Das Ziel dieser Arbeit war die vollständige Sequenzierung des Genoms von *G. oxydans* 621H (Abb. 1). Darüber hinaus sollte anhand der gewonnenen Daten eine Rekonstruktion des Stoffwechsels von *G. oxydans* vorgenommen werden. Die Entschlüsselung der genomischen Sequenz ist ein wichtiges Instrument für weitere Untersuchungen bezüglich der bereits bestehenden biotechnologischen Prozesse, die von diesem Organismus katalysiert werden. Wie oben bereits erwähnt, ist davon auszugehen, dass das oxidative Potential von *G. oxydans* keinesfalls vollständig bekannt ist. Es wird angenommen, dass das Genom eine Vielzahl bislang unbekannter und noch

nicht charakterisierter Dehydrogenasen aufweist. Neben der Möglichkeit, neue, innovative biotechnologische Verfahren zu entwickeln, bietet die Entschlüsselung des Genoms von *G. oxydans* auch die Möglichkeit, Regulationsmechanismen, die bislang kaum erforscht sind, zu verstehen.

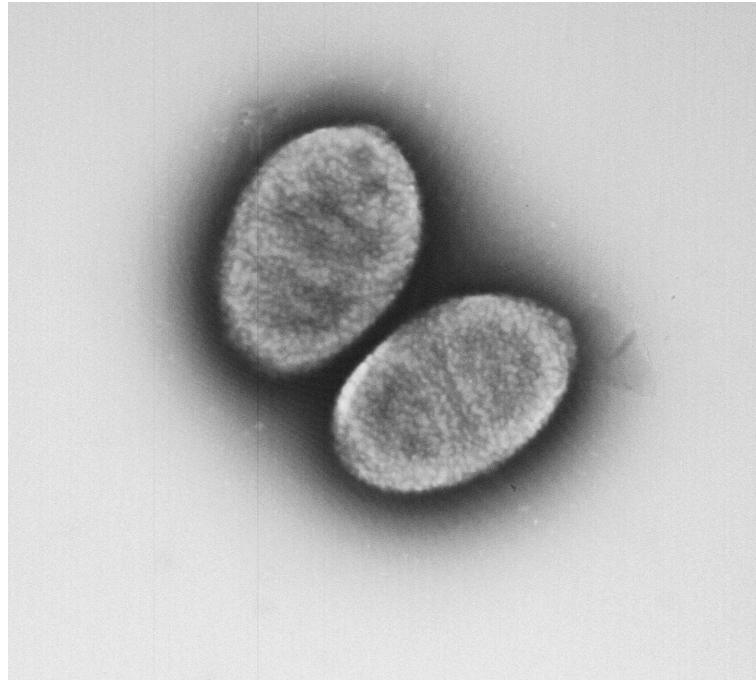


Abb. 1: Elektronenmikroskopische Aufnahme von *G. oxydans* (zur Verfügung gestellt von M. Hoppert und A. Ehrenreich, Institut für Mikrobiologie und Genetik, Universität Göttingen)

2 MATERIAL UND METHODEN

2.1 Organismen und Plasmide

In der Tabelle 1 sind die verwendeten Organismen aufgeführt. Tabelle 2 zeigt die verwendeten Klonierungsvektoren.

Tab. 1: Organismen und Plasmide

Stamm	Relevanter Geno- bzw. Phänotyp ^a	Referenz-Quelle
<i>Gluconobacter oxydans</i> 621H, DSM-Nr. 2343	Wildtyp	GILLIS & DE LEY, 1980
<i>Escherichia coli</i> DH5 α	F ⁻ , <i>lacZ</i> Δ M15, <i>recA</i> 1, π , <i>hsdR</i> 17, <i>supE</i> 44, Δ (<i>lacZYA</i> , <i>argF</i>)	HANAHAAN, 1983

^a: Genotypbezeichnungen für *Escherichia coli* siehe BACHMANN, 1990.

Tab. 2: Vektoren

Plasmid	Relevanter Geno ^a - bzw. Phänotyp ^b	Größe (kb)	Referenz-Quelle
pGEM [®] -3Z	Amp ^r , <i>lacPOZ'</i>	3,0	Promega Serva GmbH, Heidelberg
Lorist6	Amp ^r	5,6	GIBSON et al., 1987

^a: Genotypbezeichnung: *lacPOZ'*: Promotor, Operator und α -Peptid der β -Galaktosidase aus *Escherichia coli*; siehe BACHMANN, 1990.

^b: Phänotypbezeichnung: Amp^r: Ampicillinresistenz.

2.2 Nährmedien, Puffer und Stammlösungen

2.2.1 Komplexmedien

Im Folgenden sind die verwendeten Nährlösungen aufgeführt. Alle Medien wurden durch Autoklavieren für 30 min bei 121 °C sterilisiert. Feste Medien wurden durch den Zusatz von 1,5 % [w/v] Agar zu den entsprechenden Nährlösungen hergestellt.

Luria Bertani Medium (LB) (SAMBROOCK et al., 1989)

Trypton	10	g
Hefeextrakt	5	g
NaCl	10	g
H ₂ O _{dem.}	ad 1000	ml

SOB-Medium (HANAHAN, 1983)

Trypton	2 %	[w/v]
Hefeextrakt	0,5 %	[w/v]
NaCl	10	mM
KCl	2,5	mM
MgSO ₄	10	mM
MgCl ₂	10	mM

Die Magnesiumsalze wurden als 1 M Stammlösungen angesetzt und getrennt autoklaviert.

Mannitol-Medium

Hefeextrakt	5	g
Trypton	3	g
D-Mannitol	10	g
H ₂ O _{dem.}	ad 1000	ml
pH 6,0		

2.2.2 Antibiotika und Medienzusätze

Antibiotika-Stammlösungen sowie IPTG und X-Gal wurden dem autoklavierten und auf mindestens 50 °C abgekühlten Medium zur Selektion entsprechender *E. coli* Stämme zugesetzt (Tab. 3).

Tab. 3: Verwendete Antibiotika in ihren Stamm- und Endkonzentrationen

Substanz	Stammlösung [mg/ml]	Lösungsmittel	Endkonzentration [µg/ml]
Ampicillin	100	H ₂ O _{dest.}	100
Kanamycin	20	H ₂ O _{dest.}	20 - 40
Cefoxitin	50	H ₂ O _{dest.}	50
Streptomycin	75	H ₂ O _{dest.}	75
IPTG	100	H ₂ O _{dest.}	40
X-Gal	40	N,N-Dimethylformamid ^a	40

^a Die Lösung wurde unmittelbar vor Gebrauch frisch angesetzt.

2.3 Kultivierung und Stammhaltung von Mikroorganismen

2.3.1 Zellanzucht

G. oxydans 621H und *E. coli* wurden aerob in Flüssigkultur in Reagenzgläsern oder Erlenmeyerkolben angezogen, deren Volumina dem zehnfachen bis zwanzigfachen des Kulturvolumens entsprachen. *G. oxydans* wurde in Mannitol-Medium angezogen und bei 30 °C auf dem Rotationsschüttler inkubiert. *E. coli* wurde in LB-Medium kultiviert und bei 37 °C auf dem Rotationsschüttler inkubiert. Die Anzucht auf festen Medien erfolgte bei den oben angegebenen Temperaturen

2.3.2 Stammhaltung

Für die kurzfristige Stammhaltung wurden *E. coli* und *G. oxydans* auf entsprechenden Agarplatten gelagert. Diese Kulturen wurden alle vier Wochen überimpft. Für das Anlegen einer Stammkultur wurden in einem sterilen, verschraubbaren 1,5 ml

Reaktionsgefäß 900 µl einer Übernachtskultur mit 100 µl Dimethylsulfoxid (DMSO) gemischt. Die Lagerung erfolgte bei -70 °C .

2.3.3 Reinheitskontrolle

Die Reinheit der Bakterienkulturen wurde durch Vereinzlung auf Agarplatten mit anschließender makroskopischer Kontrolle und durch mikroskopische Kontrolle der Zellmorphologie im Phasenkontrastmikroskop (Fa. Carl Zeiss, Standard RA, Oberkochen) sichergestellt.

2.4 Messung der optischen Dichte

Die Bestimmung der Zelldichte in Flüssigkultur erfolgte im Spektralphotometer (Uvicon 860, Fa. Kontron Inst. Neufarn) bei 600 nm. Als Leerwert diente das entsprechende unbewachsene Medium. Proben, die eine größere OD_{600} als 0,3 aufwiesen wurden mit Medium verdünnt.

2.5 Standardtechniken für das Arbeiten mit Nukleinsäuren

2.5.1 Behandlung von Geräten und Lösungen

Zur Inaktivierung von Nukleasen wurden alle hitzestabilen Geräte und Lösungen durch Autoklavieren (20 min bei 121 °C) sterilisiert. Nicht autoklavierbare Materialien wurden mit 70 %igem Ethanol [v/v] gespült oder mit 96 %igem Ethanol [v/v] abgeflammt. Hitzelabile Substanzen (z. B. Antibiotika, Lysozym) wurden sterilen Lösungen in pulverisierter Form zugegeben, oder entsprechende Lösungen wurden sterilfiltriert.

2.5.2 Reinigung und Konzentrierung von Nukleinsäuren

2.5.2.1 Phenol/Chloroform-Extraktion

Zur Entfernung von Proteinen aus DNA-Lösungen, wurden diese mit 0,5 Vol. Phenol versetzt, kräftig geschüttelt, mit 0,5 Vol. Chloroform/Isoamylalkohol (24:1) versetzt, erneut emulgiert und zur Phasentrennung zentrifugiert (4 °C, 13000 Upm, 5 min in Heraeus 17 RS Biofuge, Heraeus, Osterode/Harz). Die DNA-haltige Phase (obere Phase) wurde erneut mit 1 Vol. Chloroform/Isoamylalkohol extrahiert. Anschließend wurde die DNA mit Isopropanol oder Ethanol gefällt.

2.5.2.2 Fällung von Nukleinsäuren

Zur Fällung wurde die DNA-Lösung zunächst mit 1/10 Vol. [v/v] einer 5 M Ammoniumacetatlösung versetzt, um die DNA zu dehydratisieren und den Gehalt an monovalenten Kationen zu erhöhen. Das Gemisch wurde geschwenkt und kurz bei Raumtemperatur inkubiert. Sollte mit Isopropanol gefällt werden, wurden davon 0,7-0,8 Vol. zugegeben, gemischt und zentrifugiert (30 min, 13000 Upm, RT in Heraeus 17 RS Biofuge, Heraeus, Osterode/Harz). Der Überstand wurde verworfen und das Pellet mit 70 % [v/v] Ethanol (eiskalt, unvergällt) gewaschen. Bei der Fällung mit Ethanol wurden 2,5 Vol. 96 % [v/v] Ethanol (eiskalt, unvergällt) zu der Lösung gegeben, gemischt und zentrifugiert (30 min, 13000 Upm, 4 °C in Heraeus 17 RS Biofuge, Heraeus, Osterode/Harz). Das bei der Fällung erhaltene Pellet wurde in der Vapofuge getrocknet und in dem gewünschten Volumen $H_2O_{\text{dest.}}$ aufgenommen.

2.5.3 Konzentrationsbestimmung und Reinheitskontrolle von DNA

Die Konzentration von DNA-Lösungen konnte durch photometrische Messung bestimmt werden. Hierzu wurde die Extinktion einer mit $H_2O_{\text{dest.}}$ verdünnten Nukleinsäure-Lösung bei 260 nm gegen Wasser in einer Quarzküvette der Schichtdicke $d = 1 \text{ cm}$ ermittelt. Für eine OD_{260} von 1,0 wurden folgende Konzentrationen angenommen (SAMBROOK et al., 1989):

doppelsträngige DNA	50 µg/ml
einzelsträngige DNA	40 µg/ml
Oligonukleotide	31 µg/ml

Eine Abschätzung der DNA Konzentration konnte auch visuell anhand der Intensität der Banden in einem Agarose-Gel (s. 2.5.4) erfolgen. Als Referenz diente dabei die 1 kb - Leiter (Fa. MBI Fermentas, Wilna Litauen), von der jeweils 0,2 µg aufgetragen wurden. Zusätzlich konnte die Reinheit einer DNA-Lösung durch OD-Bestimmung bei 230 nm und 280 nm überprüft werden. Für reine DNA gilt:

$$OD_{260} : OD_{280} = 1,8$$

$$OD_{230} : OD_{260} : OD_{280} = 0,45 : 1,0 : 0,515$$

Verunreinigungen durch Proteine oder Phenol sind an einer deutlichen Steigerung der Absorption bei 280 nm erkennbar, Polysaccharide hingegen erhöhen die Absorption bei 230 nm.

2.5.4 Standard-Agarosegelelektrophorese

Sowohl zur analytischen als auch zur präparativen Auftrennung von DNA- und DNA-Fragmenten wurden horizontale Gelelektrophoresen in Kammern der Größe 11,0 x 6,5 cm (kleine Gele) und 14 x 12 cm (große Gele) durchgeführt. Die Agarose wurde in Abhängigkeit der zu untersuchenden Fragmentgrößen in der Regel 0,8 %ig [w/v] in TAE-Puffer angesetzt, welcher gleichzeitig auch als Elektrophoresepuffer diente. Vor dem Auftragen in die Geltaschen wurden die DNA-Proben mit 0,2 Vol. Stop-Mix (= Beschwerer-Lösung) versetzt. Dadurch konnte man während der Elektrophorese die Lauffront verfolgen, und eventuell noch vorhandene Proteine wurden denaturiert. Zur späteren Auswertung und Größenbestimmung wurden zusätzlich noch 4 µl (bei einem großen Gel 10 µl) DNA-Längenstandard aufgetragen (s. 2.5.5). Nach der Elektrophorese (25-70 V, 0,5-2 h für Minigele; 40-120 V; 4-6 h für präparative Gele) wurden die Gele für 15 min im Ethidiumbromidbad (1 µg/ml H₂O) angefärbt, gewässert und die Nukleinsäurebanden mittels eines UV-Transilluminators in der Gel-Doc-Anlage (Gel-doc, Fa. Biorad, München) bei 254 nm sichtbar gemacht. Bei Bedarf wurden die Gele zur Dokumentation in der entsprechenden Auswertungssoftware (Fa. Biorad,

München) digitalisiert. Wahlweise war eine Übertragung des Bandenmusters auf Thermopapier mittels eines Thermaldruckers möglich (Fa. Biorad, München).

50x TAE-Puffer (SAMBROOK et al., 1989)

Tris	242	g	2	M
EDTA (0,5 M ; pH 8,0)	100	ml	50	mM
Essigsäure (konz.)	57	ml	1	M
H ₂ O _{dest.}	ad 1000 ml			
pH 8,5				

Stop-Mix (SAMBROOK et al., 1989)

EDTA	19	mg	1,0	mM
Harnstoff	21	g	7,0	M
Saccharose	25	g	50,0 %	[w/v]
Bromphenolblau	50	mg	0,1 %	[w/v]
H ₂ O _{dest.}	ad 50 ml			
pH 7,0				

Aliquots der autoklavierten Lösung wurden bei -20 °C gelagert.

Ethidiumbromid-Stammlösung

Die Stammlösung wurde in einer Konzentration von 10 mg/ml in Wasser angesetzt und lichtgeschützt bei 4 °C gelagert.

2.5.5 Größenbestimmung von Nukleinsäuren

Zur Größenbestimmung elektrophoretisch aufgetrennter, linearer DNA-Fragmente, wurden entweder die GeneRuler™ 1 kb DNA Ladder oder die GeneRuler™ 100 bp DNA Ladder (Fa. MBI Fermentas) parallel auf das Gel aufgetragen (Abb. 2). Zur Auswertung wurden die bekannten Fragmentgrößen halblogarithmisch gegen ihre Wanderungsstrecke aufgetragen. Diese Eichkurve wurde zur Größenbestimmung herangezogen.

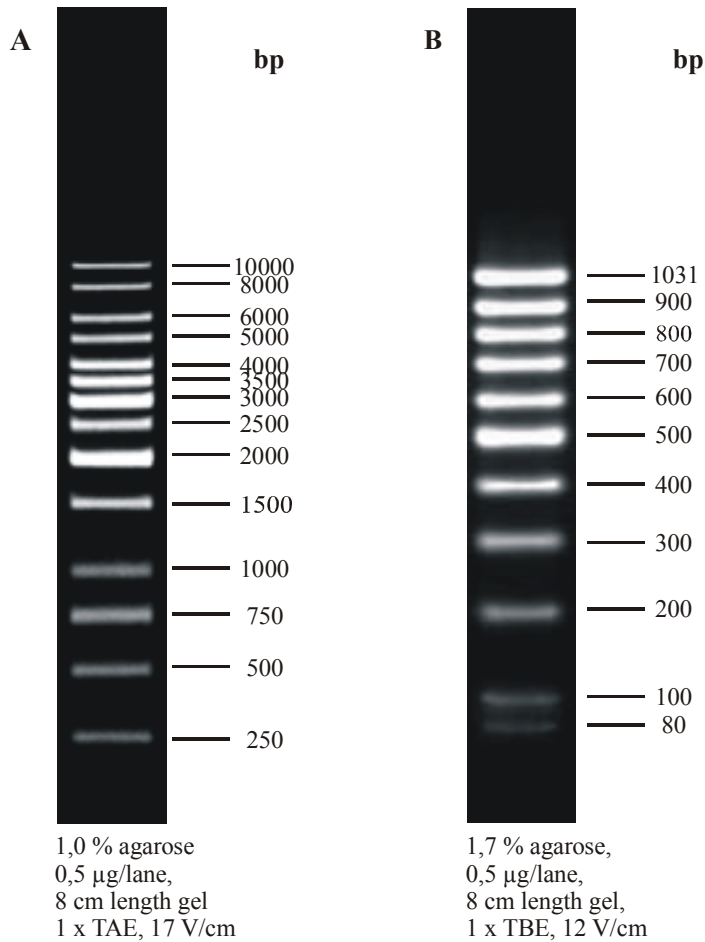


Abb. 2: A) GeneRuler™ 1 kb DNA Ladder, B) GeneRuler™ 100 bp DNA Ladder

2.6 Isolierung von Nukleinsäuren

2.6.1 Isolierung genomischer DNA mit dem AquaPure GenomicDNA Kit

Zur Isolierung hochreiner genomischer DNA von *G. oxydans* wurde diese mit dem AquaPure GenomicDNA Kit (Fa. Biorad, München) isoliert. Die verwendeten Lösungen waren Bestandteil des Kits. Die so gewonnene DNA konnte ohne weitere Aufreinigungsschritte als Template für die PCR eingesetzt werden. Die Zellen wurden in einer 5 ml Kultur in Mannitol-Medium angezogen. 500 µl der Zellsuspension wurden in ein 1,5 ml Reaktionsgefäß überführt und pelletiert, der Überstand wurde verworfen. Die Lyse der Zellen erfolgte durch Resuspendieren in 300 µl Genomic DNA Lysis Solution

und einer 5-minütigen Inkubation bei 80 °C. Zum Entfernen von RNA wurden 1,5 µl der im Kit enthaltenen RNase A-Lösung (4 mg/ml) zugegeben, die Zellen wurden vorsichtig geschwenkt und für 45 min bei 37 °C inkubiert. Die Fällung der Proteine erfolgte durch Zugabe von 100 µl Protein Precipitation Solution zu dem Zellysat, Mischen und anschließender Zentrifugation (15000 Upm, 3 min, RT). Der Überstand wurde vorsichtig in ein neues 1,5 ml Reaktionsgefäß überführt und die DNA durch Zugabe von 100 µl Isopropanol, vorsichtigem Schwenken und 1-minütiger Zentrifugation (15000 Upm, RT) gefällt. Die DNA wurde mit 70 % [v/v] Ethanol gewaschen, etwa 10 min an der Luft getrocknet und in 50 µl sterilem H₂O_{dest.} aufgenommen. Die Lagerung erfolgte bei 4 °C.

2.6.2 Isolierung genomischer DNA mit CTAB

Die Bakterienstämme wurden über Nacht in einer 100 ml Kultur angezogen und in GSA-Zentrifugenbechern durch 10-minütige Zentrifugation bei 8000 Upm geerntet. Das Zellpellet wurde in 9,5 ml TE-Puffer resuspendiert und in ein SS34-Röhrchen überführt; unter Invertieren wurden 0,5 ml 10 % [w/v] SDS und 50 µl Proteinase K (20 mg/ml) zugegeben. Es erfolgte eine Inkubation für 1 h bei 37 °C.

Anschließend wurden 1,8 ml NaCl (5 M) und 1,5 ml Cetyltrimethylammoniumbromid/NaCl-Lösung (CTAB) zu dem Lysat gegeben. Es erfolgte eine weitere Inkubation für 20 min bei 65 °C. Anschließend erfolgte eine Phenol-Chloroform-Extraktion (2.6.6) zur Entfernung der Proteine. Nach der Extraktion wurde die Oberphase in ein frisches SS34-Röhrchen überführt und die hier erhaltene DNA durch Zugabe von 0,8 Vol. Isopropanol gefällt. Nach Zentrifugation (30 min, RT, 14000 Upm) wurde der Überstand entfernt und das Pellet in der Vapofuge (UNIVAPO 150H, Fa. Zirbus, Osterode/Harz) getrocknet. Es wurde in 200-500 µl H₂O_{dest.} resuspendiert. Anschließend wurde die DNA-Lösung noch mit RNase behandelt.

TE-Puffer

Tris-HCl	50	mM
EDTA	10	mM

RNase

Der pH-Wert wurde mit HCl auf 8,0 eingestellt.

Cetyltrimethylammoniumbromid/NaCl-Lösung

NaCl	4,1	g
Cetyltrimethylammoniumbromid	10	g
H ₂ O _{dest.}	ad 100	ml

Das NaCl wurde zunächst in 80 ml H₂O_{dest.} gelöst, anschließend wurden 10 g Cetyltrimethylammoniumbromid langsam unter Rühren zugegeben und bis 65 °C erhitzt. Mit H₂O_{dest.} wurde auf 100 ml aufgefüllt.

2.6.3 Isolierung von Plasmid-DNA

Um Plasmid-DNA zu erhalten, die auch für weiteres molekulares Arbeiten verwendet werden konnte (z. B. Restriktionsverdau, Ligation), wurde die Anionenaustausch-Chromatographie des QIAprep Spin Miniprep Kits genutzt (QIAGEN GmbH, Hilden). Die Durchführung erfolgte nach den Angaben des Herstellers: 1,5-3,0 ml einer 5 ml Übernachtskultur in Selektivmedium wurden geerntet und in 250 µl P1 Puffer resuspendiert. Zur Zellyse wurden 250 µl P2 Puffer hinzugegeben und maximal 5 min bei Raumtemperatur inkubiert. Nach Neutralisation durch Zugabe von 350 µl N3 Puffer wurden die ausgefallenen Proteine durch Zentrifugation (15000 Upm, 10 min, RT) pelletiert. Der Überstand wurde auf eine QIAprep Spin Säule aufgetragen, und durch Zentrifugation (1 min, 13000 Upm, RT) wurde die DNA an das Säulenmaterial gebunden. Um besonders saubere DNA zu erhalten wurde diese zunächst mit 500 µl PB Puffer gewaschen. Anschließend wurde die gebundene DNA durch Zugabe von 750 µl PE Puffer gereinigt. Zur Elution wurden 50 µl H₂O_{dest.} eingesetzt. Sollten low-copy-Plasmide oder Cosmide isoliert werden, wurde das zur Elution verwendete H₂O_{dest.} auf 70 °C erhitzt.

2.6.4 Isolierung von DNA-Fragmenten aus Agarosegelen

Zur Isolierung von DNA-Fragmenten aus Agarosegelen wurden die entsprechenden mit Ethidiumbromid angefärbten Banden unter UV-Licht aus dem Gel ausgeschnitten. Die Elution fand mittels QIAquick Gel Extraction Kit (50) statt. Das Gewicht des ausgeschnittenen Gelstückes wurde mit der Feinwaage bestimmt. Anschließend wurde

das dreifache Gelvolumen des Puffers QX1 hinzugefügt und das Gemisch zum Auflösen des Gels unter mehrmaligem Schütteln bei 50 °C für 10 min inkubiert. Der Puffer QX1 enthält hohe Konzentrationen chaotroper Salze, was die Adsorption der zu eluierenden DNA an die Silicagel-Membran der QIAquicksäulen ermöglicht. Es folgten die Zugabe von 1 Gelvolumen Isopropanol, vorsichtiges Mischen und das Beladen der QIAquicksäule mit der Lösung. Nach Zentrifugation (RT, 13000 Upm, 1 min in Biofuge 15, Fa. Heraeus, Osterode/Harz) wurde die an die Säule gebundene DNA mit 0,75 ml PE-Puffer gewaschen (RT, 13000 Upm). Zur vollständigen Entfernung des Waschpuffers wurde nochmals zentrifugiert und die Säule anschließend in ein steriles 1,5 ml Reaktionsgefäß überführt. Die Elution der DNA erfolgte nach 1 minütiger Inkubation mit 30-50 µl EB-Puffer oder H₂O_{dest.} und anschließender Zentrifugation (RT, 13000 Upm, 1 min).

2.6.5 Aufreinigung von PCR-Produkten

Amplifizierte PCR-Produkte mussten, ehe sie sequenziert werden konnten, aufgereinigt werden. Dieses erfolgte mit dem QIAquick[®] PCR Purification Kit (Fa. Qiagen, Hilden) in den vom Hersteller mitgelieferten Puffersystemen. Dazu wurden 5 Volumen des Puffers PB mit dem PCR-Produkt vermischt und auf eine QIAquick-Säule gegeben. Durch Zentrifugation (1 min, 12000 Upm, RT) wurde die DNA an das Säulenmaterial gebunden. Der Durchfluss wurde verworfen und die Säule wurde mit 750 µl PE-Puffer gewaschen. Um sämtliche Reste des Waschpuffers von der Säule zu entfernen, wurde diese zweimal zentrifugiert (14000 Upm, RT). Zur Elution der PCR-Produkte wurden 50 µl H₂O_{dest.} auf die Säule gegeben, 1 min bei RT inkubiert und erneut 1 min zentrifugiert (14000 Upm, RT). Die PCR-Produkte wurden bis zur weiteren Verwendung bei -20 °C gelagert.

2.6.6 Phenol-Chloroform-Extraktion

Zur Entfernung von Proteinen aus DNA-Lösungen, wurden diese mit 0,5 Vol. Phenol versetzt, kräftig geschüttelt, mit 0,5 Vol. Chloroform/Isoamylalkohol (24:1) versetzt, erneut emulgiert und zur Phasentrennung zentrifugiert (4 °C, 13000 Upm, 5 min in Heraeus 17 RS Biofuge, Heraeus, Osterode/Harz). Die DNA-haltige Phase (obere

Phase) wurde erneut mit 1 Vol. Chloroform/Isoamylalkohol extrahiert. Anschließend wurde die DNA mit Isopropanol gefällt.

2.7 Enzymatische Modifikation von DNA

2.7.1 Spaltung von DNA durch Restriktionsendonukleasen

Sequenzspezifische Spaltung von DNA durch Restriktionsendonukleasen liefert lineare DNA-Fragmente mit definierten Enden. Der Verdau der DNA (1-3 U/ μ g) erfolgte in den vom Hersteller empfohlenen Puffersystemen. Die Inkubation erfolgte in der Regel bei 37 °C (siehe Empfehlung des Herstellers) für 3-12 h. Das Gesamtvolumen des Ansatzes betrug 10-50 μ l, wobei das Enzym (3-5 U Enzym/ μ g DNA) mindestens 1:10 verdünnt wurde, um eine Gesamtkonzentration von 5 % Glycerin nicht zu überschreiten, die sich störend auf die Spaltreaktion auswirken kann. Sollte mit zwei verschiedenen Enzymen, die unterschiedliche Pufferbedingungen benötigen, verdaut werden, so wurde erst die Restriktion in dem System mit niedrigerer Ionenstärke durchgeführt und diese anschließend für das zweite Enzym erhöht. Gegebenenfalls wurde eine Inaktivierung der Verdauansätze für 20 min bei 65 °C durchgeführt. Die Spaltprodukte wurden im Agarosegel überprüft.

2.7.2 Dephosphorylierung von DNA-Fragmenten

Vor der Ligation von Vektor-DNA mit dem zu klonierenden DNA-Fragment wurde die 5'-Phosphatgruppe der Vektor- oder Insert-DNA durch Behandlung mit alkalischer Phosphatase aus Kälberdarm (CIP) (SAMBROOK et al., 1989) abgespalten. Dadurch sollte die Ligation von Vektor-DNA oder die Ligation diskontinuierlicher, genomischer DNA-Fragmente verhindert werden. Die Dephosphorylierung wurde direkt im Verdauansatz durchgeführt. Nach dem Verdau wurde ein U alkalische Phosphatase pro 10 μ g zu dephosphorylierender DNA zugegeben und 30 min bei 37 °C inkubiert. Anschließend wurde erneut ein U Enzym zugegeben und erneut 30 min bei 37 °C inkubiert. Die Reaktion wurde durch Inkubation bei 65 °C für 15 min abgestoppt. Die DNA wurde mittels dem QIAquick[®] PCR Purification Kit (Fa. Qiagen, Hilden, s. 2.6.5) aufgereinigt.

2.7.3 Herstellung von stumpfen Enden

Zur Herstellung von stumpfen Enden wurde das Klenow-Fragment der DNA-Polymerase I eingesetzt. Ein Ansatz von 30 μl in $\text{H}_2\text{O}_{\text{dest.}}$ gelöster DNA (100 $\text{ng}/\mu\text{l}$), 1 μl Klenow-Fragment (Fa. MBI Fermentas, Wilna, Litauen) sowie 3,9 μl vom Hersteller mitgelieferter Puffer wurde für 3 min bei 37 °C inkubiert. Es erfolgte die Zugabe von 4 μl dNTPs (10 mM) und eine weitere Inkubation bei 37 °C für 10 min. Anschließend erfolgte eine Hitzeinaktivierung für 10 min bei 75 °C.

2.7.4 Ligation von DNA Fragmenten

Bei einer Ligation wird eine Phosphodiesterbindung zwischen doppelsträngigen DNA-Fragmenten, die mindestens ein freies 3'-Hydroxyende und 5'-Phosphatende besitzen, gebildet. Da der Erfolg einer Ligation unter anderem von Länge und Konzentration der Vektor- und der zu inserierenden DNA abhängt (DUGAICZYK et al., 1975), wurde das zu ligierende DNA-Fragment in 3- bis 5fachem Überschuss eingesetzt. Ligationen wurden in einem Gesamtvolumen von 15-20 μl mit 1 μl T4-DNA-Ligase (1 U/ μl , Fa. Boehringer Mannheim, Mannheim) in dem vom Hersteller gelieferten Puffer durchgeführt. Der Ligationsansatz wurde für etwa 16 h bei 16 °C inkubiert oder für 3 bis 16 h bei 22 °C. Anschließend wurde er direkt zur Transformation von *E. coli* eingesetzt (s. 2.8.2, 2.8.3).

2.8 Übertragung von DNA in *E. coli* und Selektion rekombinanter Klone

2.8.1 Herstellung von transformationskompetenten *E. coli*-Zellen (INOUE et al., 1990)

Die Anzucht der Zellen erfolgte bei 18 °C in 250 ml SOB-Medium in einem Schüttelwasserbad. Die Hauptkultur wurde mit einer 5 ml Vorkultur von *E. coli* DH5 α angeimpft (37 °C, über Nacht). Unter den oben genannten Bedingungen hat *E. coli* DH5 α eine Verdopplungszeit von etwa 7 h, so dass die Hauptkultur nach ungefähr 2 Tagen geerntet werden kann: die $\text{OD}_{600 \text{ nm}}$ liegt dann im Bereich zwischen 0,4-0,6. Zur

Zellernte wurde die Kultur 10 min auf Eis gehalten, bei 4 °C in vorgekühlten GSA-Bechern bei 7000 Upm für 10 min zentrifugiert und in 80 ml TB-Puffer resuspendiert. Die Zellsuspension wurde weitere 10 min auf Eis gehalten, anschließend zentrifugiert, das Pellet in 20 ml TB resuspendiert und 1,5 ml DMSO unter leichtem Schwenken zugegeben. Nach 10 min auf Eis wurden zügig in vorgekühlte Reaktionsgefäße je 200 µl aliquotiert und bei -70 °C gelagert.

Transformationspuffer (TB)

Pipes	10	mM
MnCl ₂	55	mM
CaCl ₂	15	mM
KCl	250	mM

pH 6,7; das Mangansalz wurde getrennt autoklaviert

2.8.2 Transformation von *E. coli* DH5α durch Hitzeschock

Die kompetenten *E. coli*-Zellen wurden 5 min auf Eis aufgetaut, vorsichtig mit dem Ligationsansatz, bzw. dem zu transformierenden Plasmid, vermischt (1-20 µl DNA/200 µl Zellen) und zur DNA-Bindung 30 min auf Eis inkubiert. Nach einem Hitzeschock (45 s bei 42 °C) wurden die Zellen in 400 µl LB-Medium aufgenommen und 1 h unter Schütteln bei 37 °C inkubiert. Der Transformationsansatz wurde in 150 µl Portionen auf LB-Selektionsplatten ausplattiert und nicht länger als 20 h inkubiert, um die Bildung von Satelliten-Kolonien gering zu halten.

2.8.3 Transformation von *E. coli* durch Elektroporation

Eine andere Möglichkeit DNA in *E. coli* zu transferieren ist die sogenannte Elektroporation. Dabei werden durch das Anlegen einer hohen Spannung die Membranen kurzzeitig für Plasmid- oder Cosmid-DNA durchlässig (FIEDLER & WIRTH, 1988). Für die Elektroporation von *E. coli* DH5α wurde der Stamm über Nacht in einer 5 ml Vorkultur angezogen. Mit dieser wurden 500 ml LB-Medium angeimpft, diese Hauptkultur wurde unter Schütteln bei 37 °C bis zu einer OD_{600nm} von 0,7-0,8 angezogen. Nach einer 15-minütigen Inkubation auf Eis wurden die Zellen geerntet (5000 Upm, 10 min, 4 °C) und anschließend zweimal mit eiskaltem H₂O_{dest.} gewaschen.

Das Pellet wurde mit 20 ml einer 10 %igen, eiskalten Glycerin-Lösung gewaschen und dann in 800 μ l derselben Lösung aufgenommen. Die Zellen wurden in 40 μ l Portionen aliquotiert und entweder sofort verwendet oder bei -70 °C gelagert.

Die Elektroporation wurde an einem „Gene Pulser“ (Fa. BioRad Laboratories, München) durchgeführt. Dazu wurden 0,5-2 μ l der zu transformierenden DNA mit einem Aliquot (40 μ l) der elektrokompenten Zellen in vorgekühlten Elektroporationsküvetten vorsichtig gemischt. Es wurde eine Spannung von 2,5 kV, eine Kapazität von 25 μ F und ein Widerstand von 200 Ω gewählt. Direkt nach der Elektroporation wurde 1 ml LB-Medium zugegeben, gemischt, der Ansatz in ein 1,5 ml Reaktionsgefäß überführt und 1 h bei 37 °C unter gelegentlichem Schwenken inkubiert. Anschließend wurden je 100 μ l auf Selektivplatten ausplattiert. Gegebenenfalls wurden die Zellen erst abzentrifugiert (12000 Upm, 1 min), in 50-100 μ l LB-Medium aufgenommen und dann ausplattiert.

2.8.4 Selektion rekombinanter Klone durch den Blau-Weiß-Test

Neben der Selektion rekombinanter Klone durch plasmidvermittelte Antibiotikaresistenzen gibt es bei manchen Plasmiden (z. B. pTZ19R) in *E. coli*-Wirtsstämmen wie DH5 α ein weiteres Testsystem, den Blau-Weiß-Test. Dieser Test zeigt an, ob das aufgenommene Plasmid ein Insert trägt oder rezirkularisiert vorliegt. Kolonien mit Plasmid-Insert behalten auf LB-Platten, die mit X-Gal und IPTG versetzt sind, ihre helle Farbe bei, insertlose Plasmide bewirken eine Blaufärbung der Kolonien. *E. coli* DH5 α trägt eine aminoterminal Deletion im β -Galaktosidase-Gen. Nach der Induktion durch IPTG wird eine verkürzte enzymatisch inaktive β -Galaktosidase gebildet. pTZ19R besitzt das *lacZ'*-Gen, welches für die fehlenden 146 Aminosäuren kodiert. Es vermag mit der verkürzten β -Galaktosidase zum aktiven Enzym zu assoziieren, welches X-Gal zum blauen Indigo-Derivat umsetzt. Die Kolonien färben sich blau. Da sich innerhalb des *lacZ'*-Gens die multiple Klonierungsstelle befindet, wird dieses bei Einlagerung von Inserts unterbrochen und es kann keine aktive β -Galaktosidase gebildet werden. X-Gal wird nicht umgesetzt und die Kolonien bleiben weiß.

2.9 *In vitro*-Amplifikation von DNA durch Polymerasekettenreaktion (PCR)

Mit Hilfe der PCR wurden zum einen Sequenzbereiche hinsichtlich ihrer richtigen Assemblierung überprüft, zum anderen wurden durch die Sequenzierung generierter PCR-Produkte Sequenzlücken (*Gaps*) innerhalb des Genoms geschlossen. Als Matrize dienten genomische DNA aus *G. oxydans* und Cosmide, die als Insert genomische DNA von diesem Organismus trugen (ca. 40 kb). Aufgrund des mit etwa 61 % hohen GC-Gehalts von *G. oxydans* wurde als Enzym die Ex-*Taq*-DNA-Polymerase (Fa. Takara, Apen) eingesetzt, die sich vor allem durch eine hohe Thermostabilität auszeichnet. Die verwendeten Starteroligonukleotide (Primer) hatten in der Regel eine Länge von 17-18 Basen. Eine optimale Amplifikation war gewährleistet, wenn beide Primer einen ähnlichen Schmelzpunkt (T_m) aufwiesen und in gleicher Konzentration vorlagen. Der Schmelzpunkt wurde nach folgender Formel berechnet (CHESTER & MARSHAK, 1993):

$$T_m = 69,3 + 0,41 (\% \text{ GC-Gehalt des Primers}) - 650/L$$

L = Länge des Primers

Als Faustregel zur Berechnung von T_m der Primer kann man auch 2 °C für A oder T und 4 °C für G oder C berechnen.

Die Standard-PCR-Reaktionen fanden in 100 µl-Ansätzen in 0,5 ml Reaktionsgefäßen statt. Folgende Komponenten wurden auf Eis zusammenpipettiert:

Primer 1 (5 pmol/µl)	8	µl
Primer 2 (5 pmol/µl)	8	µl
Matrizen-DNA (100 ng/µl)	1	µl
dNTP-Gemisch (2,5 mM)	8	µl
10 x Reaktionspuffer	10	µl
H ₂ O _{dest.}	64,6	µl
DNA-Polymerase (5 U/µl)	0,4	µl

Das Enzym wurde grundsätzlich als letztes zu dem Ansatz pipettiert, anschließend wurden alle Komponenten noch einmal gemischt. Die Amplifizierung erfolgte im Mastercycler[®] Gradient (Fa. Eppendorf, Hamburg), wobei folgendes Programm verwendet wurde:

1. Initiale Denaturierung	95 °C	4 min
2. Denaturierung	96 °C	1,5 min
3. Anlagerung	x °C	1 min
4. Polymerisierung	72 °C	y min
5. Zyklen	30 mal Schritt 2-4	
6. Finale Elongation	72 °C	10 min

Die Anlagerungstemperatur x war abhängig von der Schmelztemperatur der Primer und lag idealerweise 5-7 °C tiefer als T_m . Die Dauer y der Polymerisierung betrug 1 min pro 1 kb zu amplifizierender DNA. Bei kombinatorischer PCR zur Überwindung von Sequenzlücken, also Reaktionen, bei denen die erwartete Produktgröße unbekannt war, betrug die Polymerisierung 8 min. Das bedeutet, dass das Amplifikat maximal 8 kb groß sein kann, was die Grenze der Polymerisierung, welche die *Ex-Taq*-Polymerase leisten kann, darstellt. Zur Optimierung der PCR konnten einige Parameter, wie z. B. Anlagerungstemperatur und die Anzahl der Zyklen, variiert werden. Darüber hinaus konnte durch Zugabe von DMSO [4 %, v/v] die Spezifität erhöht werden. Nach Abschluß der Amplifikation wurde ein Aliquot im Agarosegel überprüft. Anschließend erfolgte gegebenenfalls die Aufreinigung (s. 2.6.5) und Sequenzierung (s. 2.10) der Produkte.

2.10 Sequenzierung von DNA

Die zum Schließen des Genoms notwendigen DNA-Sequenzierungen erfolgten im Göttinger Genom-Analyselabor. Es wurde mit dem ABI PRISM 377 DNA Sequencer (Fa. Perkin Elmer/Applied Biosystems) sequenziert. Grundlage des Verfahrens ist die von SANGER et al., (1977) entwickelte Didesoxynukleotid-Kettenabbruchmethode. Abweichend von SANGER et al. (1977) werden die Abbruchfragmente bereits während der gelelektrophoretischen Auftrennung durch ihre Fluoreszenzsignale detektiert, die durch das Anregen mit einem Laserstrahl emittiert werden (VOSS et al., 1992). Im

Unterschied zum A.L.F. oder Licor werden beim Sequenzieren mit dem ABI PRISM 377 DNA Sequencer nicht markierte Primer, sondern markierte Oligonukleotide verwendet. Ein optimal hergestellter Sequenzieransatz lieferte eine auflösbare DNA-Leseweite von ca. 800 bp.

Nach einem Sequenzierlauf werden die Rohdaten, also die Fluoreszenzsignale automatisch in DNA Basen umgewandelt. Dieser Vorgang wird als *base calling* bezeichnet. Hierzu wurde der Basecaller Phred verwendet (EWING et al., 1998). Phred liest die Daten von den Sequenziermaschinen (*trace files*), wandelt sie in Basen um und ordnet jeder einzelnen Base darüber hinaus noch einen Qualitätswert zu. Die so erhaltenen Daten können dann in Genomassemblierungsprogramme, wie z. B. Consed oder das Staden Package, eingespeist werden (BONFIELD & STADEN, 1996; GORDON et al., 1998; STADEN et al., 2000).

2.11 Die Prozessierung von Rohsequenzen

2.11.1 Einführung in das Staden Package

Im Zuge einer Sequenzierung eines Genoms ist eine Genomassemblierungssoftware, die letztendlich die erhaltenen Sequenzdaten visualisiert, unerlässlich. Die geläufigsten Programme, die dieses auch für große Datenmengen ermöglichen sind Consed (GORDON et al., 1998) und das STADEN-Software-Package (STADEN et al., 2000). In dieser Arbeit wurde das Staden Package verwendet, das im folgenden kurz vorgestellt werden soll. Im wesentlichen umfasst diese Genomassemblierungssoftware zwei Programme, nämlich Pregap4 und Gap4. Eine detaillierte Beschreibung ist unter <http://www.mrc-lmb.cam.ac.uk/pubseq/> (inklusive Benutzerhandbuch) zu finden.

2.11.2. Das Pregap4 Interface

Pregap4 ist ein Interface, das den Hintergrund der Daten, die prozessiert werden müssen, definiert. Erst nachdem neue DNA-Sequenzen dieses Interface durchlaufen haben, können sie in die Datenbank des Gap4 eingebracht werden. Der erste Schritt dabei ist, die *trace files* einschließlich der Werte für die Qualität der einzelnen Basen, die bei der Sequenzierung generiert werden, in ein für Gap4 lesbares Format zu

konvertieren. Im nächsten Schritt werden die sogenannten *experimental files* erstellt. Dabei handelt es sich im Prinzip um Text-Dateien, welche die einzelnen Basen wiedergeben. Darüber hinaus wird ein Namensschema erstellt, aus dem eindeutig hervorgeht, ob es sich bei der Sequenz um Plasmide, Cosmide oder PCR-Produkte handelt und welcher Primer verwendet wurde. Diese Informationen sind bei der Bearbeitung eines Genoms unerlässlich. Es folgt ein Quality Clip, bei dem Basen schlechter Qualität maskiert werden und Sequenzläufe mit zu schlechter Qualität fallen aus dem Datensatz heraus. Anschließend finden noch mehrere Screenings statt. Die Sequenz wird auf das Vorhandensein von Sequenzierungsvektoren überprüft, die Vektorsequenz wird gegebenenfalls maskiert oder entsprechend markiert. Außerdem findet ein Screen auf *E. coli* Sequenzen statt, da diese die häufigsten Verunreinigungen von Sequenzen darstellen.

Die oben beschriebenen Arbeitsschritte werden als das Prozessieren der Rohdaten bezeichnet. Zur Vereinfachung dieser Prozesse wurde vom Göttinger Genomanalyse-Labor (G2L) ein Skript geschrieben und in dieser Arbeit eingesetzt. Als Resultat der Prozessierung werden die Daten, also die prozessierten Sequenzen, in bestimmte Listen, z. B. *passed*, *failed*, etc. geschrieben. Die sogenannte *passed*-Liste enthält diejenigen Daten, welche das Prozessieren erfolgreich bestanden haben. Diese werden dann in die Gap4-Datenbank geladen.

2.11.3 Gap4

Das Gap4 des Staden Package ist das Programm, in dem die Bearbeitung der Daten erfolgte. Zunächst wurden die einzelnen Sequenzläufe zu sogenannten Contigs assembliert. STADEN (1980) definierte einen Contig als ein Set von Sequenzläufen, die durch das Überlappen von Sequenzen miteinander im Verhältnis stehen. Dabei kann jede Sequenz nur zu einem einzigen Contig gehören und jeder Contig beinhaltet mindestens eine Sequenz. Die Summe der Sequenzierläufe in einem Contig können eine Konsensussequenz bilden, wobei die Länge dieser Sequenz der Länge des Contigs entspricht. Das Ziel eines Genomprojekts ist es, am Ende einen Contig zu erhalten, der dem Chromosom entspricht. Sind Plasmide vorhanden so bildet jedes Plasmid ebenfalls einen Contig.

2.11.3.1 Der *Contig Editor*

Der *Contig Editor* ist die Visualisierung der einzelnen Basen, einschließlich der Qualitätswerte und der *Trace files*. Dieser diente als Grundlage zur Editierung, also zur Überprüfung jeder einzelnen Base der Sequenz. Das Ziel des Editierens war es, die Konsensussequenz zu verbessern und fehlerhafte Assemblierungen zu detektieren bzw. aufzulösen. Assemblierungsfehler (*Misassemblies*) treten vor allem in repetitiven Sequenzbereichen auf, wie z. B. in Transposasen-Genen oder Insertionselementen. Entsprechende Einzelsequenzen wurden markiert und aus dem Contig extrahiert. Erst wenn die Qualität der aus den Einzelsequenzen resultierenden Konsensussequenz ausreichend gut war (der dazugehörige Qualitätswert sollte größer als 50 % sein), wurde der entsprechende Bereich als korrekt angesehen. Bereiche schlechter Qualität wurden durch das Wiederholen von Sequenzläufen oder auch durch *Walkings* korrigiert und verifiziert. Die Fehlerrate der Sequenz sollte nicht höher als ein Fehler in 10000 Basen sein.

2.11.3.2 Das *Template Display*

Das *Template Display* enthält wichtige Informationen über die korrekte Assemblierung der einzelnen Sequenzen. Abb. 3 zeigt schematisch die Darstellung des *Template Displays*.

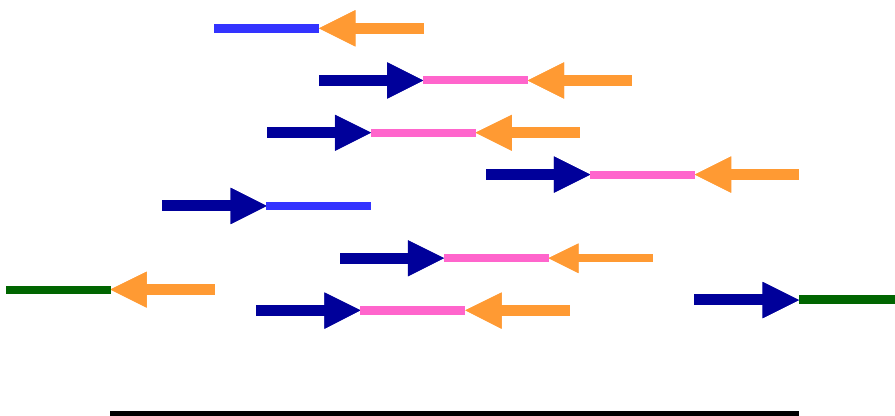


Abb. 3: Schematische Darstellung des *Template Displays*

Orange, bzw. blaue Pfeile: tatsächlich generierte Sequenz mit universellem *forward* (for) bzw. *reverse* (rev) Primern. Rosa Linien zeigen an, dass die relative Lage von for zu rev der durchschnittlichen Insertgröße entspricht. Grüne Linien: der Gegenlauf ist in einem anderen Contig, blaue Linien: der Gegenlauf existiert nicht in der Datenbank.

Der schwarze Strich symbolisiert einen Contig, die Pfeile spiegeln jeweils *forward*-, bzw. *reverse*-Läufe wieder. Entsprechen *forward* und *reverse*-Lauf eines Plasmides den Erwartungen, ist die Linie dazwischen rosa eingezeichnet. Existiert einer der beiden *Readings* nicht, folgt dem entsprechenden Pfeil eine blaue Linie. Liegt der entsprechende Gegenlauf in einem anderen Contig ist dies durch eine grüne Linie gekennzeichnet. Die Analyse der Sequenzierläufe mit Hilfe des *Template Displays* war ein wichtiges Werkzeug zur Überwindung von Sequenzlücken. Darüber hinaus konnte der logische Aufbau der Datenbank überprüft, und gegebenenfalls im *Contig Editor* korrigiert werden. Lag z. B. ein *Misassembly* vor, so konnte dieses im *Template Display* überprüft werden. Fand man innerhalb eines Contigs nur Sequenzläufe deren Gegenlauf in einem anderen Contig lag, so war dies ein sicherer Hinweis auf eine Fehlassemblierung.

2.12 Methoden zur Überwindung von Sequenzlücken in einem Genomsequenzierungsprojekt

Grundsätzlich sind die Methoden zur Überwindung von Sequenzlücken natürlich sehr vielfältig und müssen daher auf das jeweilige Genomsequenzierungsprojekt abgestimmt werden. Liegen bei einem Organismus z. B. schon die Daten naher Verwandter vor, so können auch diese Daten auf der Basis der Annotation zum Lückenschluss dienen. Da dies bei *G. oxydans* 621H jedoch nicht der Fall war, konnte diese Methode nicht angewendet werden. Eine Grundlage zum Schließen der Lücken bildeten unter anderem die beiden vorhandenen Genbanken von *G. oxydans*, die bereits von Integrated Genomics (Chicago) hergestellt wurden. Diese umfassten eine Plasmid-Genbank mit einer durchschnittlichen Insertgröße von 2 kb und eine Cosmid-Genbank mit einer durchschnittlichen Insertgröße von 40 kb. Im Folgenden werden die für dieses Projekt herangezogenen Methoden zum Lückenschluss kurz vorgestellt.

2.12.1. Identifikation von Contigüberlappungen mittels Gap4 oder BLASTN

Nachdem die erste Editierung des Genoms abgeschlossen war, wurde zunächst überprüft ob und welche Contigenden bereits überlappten. Dieses wurde unter anderem mit dem Programm Gap4 durchgeführt. Mit Hilfe der Funktion *find internal joins* konnten potentielle Überlappungen identifiziert werden. Außerdem wurden sämtliche Enden der Contigs mit Hilfe des BLAST-Algorithmus (ALTSCHUL et al., 1990) gegen die gesamte Datenbank abgeglichen. Auch hierbei wurden potentielle Überlappungen gefunden. Jede Übereinstimmung musste jedoch per Hand überprüft werden, da natürlich auch repetitive Sequenzbereiche Treffer erzielten. Auf die Behandlung der repetitiven Sequenzen, vor allen Transposase-Gene und Insertionselemente wird im nächsten Abschnitt noch eingegangen, da diese in dem Projekt ein besonderes Problem darstellten. Konnten zwischen zwei Contigenden ausreichend Übereinstimmungen gefunden werden, die auch dem logischen Aufbau der Datenbank entsprachen, wurden diese zusammengefügt. Dies wurde mit der Funktion *join contigs* durchgeführt.

2.12.2 Lückenschluss durch *Walkings* auf Plasmiden

Unter einem *Walking* versteht man das „Laufen“ auf einem vorhandenen *Template*. Ein Sequenzlauf liefert durchschnittlich 600-800-bp. Bei einer erwarteten Insertgröße der Plasmide von 2 kb bedeutet dies, dass nicht die gesamte Sequenz des Inserts vorliegt. In Abbildung 3 symbolisiert ein Pfeil jeweils einen Sequenzlauf, der mit einem universellen Primer, also einem Primer, der in der multiplen Klonierungsstelle des Vektors bindet, erstellt wurde. Leitet man nun auf der erhaltenen Sequenz spezifische Oligonukleotide ab und macht einen Sequenzlauf mit dem entsprechenden *Template* (Plasmid), erhält man neue Informationen. Diese Technik kann man sich beim Lückenschluss zunutze machen. Im Template Display des Gap4 symbolisieren grüne Striche (vgl. Abb. 3), dass der entsprechende Gegenlauf in einem anderen Contig liegt. Unter der Annahme, dass die Assemblierung der Sequenzen richtig ist, würde man an den Enden eines Contigs X erwarten, dass viele *Readings* dort vorliegen, deren Gegenlauf in einem anderen Contig liegt (im *Template Display* durch einen grünen Strich markiert). Ist die Sequenzlücke kleiner als 2 kb, also kleiner als die durchschnittliche Insertgröße der Plasmid-Genbank, so würde man erwarten, dass die

entsprechenden Gegenläufe der Plasmide alle auf demselben Contig Y liegen. Durch die Verwendung spezifischer Primer konnten einige Lücken zwischen zwei Contigs durch Plasmid-*Walking* geschlossen werden.

2.12.3 Lückenschluss mit Hilfe von PCR

2.12.3.1 Auf der Cosmid-Genbank basierende PCR

War eine Lücke zwischen zwei Contigs größer als 2 kb, also größer als die durchschnittliche Insertgröße der Plasmid-Genbank, konnte diese nicht durch einfaches Plasmid-*Walking* überwunden werden. Hier kam die schon vorhandene Cosmid-Genbank zum Einsatz. Betrachtet wurden wieder die Läufe, deren Gegenlauf in einem anderen Contig lagen, die also im *Template Display* grün dargestellt sind. Es wurden dann Cosmide ausgewählt, deren Gegenlauf in einem anderen Contig waren, wobei Lauf A in Contig X und Lauf B in Contig Y liegen musste und zwar bei allen ausgewählten Cosmiden. Wie unter 2.12.2 beschrieben wurden spezifische Primer ausgewählt und eine PCR durchgeführt. Ein Nachteil dieser Methode war, dass die zu amplifizierenden Fragmente sehr groß sein konnten. Darüber hinaus war der GC-Gehalt dieses Organismus mit 61 % recht hoch. Diese Gegebenheiten forderten eine äußerst hitzestabile Polymerase, da zum einen die Polymerisationszeit der zu amplifizierenden Fragmente lang war (vgl. 2.9), zum anderen musste aufgrund des hohen GC-Gehalts eine hohe und lange Denaturierungstemperatur und –zeit gewählt werden. Mit Hilfe der verwendeten Takara Ex-Taq-Polymerase™ (Fa. Takara, Apen) konnten Fragmente einer Größe bis zu 8 kb amplifiziert werden. Die PCR Produkte wurden anschließend mit den verwendeten spezifischen Primern sequenziert. Gegebenenfalls wurden nach erfolgreicher erster Sequenzierung der PCR-Produkte weitere spezifische Primer abgeleitet, bis die gesamte Sequenz des entsprechenden Amplifikats vorlag und die Lücke zwischen zwei Contigs geschlossen werden konnte.

2.12.3.2 Kombinatorische PCR

Die kombinatorische PCR wurde nur dann eingesetzt, wenn die zuvor erwähnten Methoden zur Überwindung von Sequenzlücken kein Resultat erzielten. Dabei wurden von allen Enden der Contigs spezifische Primer abgeleitet. Bei der anschließend

durchgeführten PCR wurden sämtliche Primerkombinationen getestet. Erhielt man bei einer Kombination ein Produkt, wurde dieses mit den für die PCR verwendeten Oligonukleotiden sequenziert. Voraussetzung für diese Methode war allerdings, dass die Lücke zwischen zwei Contigs kleiner als 8 kb sein muss, da die zu amplifizierende DNA Sequenz sonst zu groß war und die PCR keine spezifischen Produkte lieferte.

2.13 Auffinden repetitiver Sequenzen

Sequenzbereiche, die sich innerhalb eines Genoms wiederholen, also repetitive Sequenzen, stellen bei jedem Genomprojekt ein Problem dar, dem besondere Beachtung geschenkt werden muss. Sämtliche sich wiederholende Bereiche stellen Quellen für potentielle *Misassemblies* dar und müssen daher einzeln überprüft werden. Häufig vorkommende Sequenzbereiche sind unter anderem Transposase-Gene oder RNA-Bereiche. Dabei kamen die unter 2.13.1 und 2.13.2 beschriebenen Methoden zum Einsatz.

2.13.1 Identifizierung repetitiver Sequenzen mit Hilfe von Gap4

Innerhalb der Genomassemblierungssoftware Gap4 erfolgte die Identifizierung repetitiver Sequenzen (*repeats*) mit der Funktion *find repeats*. Sämtliche dabei gefundenen Bereiche des Genoms wurden markiert und hinsichtlich ihrer Assemblierung überprüft. Dieses erfolgte zunächst im *Template Display*. Unsichere Bereiche wurden darüber hinaus durch geeignete PCR Reaktionen überprüft. Gegebenenfalls wurden einzelne Sequenzen aus einem Contig entfernt. Zum Teil waren auch Contigs zusammengefügt, die nicht zusammen gehörten. In diesen Fällen musste der entsprechende Contig mit der Funktion *break a Contig* aufgebrochen werden. Dies führte dazu, dass zunächst eine neue Lücke entstand. In den meisten Fällen konnten diese neu entstandenen Lücken jedoch sofort durch die richtige Assemblierung wieder geschlossen werden.

2.13.2 Auffinden repetitiver Sequenzen mit Hilfe von RepVis und BLASTN

RepVis ist ein Computerprogramm, das speziell für das Auffinden und die Visualisierung von *Repeats* in einem Genom entwickelt wurde (KURTZ & SCHLEIERMACHER 1999). Die Anwendung dieses Programms ist jedoch nur dann sinnvoll, wenn die Anzahl der vorhandenen Contigs in dem Genomprojekt schon relativ gering ist (es sollten weniger als zehn Contigs vorhanden sein), weil die *Repeats* für jeden einzelnen Contig überprüft werden müssen. Als Inputfile diente jeweils die Konsensussequenz eines Contigs. Man erhielt als Resultat dann jede Sequenz, die mehr als 500 bp lang war und mehr als einmal in dem Datensatz vorhanden war. Jede dieser Sequenzen wurde als Text-Datei in Gap4 geladen und über *find internal joins* wurden die entsprechenden Bereiche der Konsensussequenz markiert. Die Text-Dateien wurden dabei jedoch nicht an die entsprechenden Stellen assembliert, sondern dienten lediglich zum einfachen Auffinden der Bereiche. Darüber hinaus erfolgte mit jedem ermittelten *Repeat* ein BLAST gegen die gesamte Datenbank (ALTSCHUL et al., 1990, vgl. 2.12.1) um zu gewährleisten, dass auch jeder repetitive Bereich markiert war. Jede dieser Markierungen wurde anschließend eingehend überprüft. Gegebenenfalls wurden wieder einzelne Sequenzen aus dem Datensatz extrahiert, Contigs gebrochen und neu zusammengefügt oder aber Bereiche durch PCR-Reaktionen amplifiziert und die entsprechenden Produkte sequenziert, um die Assemblierung zu überprüfen.

2.14 Sequenzanalyse

Nachdem in dem Projekt die Phase der Sequenzierung, Editierung und Überprüfung der Assemblierung abgeschlossen war, wurde versucht alle in *G. oxydans* 621H vorkommenden Gene zu identifizieren. Die Vorhersage aller potentiellen Gene in einem Organismus wird als ORF-Vorhersage bezeichnet. Dieses wurde mit entsprechenden Computeralgorithmen durchgeführt. Anschließend erfolgte die Annotation, das heißt jedem ORF wurde versucht eine Funktion zuzuordnen, die das aus dem Gen resultierende Protein beschreibt.

2.14.1 ORF-Vorhersage

Die ORF-Vorhersage ist der erste kritische Schritt, der die Grundlage der Annotation bildet. Da die Daten eines Genoms zu groß und zu vielfältig sind um dieses per Hand durchzuführen, wurden dafür Computer Algorithmen herangezogen. Sämtliche ORF-Vorhersageprogramme für Prokaryoten basieren auf dem Auffinden von Start- (ATG, GTG, TTG) und Stopcodons (TAA, TGA, TAG) in allen sechs Leserahmen. Mit Hilfe der Algorithmen werden solche Bereiche festgelegt, die zwischen Start und Stop auf demselben Leserahmen eine bestimmte Größe erreichen (z. B. 90 bp, also 30 Codons). Darüber hinaus wird überprüft, ob eine potentielle Ribosomenbindestelle vorhanden ist. Alle ORF-Vorhersageprogramme haben gemein, dass entweder zu viele (*overprediction*) oder zu wenig ORFs identifiziert werden, was eine manuelle Nachbearbeitung der identifizierten offenen Leserahmen unerlässlich macht.

Für das ORF-Finding auf dem Chromosom wurde das Programm YACOP (yet another combination of prediction, TECH & MERKL, 2003) eingesetzt. Dieses Programm wurde hausintern entwickelt und stellt eine Kombination aus den schon bekannten ORF-Finding Programmen Critica (BADGER & OLSEN, 1999), Glimmer (SALZBERG et al., 1998) und ZCURVE (GUO et al., 2003) dar. Der Anspruch dieses Programms ist, dass möglichst kein potentielles Gen unidentifiziert bleibt, aber auch die *overprediction* gering ist. Die Fehlerrate von YACOP ist zwar im Vergleich zu anderen ORF-Finding Programmen geringer, dennoch ist auch hier eine manuelle Nachbearbeitung notwendig. Das ORF-Finding auf den in *G. oxydans* identifizierten Plasmiden erfolgte aufgrund der geringeren Datenmenge (im Vergleich zum Chromosom) zum größten Teil manuell. Dafür wurde die im Internet frei erhältliche Software Artemis (RUTHERFORD et al., 2000; <http://www.sanger.ac.uk/Software/Artemis>) eingesetzt. Artemis ist ein verhältnismäßig kleines Programm zur Sequenzvisualisierung und Annotation. Es ist möglich, sich damit alle potentiellen ORFs darstellen zu lassen. Berücksichtigt wird dabei aber lediglich ein definierter Bereich (z. B. 30 Codons) zwischen einem Start- und einem Stopcodon, der in einem einzigen der sechs möglichen Leserahmen liegt. Mit Artemis ist es jedoch auch möglich, sich die aus dem potentiellen ORF ergebende Proteinsequenz anzeigen zu lassen und diese mit der NCBI-Sequenzdatenbank

(<http://www.ncbi.nlm.nih.gov/BLAST/>) abzugleichen. Anschließend erfolgte die manuelle Überprüfung.

2.14.2 Genomweite Annotation mit ERGO

Die Annotation hat zum Ziel, jedem aus einem Gen resultierenden Protein einen Namen zu geben, bzw. diesem eine Funktion zuzuordnen. Zur Erleichterung dieser Arbeit gibt es verschiedene Computerprogramme wie PEDANT, MAGPIE, oder ERGO. Im Institut für Mikrobiologie und Genetik der Universität Göttingen hat man das Programm ERGO etabliert (Fa. Integrated Genomics, Chicago; <http://www.integratedgenomics.com/>; Overbeek et al., 2003). Es handelt sich dabei um eine lokal installierte Version, so dass auch noch nicht veröffentlichte und sich noch in Bearbeitung befindende Projekte dort eingespeist werden können. Neben der Möglichkeit zur Annotation bietet ERGO noch eine Vielzahl weiterer Funktionen zur Sequenzanalyse. Zur Zeit umfasst die Datenbank Sequenzen aus insgesamt 402 Organismen, darunter 27 Archaeen, 81 Eukaryoten, 89 Viren und 205 Bakterien.

Bei der Annotation von *G. oxydans* wurden die von YACOP, bzw. Artemis ermittelten ORFs in ERGO eingespeist. Zum einen berechnet das Programm mittels des FASTA Algorithmus die Ähnlichkeiten aller *G. oxydans*-ORFs zu Sequenzen der internen ERGO-Datenbank, wobei dieses sowohl auf Nukleotid-, als auch auf Proteinebene geschieht. Darüber hinaus führt ERGO auch eine Ähnlichkeitssuche aller ORFs mit der externen NCBI-Datenbank durch. Der externe NCBI-Sequenzabgleich umfasst eine Vielzahl von Nukleotid- und Proteinsequenz-Datenbanken wie die PDB-, SwissProt-, Pir-, DDBJ-, EMBL- und die GenBank-Datenbank. Auf diese Art und Weise wurde jedem ORF aus *G. oxydans* eine Funktion zugeordnet, basierend auf den Ähnlichkeiten zu bereits bekannten Genen und Proteinen. Wenn keine Ähnlichkeiten zu bereits bekannten Sequenzen vorhanden waren, bezeichnete das Programm den entsprechenden ORF als hypothetisch. Diese automatische Annotation ist in der Regel jedoch nicht ausreichend und mit zu vielen Fehlern behaftet. Daher war es notwendig jeden einzelnen ORF zu überprüfen. Eine Fehlerquelle der automatischen Annotation ist z. B., dass ERGO dem betrachteten ORF die Funktion desjenigen Proteins zuwies, das den höchsten Ähnlichkeitswert (Smith-Waterman-Score) zum betrachteten ORF aufwies. Gerade bei geringer Homologie, also einem niedrigen Smith-Waterman-Score, zeigte

sich jedoch, dass die dem ORF zugewiesene Funktion oft zu ungenau oder aber zu speziell war (WATERMAN, 1984).

Bei der Überprüfung der automatischen Annotation wurde die dem ORF zugewiesene Funktion mit den Funktionen der homologen Proteine verglichen. Dabei wurden zunächst die Ähnlichkeitswerte überprüft. Eine signifikante Homologie zwischen zwei Proteinsequenzen wird im allgemeinen definiert als eine Proteinidentität größer als 30 % bei einer Überlappung (*Alignment*), von mehr als 60 % der Eingabe (*query*)- und der Vergleichssequenz (*subject*). Insbesondere wurde darauf geachtet, ob sich unter den homologen Proteinen eines betrachteten *G. oxydans*-ORFs auch eines (oder mehrere) befand, das bereits biochemisch charakterisiert war. Bei einer entsprechend hohen Homologie wurde dem ORF dessen Funktion vorrangig zugewiesen.

Oft wurde auch noch einmal ein externer Sequenzabgleich mit der NCBI-Datenbank durchgeführt und das Ergebnis mit der Annotation verglichen. Gegebenenfalls erfolgte eine manuelle Änderung der bestehenden Annotation. Konnte dem ORF eine Funktion zugewiesen werden, musste dennoch zusätzlich das Startcodon überprüft werden, einschließlich der Suche nach einer vorhandenen Ribosomenbindestelle etwa 6-12 Nukleotide stromaufwärts vom Start des Gens.

Mittels der ERGO-Funktion *Pinned Regions* oder *Contig Regions* wurde die genetische Umgebung des betrachteten ORFs mit der homologer ORFs betrachtet. Dadurch konnte vor allem bei einem Operon überprüft werden, ob auch alle Gene vorhanden waren. Fehlte ein Gen bei einer solchen Struktur, kann dieses unterschiedliche Ursachen haben, die letztendlich immer zu einer erneuten Prüfung der Rohdaten führten. Zum einen kann ein Fehler in der ORF-Vorhersage vorliegen, möglich ist aber auch ein Fehler in der Konsensussequenz, was eine erneute Editierung des entsprechenden Bereichs fordert. In seltenen Fällen besteht auch die Möglichkeit, dass dieses Gen in dem betrachteten Operon tatsächlich nicht vorhanden ist.

Traten Überlappungen bei ORFs auf, musste derjenige identifiziert werden, der mit hoher Wahrscheinlichkeit ein Gen darstellte. Dabei wurden die schon erwähnten Indizien wie Homologien zu anderen ORFs, Vorhandensein einer Ribosomenbindestelle u.s.w. herangezogen. Besonders schwierig gestaltete sich diese Entscheidung bei

hypothetischen ORFs, die keine oder kaum Homologien zu bekannten Genen oder Proteinen zeigten. In diesen Fällen konnten nur allgemeine Kriterien eines Gens, wie Startcodon und Ribosomenbindestelle herangezogen werden. Bei hypothetischen und konservierten ORFs erfolgte ein Abgleich mit der COG-Datenbank (COG: clusters of orthologous genes), um diesen eventuell einer Proteinfamilie zuordnen zu können. Durch die Vorhersage transmembraner Helices (TMPred) konnte überprüft werden, ob es sich bei dem aus dem Gen resultierenden Protein um ein cytoplasmatisches oder membrangebundenes Protein handelte.

Durch diese Nachbearbeitung der automatischen Annotation wurde das Ergebnis deutlich verbessert. Vor allem der Vergleich mit externen Datenbanken war von entscheidender Bedeutung. Ist in der internen ERGO-Datenbank ein ORF einmal falsch annotiert und zeigt Ähnlichkeiten zu einem betrachteten ORF, wird ihm diese (falsche) Funktion zunächst zugewiesen. Damit findet in der internen ERGO-Datenbank eine Fehlerfortpflanzung statt. Neben der Möglichkeit zur Annotation bietet ERGO noch eine Vielzahl weiterer Funktionen zur Sequenzanalyse, so wurde z. B. auch die Stoffwechselrekonstruktion mit Hilfe der ERGO-Datenbank durchgeführt.

2.15 Bezugsquellen für Biochemikalien und Enzyme

Soweit nicht anders angegeben, wurden Chemikalien der Firmen Fluka Chemie AG (Neu-Ulm), Merck (Darmstadt), Riedel de Haen (Hannover) oder Roth (Karlsruhe) mit den Reinheitsgraden „reinst“ oder „pro analysis“ verwendet. Restriktionsendonukleasen stammten, soweit nicht anders angegeben, von den Firmen Gibco BRL (Eggenstein), MBI Fermentas (Wilna, Litauen) und Pharmacia (Freiburg). Spezielle Chemikalien und Biochemikalien wurden von folgenden Firmen bezogen:

Biomol Feinchemikalien, Hamburg:	IPTG
Biorad, München:	AquaPure GenomicDNA Kit
Difco Laboratories, Detroit:	Bacto-Agar
GERBU Biotechnik, Gaiberg:	Ampicillin, X-Gal
Gibco BRL, Eggenstein:	Agarose
MBI Fermentas, Wilna, Litauen:	DNA-Größenstandard (1 kb-Leiter), T4-DNA-Polymerase, Alkalische Phosphatase

	(CIP), T4-DNA-Ligase, PEG, Klenow-Fragment
Merck, Darmstadt:	Proteinase K, Lysozym
New England Biolabs GmbH, Schwalbach:	Restriktionsenzym <i>Bst</i> XI
Oxoid, Basingstoke, England:	Trypton, Hefeextrakt
Takara, Apen:	Ex- <i>Taq</i> -DNA-Polymerase
Qiagen, Hilden:	QIAquick Gel Extraction Kit, QIAprep Spin Mini Kit (50)
Serva Feinbiochemika, Heidelberg:	EDTA, Ethidiumbromid, SDS, Tween 20,
Sigma Aldrich Chemie, Deisenhofen:	Kanamycin Cefoxitin
Whatman & Co, Clifton, USA:	Na-N-Lauroylsarcosin, ATP, Tris

Enzyme wurden in den von den Firmen angegebenen Puffersystemen eingesetzt.

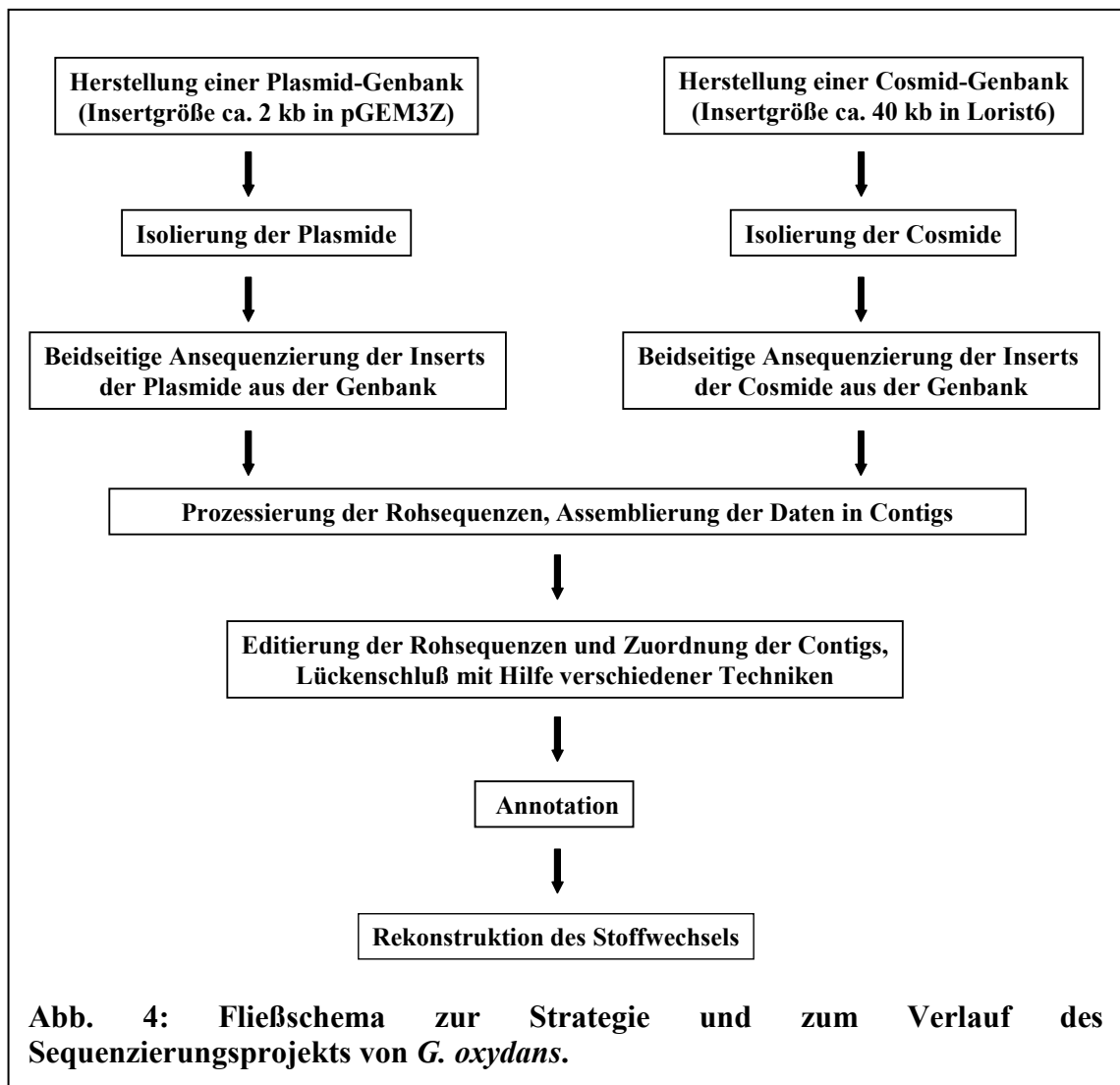
3 EXPERIMENTE UND ERGEBNISSE

Ziel dieser Arbeit war die vollständige Entschlüsselung des Genoms und die Analyse des genetischen Potentials von *G. oxydans* 621H. Die Sequenzierung des Genoms dieses industriell bedeutsamen Organismus sollte dazu beitragen, bereits bestehende biotechnologische Verfahren, die auf Enzymaktivitäten von *G. oxydans* beruhen, zu optimieren und neue fermentative Prozesse zu entwickeln. Im folgenden wird der Ablauf des Genomsequenzierungsprojekts beschrieben. Darüber hinaus werden allgemeine Sequenzmerkmale von *G. oxydans* vorgestellt und es werden Einblicke in die Analyse des Genoms gegeben. Aufgrund der Fülle der gewonnenen Daten können bei der Analyse jedoch nur ausgewählte Eigenschaften betrachtet werden.

3.1 Vorarbeiten und Überblick über die Strategie des Sequenzierungsprojekts

Die Generierung der Sequenz-Rohdaten wurde von der Firma Integrated Genomics, Chicago durchgeführt. In Abb. 4 ist ein Fließschema zur Strategie und zum Verlauf des Sequenzierungsprojekts von *G. oxydans* dargestellt. Die Arbeiten umfassten das Anlegen einer Plasmid- und Cosmid-Genbank, sowie die Isolierung und das Ansequenzieren der Plasmide und Cosmide mit Oligonukleotiden, die ihre Erkennungssequenz in der multiplen Klonierungsstelle des Vektors hatten. Die Insertgröße der Plasmide betrug etwa 2 kb und die der Cosmide etwa 40 kb. Die durchschnittlichen Sequenzläufe lieferten im Mittel Informationen von 600 bp. Das bedeutet, dass jeweils nur die Enden der Inserts sequenziert wurden. Die Sequenzierung erfolgte mit einer durchschnittlich 8fachen Abdeckung, so dass statistisch jeder Bereich des Genoms mit acht Sequenzen abgedeckt sein sollte. Das erste (automatische) Zusammenfügen (*Assembly*) der Daten wurde ebenfalls von Integrated Genomics, Chicago mit dem „Consed-Paket“ durchgeführt. Unter der Assemblierung der Daten versteht man die Anordnung von überlappenden Einzelsequenzen in sogenannte Contigs (Abb. 4). Ein Contig setzt sich somit aus mehreren überlappenden Sequenzläufen zusammen. Die Summe der Einzelsequenzen in einem Contig bildet die Konsensussequenz. Die so generierten Daten wurden nachfolgend dem Göttinger

Genomanalyse-Labor übergeben und in das Software-Paket „Staden Package“ geladen. Das Ziel in einem Genomsequenzierungsprojekt ist es letztendlich einen Contig zu erhalten, der die Sequenz des Chromosoms widerspiegelt. Falls in dem zu untersuchenden Organismus noch Plasmide vorhanden sind, stellt jedes Plasmid einen weiteren Contig dar. Um dieses Ziel für *G. oxydans* zu erreichen, mussten die Rohsequenzen verifiziert werden. Das bedeutet, dass die Anordnung der Sequenzen in den Contigs und jede einzelne Base auf ihre Richtigkeit überprüft wurden (Editierung). Anschließend war es notwendig, die Lücken zwischen den Contigs zu schließen. Im nächsten Stadium des Sequenzierungsprojekts wurden anhand der Basenabfolge ORFs definiert (s. 2.14.1), denen anschließend Funktionen zugeordnet wurden (Annotation). Auf der Basis der Annotation erfolgte nachfolgend eine Rekonstruktion des Stoffwechsels.



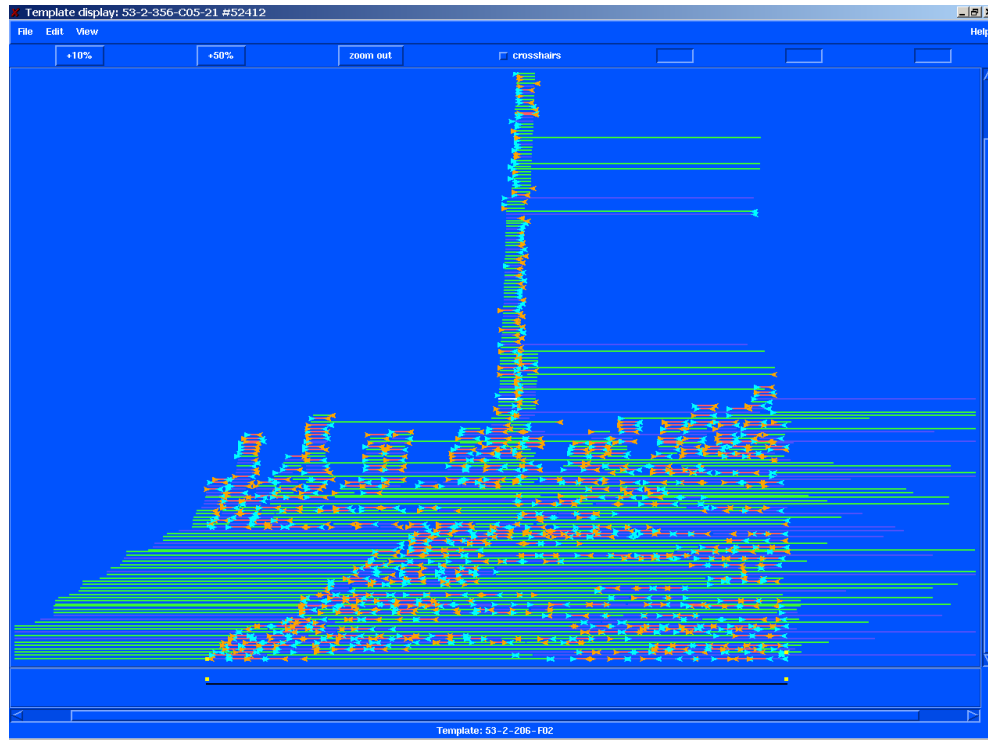
3.2 Editieren der Rohsequenzen

Die Daten der Rohsequenzierung, welche von Integrated Genomics, Chicago erstellt wurden, lieferten 53.124 Sequenzläufe, die in 138 Contigs assembliert waren. Die Gesamtqualität der Datenbank war zu diesem Zeitpunkt noch nicht zufriedenstellend. Aufgrund vieler repetitiver Sequenzen in dem Genom war ein großer Teil der Assemblierung fehlerhaft. Hinzu kam die geringe Qualität der ersten Sequenzierungen, die mit vielen Fehlern behaftet war. Ein Sequenzlauf bestand im Prinzip aus einer Vielzahl von Fluoreszenzsignalen (s. 2.10). Die Umwandlung dieser Signale in Basen erfolgte mit dem sogenannten Basecaller Phred. Dieses Programm wandelte die Fluoreszenzsignale in Basen um, zum anderen wurde jeder Base noch ein Qualitätswert zugeordnet. Die Summe der einzelnen Sequenzen, die an einer Stelle überlappten, ergab einen Qualitätswert (*base confidence value*) für die Konsensussequenz des Genoms von *G. oxydans*.

In Abb. 5 ist eine fehlerhafte Assemblierung, wie sie gehäuft in dem Datensatz zu finden war, gezeigt. Ein wichtiges Hilfsmittel, um derartige Fehler zu detektieren, war das sogenannte *Template Display* in Gap4 (Abb. 5 A). Die kleinen orangen und hellblauen Pfeile markieren jeweils die tatsächliche Sequenz, welche mit universellen *forward* bzw. *reverse* Primern generiert wurde. Die roten Linien zwischen einem *forward* und einem *reverse* Lauf symbolisieren, dass die Lage und Orientierung der korrespondierenden Läufe zueinander der erwarteten Insertgröße entsprach (2 kb bei Plasmiden, 40 kb bei Cosmiden). Ein grüner Strich zeigt, dass der entsprechende Gegenlauf in einem anderen Contig lag. Dieses bedeutete zwangsläufig, dass entweder der *forward* oder der *reverse* Lauf nicht an der richtigen Position im Genom assembliert war, was als *Misassembly* bezeichnet wird. Abb. 5 A zeigt ein solches *Misassembly* mit einem *Pile-up*-Bereich, bei dem es sich um ein Insertionselement handelt. Als *Pile-up* werden Sequenzen zusammengefasst, welche in großer Anzahl fälschlicherweise an derselben Stelle im Genom assembliert wurden. Diese Anhäufung von Daten wird durch repetitive Sequenzen hervorgerufen, z. B. durch Insertionselemente, die mehrfach im Genom auftreten und untereinander große homologe Bereiche aufweisen. Bei der automatischen Assemblierung wurden diese Sequenzläufe mit hoher Homologie zueinander an dieselbe Position des Genoms assembliert. Als Resultat entsteht dann das in Abb. 5 A gezeigte *Pile-up*. Auffallend waren in Abb. 5 A die vielen Sequenzläufe

(*Readings*), deren Gegenläufe in einem anderen Contig lagen (gekennzeichnet durch grüne Striche). Da es sich hier um eine repetitive Sequenz handelte, war die Qualität der Konsensussequenz natürlich sehr gut.

A:



B:



Abb. 5: Misassembly in den Rohdaten.

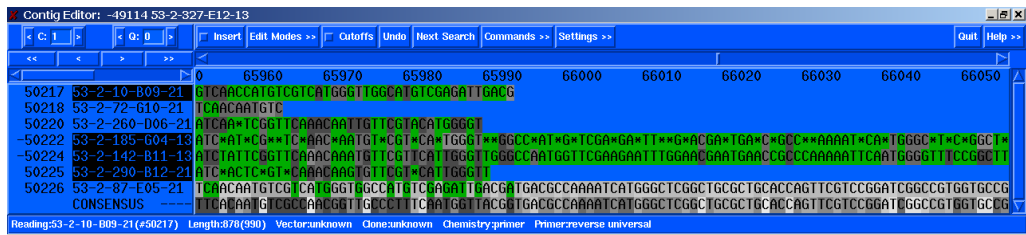
A: Dargestellt ist ein sogenanntes *Pile-up* im *Template Display*, eine Überlagerung von ähnlichen Sequenzen, die aufgrund einer Fehlassemblierung den gezeigten Turm von Sequenzen bilden. B: Die Abbildung veranschaulicht den zu A gehörenden *Contig Editor*, also die Sequenzen, die sich hinter dem *Pile-up* verbergen. Aufgrund der Überlagerung der vielen Sequenzen konnte jedoch nur ein Teil dargestellt werden.

Abb. 5 B stellt die überlappenden Sequenzen aus dem *Pile-up* Bereich dar. Gezeigt ist der sogenannte *Contig Editor* aus Gap4. Links sind jeweils die Bezeichnungen der Sequenzen angegeben, die aus einer Zahlenabfolge bestehen. Die unterste Sequenz ist die Konsensussequenz, welche sich aus den darüber gezeigten Einzelsequenzen zusammensetzt. Dabei ist jede Base in einem bestimmten Grauton hervorgehoben. Diese Markierung symbolisiert die durch den Basecaller Phred zugeordnete Qualität. Je dunkler der Grauton, desto schlechter die Qualität und desto höher war die Wahrscheinlichkeit, dass es sich bei dem zugrunde liegenden Fluoreszenzsignal um ein unspezifisches Signal handelte. Die Sterne in Abb. 5 B sind sogenannte Platzhalter, die immer dann entstanden, wenn die Einzelsequenzen nicht zu 100 % homolog waren. Derartiges kann z. B. aus qualitativ schlechten Sequenzen resultieren. Bei der Auswertung können die Fluoreszenzsignale nicht richtig zugeordnet werden und Sequenz A weist dann z. B. an einer Position die Basenabfolge „CCC“ auf, während in Sequenz B an derselben Position die Basenabfolge „CC“ zu finden ist. Beim Übereinanderlegen der beiden Sequenzen wird Sequenz B dann als „C*C“ dargestellt. Es sei an dieser Stelle noch darauf hingewiesen, dass in Abb. 5 B nur ein Teil der übereinanderliegenden Sequenzen gezeigt ist. An der Position des *Pile-ups* aus Abb. 5 A befanden sich mehr als 50 Sequenzen mit Homologien zueinander, die aufgrund ihrer Ähnlichkeit bei der automatischen Assemblierung fälschlicherweise übereinander gelegt wurden. Abb. 5 B zeigt also nur einen kleinen Ausschnitt aus dem betreffenden Bereich.

Auf der anderen Seite bedeutete die gehäufte Übereinanderlagerung von Daten aber auch, dass es innerhalb des Genoms Sequenzbereiche gab, deren Abdeckung entsprechend gering war (Abb. 6). Wenn viele Sequenzen aufgrund der hohen Ähnlichkeit zueinander fälschlicherweise an dieselbe Position im Genom assembliert werden, bedeutet dies, dass es Bereiche in dem Genom gibt, an denen genau diese Sequenzen fehlen. Ein solcher Bereich ist in Abb. 6 dargestellt. In grün sind Basen hervorgehoben, die Unterschiede zu der Konsensussequenz aufwiesen. Einige Sequenzläufe sind in Abb. 6 A schwarz markiert. Drei dieser Sequenzen sind in Abb. 6 B durch einen Pfeil gekennzeichnet. Weiter oben wurde beschrieben, dass eine rote Linie zwischen einem korrespondierenden *forward* und *reverse* Lauf darauf hinweist, dass die Lage und Orientierung der *Readings* zueinander den Erwartungen entsprach. Es wurde daher angenommen, dass ein durch eine rote Linie verbundener *forward* und

reverse Lauf das gesamte Insert des entsprechenden rekombinanten Plasmides widerspiegelte, welches eine Größe von 2 kb aufwies. In Abb. 6 A ist zu erkennen, dass die Qualität der Konsensussequenz sehr schlecht war (erkennbar an der dunkelgrauen Markierung). Dennoch schien im *Template Display* die Assemblierung dieser Sequenzläufe korrekt zu sein. Bei einer fehlerhaften Assemblierung hätte man erwarten, dass die in 6 B durch kleine Pfeile dargestellten Sequenzläufe von einer grünen Linie gefolgt würden, wie es in Abb. 6 A (das *Pile-up*) zu sehen ist.

A:



B:

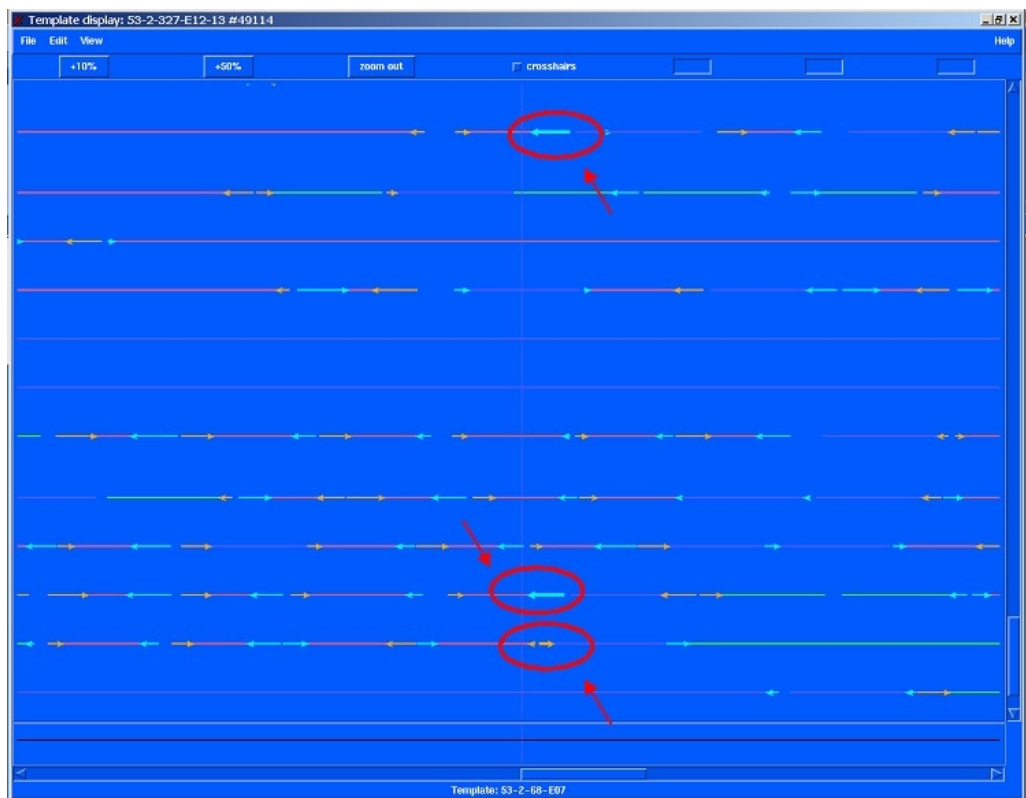


Abb. 6: Bereich schlechter Abdeckung aus Gap4.

In Abb. 6 A ist der *Contig Editor* gezeigt, die Einzelläufe passen nicht zu der Konsensussequenz (grün dargestellt). Teil B zeigt das dazugehörige *Template Display*. Die Läufe, die in A schwarz unterlegt sind, sind hier durch einen Pfeil markiert.

Beide Extreme traten in dem Datensatz mehr als 50 Mal auf und mussten durch Wiederholungen von Sequenzierläufen, *Walkings* und mühevoller Sequenzierarbeit aufgelöst werden. Wiederholungen von *Readings* dienten dazu, Sequenzen mit schlechter Qualität zu überprüfen, die Qualität zu verbessern und Fehler zu beheben. Der Nachteil an Wiederholungen ist, dass diese mit universellen Primern durchgeführt werden mussten, deren Bindestelle in der multiplen Klonierungsstelle des Vektors lag. Das bedeutete, dass damit immer nur die Enden der Inserts sequenziert werden konnten. Bei *Walkings* wurden Primer abgeleitet, die innerhalb der Inserts banden, so dass bestimmte Positionen des Genoms genau überprüft werden konnten.

Anzumerken sei auch noch, dass die *Misassemblies*, wie sie bei *G. oxydans* auftraten, nicht dem typischen Bild einer Fehlassemblierung entsprachen. Erwarten würde man einen Sequenzbereich, bei dem einzelne Basen (guter Qualität) nicht zu der Konsensussequenz passen. Die entsprechenden *Readings* würde man dann aus dem jeweiligen Contig extrahieren. Fakt war jedoch, dass oft nur mit Hilfe von spezifischen Primern und anschließender PCR-Reaktion bestimmte Sequenzbereiche überprüft werden konnten.

Nach Abschluss der ersten Editierungsrunde lagen 54.167 Einzelsequenzen vor, die in 134 Contigs assembliert waren. Die *Pile-ups* waren zum größten Teil aufgelöst. In der zweiten Phase des Genomprojekts wurde damit begonnen, die Zahl der Contigs weiter zu reduzieren und die Qualität der Konsensussequenz zu verbessern. Statistisch gesehen sollte am Ende der Sequenzierung die Fehlerrate in der genomischen Sequenz nicht höher als 1/100.000 Basen sein. Nur unter diesen Bedingungen war gewährleistet, dass es sich um verlässliche Daten handelte.

3.3 Lückenschluss

Nach Abschluss der ersten Editierungsrunde war die Gesamtqualität der Datenbank immer noch nicht zufriedenstellend, da statistisch gesehen die Fehlerrate bei 1/101 lag. Die weitere Verbesserung der Konsensussequenz erfolgte zeitgleich mit dem Schließen der ersten Lücken. Dieses Stadium der Sequenzierung gliederte sich in mehrere Phasen: In der ersten Phase wurden die Lücken mittels des BLASTN-Algorithmus (s. 2.12.1) und durch *Walkings* auf Plasmiden geschlossen. In der zweiten Phase wurde mit Hilfe

der Cosmidgenbank ein sogenannter Supercontig generiert, wobei die einzelnen Contigs so angeordnet wurden, dass jeweils die Contigenden, die eine Lücke potentiell überspannten, nebeneinander lagen.

3.3.1 Lückenschluss durch *Walkings* und BLASTN

Wie in 2.12.2 beschrieben, wurden jeweils an den Contigenden spezifische Primer abgeleitet. Mit diesen Primern wurden *Walkings* durchgeführt, wobei Plasmide als *Template* dienten. Eine schematische Darstellung für diese Lückenschluss-Strategie ist in Abb. 7 gezeigt. Ein Contig ist in Abb. 7 durch einen schwarzen Strich symbolisiert, die gelben Quadrate stellen die abgeleiteten Primer dar. Plasmide, die als *Template* dienten, sind in dunkelblau und orange dargestellt. Die rote Linie dazwischen symbolisiert, dass die Plasmide sich auf unterschiedlichen Contigs befanden. Allerdings entsprach die Insertgröße bei einer derartigen Anordnung der beiden Contigs zueinander den Erwartungen.

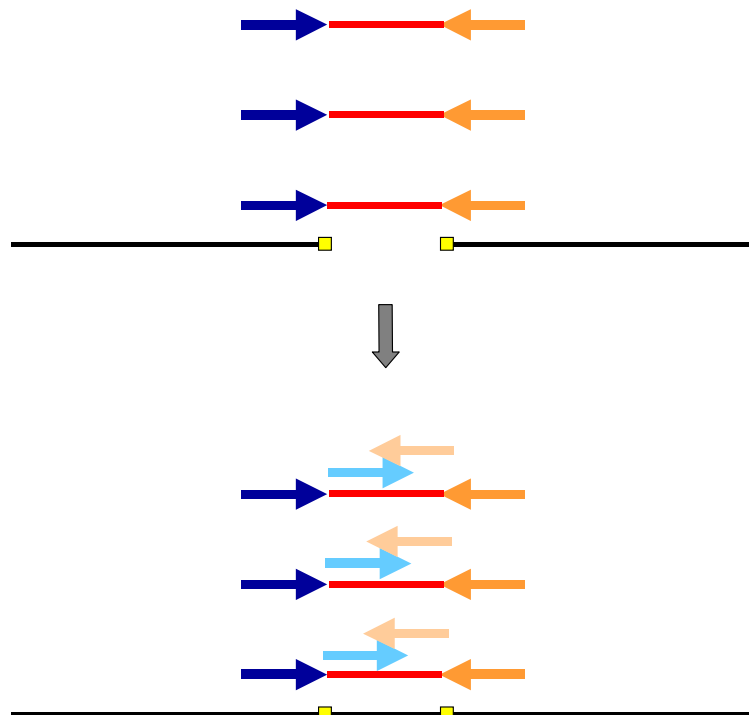


Abb. 7: Schematische Darstellung für die Strategie zum Schließen einer Lücke durch *Walkings*.

Der schwarze Strich symbolisiert jeweils einen Contig, Primer sind als gelbe Quadrate dargestellt. Die Pfeile symbolisieren jeweils einen Sequenzlauf.

Das Ziel dieser Strategie war, den jeweiligen Contig so zu vergrößern, dass eine Lücke geschlossen werden konnte. Das war dann der Fall, wenn neue Sequenzen aus zwei verschiedenen Contigs überlappten. Ein Beispiel für derartige neue Sequenzen ist in Abb. 7 in hellblau, bzw. hellorange dargestellt. Für die *Walkings* wurden von jeder Seite des entsprechenden Contigs 3-5 unterschiedliche Plasmide ausgewählt, die als Matrize (*Template*) dienten. Bei einer Überlappung der neu erhaltenen Sequenzen erhielt man anschließend für die jeweilige Lücke eine 6-10fache Abdeckung. Auf diese Art und Weise und mit Hilfe des BLASTN-Algorithmus konnte die Zahl der Lücken auf 20 reduziert werden, die durchschnittliche Größe eines Contigs betrug 150.000 bp. Beim Einsatz des BLASTN-Algorithmus wurden die Enden der Contigs mit der gesamten Datenbank abgeglichen. Zeigte dieser Abgleich Homologien zu schon bestehenden Sequenzen in der Datenbank, die z. B. aus Einzelläufen bestanden, konnten die entsprechenden Contigs zunächst verlängert und anschließend die Lücke zwischen zwei Contigs geschlossen werden.

3.3.2 Erstellen eines „Supercontigs“

Nachdem die Möglichkeiten zum Reduzieren der Lücken durch einfachen BLASTN-Abgleich und *Walkings* auf Plasmiden erschöpft waren, mussten andere Methoden herangezogen werden, um das Genom zu schließen. Hier kam die bereits vorhandene Cosmidgenbank zum Einsatz. Sämtliche Contigenden wurden dahingehend überprüft ob, und wenn ja wohin Cosmide aus diesem Contig herausführten. Das Ziel war es in dieser Sequenzierungsphase die verbleibenden Contigs so anzuordnen, dass die jeweilige Position im Genom richtig war und „lediglich“ die verbleibenden Lücken mit Sequenz gefüllt werden mussten. Abb. 8 zeigt die beiden „Supercontigs“, die mit dieser Strategie erstellt werden konnten. Dabei wurden 15 der noch bestehenden 20 Contigs integriert. Für die verbleibenden Contigs gab es zu der Zeit noch keine Anhaltspunkte bezüglich der richtigen Assemblierung im Genom von *G. oxydans*. Die schwarzen Striche symbolisieren jeweils einzelne Contigs. Zu diesem Zeitpunkt des Projekts existierten aber schon viele PCR-Produkte, welche die jeweilige Lücke überspannten. Dieses ist ebenfalls in der Abb. 8 vermerkt, inklusive der Größe der entsprechenden PCR-Produkte. An den Contigenden, die mit einem hellgrauen Kreis versehen sind, lagen bereits spezifische Primer vor, die PCR-Produkte mussten jedoch noch generiert werden. Problematisch waren die Bereiche, die mit einem weißen Kreis markiert sind.

An diesen Positionen gab es noch keine Anhaltspunkte, um den Contig zu verlängern. Die Zahlen oberhalb und unterhalb der Contigs geben die Größe der DNA Sequenz beziehungsweise die Bezeichnung der Contigs an. Zu diesem Zeitpunkt war der größte Contig bereits 870.992 bp groß.

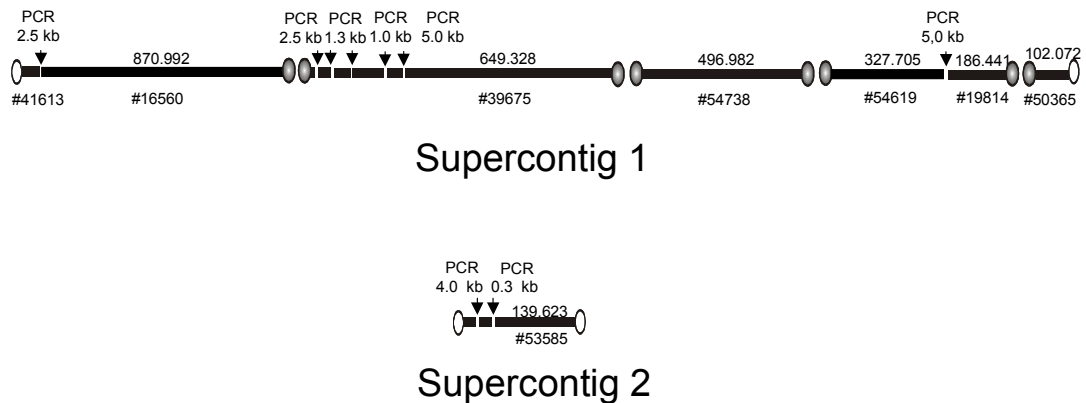


Abb. 8: Erstellte „Supercontigs“

Die oben dargestellte Vorgehensweise war jedoch nicht so trivial wie es zunächst den Anschein hatte. Auf sehr vielen Contigs gab es zwar Cosmide, deren Gegenlauf in einem anderen Contig lag, aber oft bedurften die Befunde noch weiterer Interpretation. Es war keineswegs so, dass sämtliche Gegenläufe aus Contig A in Contig B assembliert waren. Dieses konnte mehrere Gründe haben: Die durchschnittliche Insertgröße der Cosmide betrug 40 kb. Befand sich die Sequenz des betreffenden Cosmides am Ende eines Contigs, konnte die Lücke also bis zu 40 kb groß sein. Das bedeutet, dass der entsprechende Gegenlauf nicht unbedingt im benachbarten Contig liegen musste, sondern dass sich durchaus noch andere Contigs innerhalb dieser Lücke befinden konnten. Darüber hinaus war auch nicht jede Lücke gleichwertig. Besonders repetitive Bereiche des Chromosoms bereiteten Schwierigkeiten beim *Gap closure*. Dazu gehörten unter anderem Transposons, IS-Elemente aber auch rRNA-Bereiche. Die dazugehörigen Sequenzen waren sehr konserviert und tauchten in verschiedenster Kopienzahl in dem Genom auf. Die Schwierigkeit bestand vor allem darin, die flankierenden Sequenzen solcher repetitiven Bereiche richtig zu assemblieren. Zeigten also die Gegenläufe der betrachteten Cosmide eines Contig auf unterschiedliche Contigs, konnte dies auch ein Hinweis auf noch vorhandene *Misassemblies* sein.

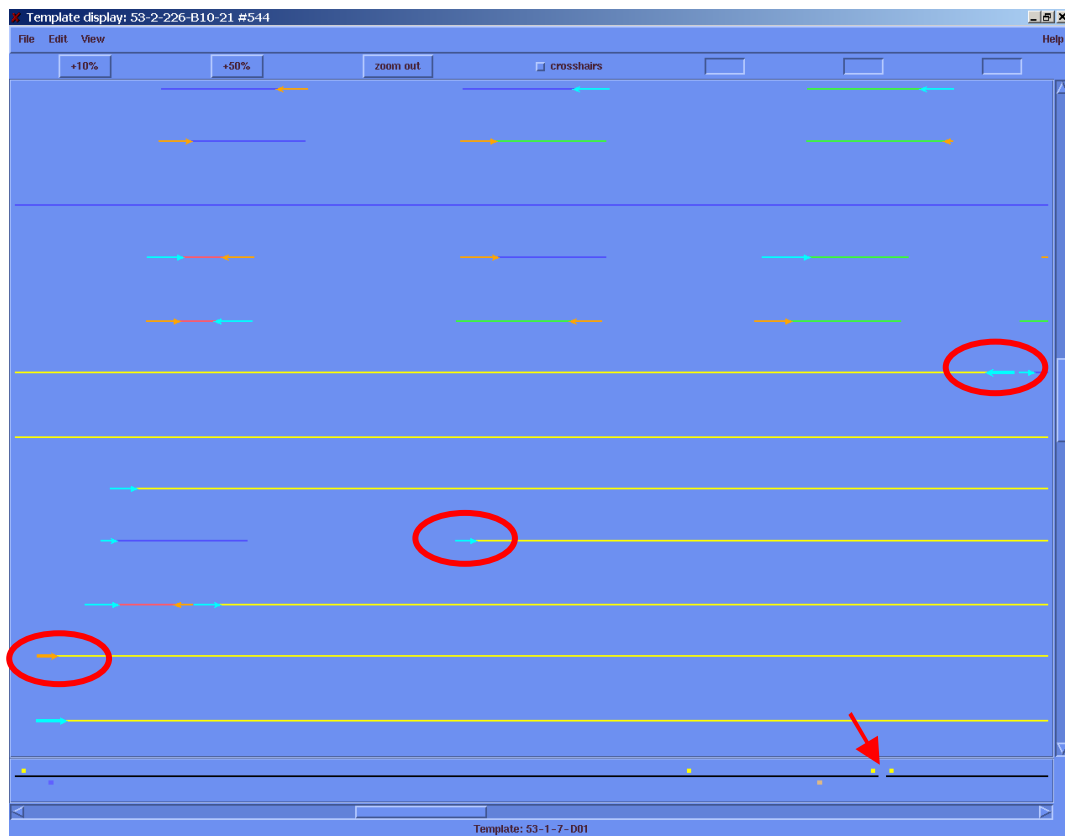


Abb. 9: Auswahl geeigneter Cosmide als PCR-Template

Dargestellt sind zwei nebeneinander liegende Contigs mit Hilfe des *Template Displays*.

Die Strategie, um die noch bestehenden Lücken innerhalb des Supercontigs zu schließen, soll anhand von Abb. 9 veranschaulicht werden. Im *Template Display* sind zwei nebeneinander liegende Contigs dargestellt. Der rote Pfeil zeigt auf die Lücke zwischen diesen beiden Contigs, die gelben Punkte rechts und links neben dem Pfeil symbolisieren, dass an dieser Stelle bereits Primer abgeleitet waren. Die Lücke sollte dadurch überwunden werden, dass eine PCR-Reaktion mit den an den Contigenden abgeleiteten Primern durchgeführt wurde. Als *Template* für die PCR wurden neben chromosomaler DNA auch Cosmide eingesetzt. Cosmide, die als *Templates* für die PCR ausgewählt wurden, sind rot eingekreist. Die gelben Striche zwischen *forward* und *reverse Readings* der Cosmide symbolisieren, dass bei dieser Anordnung der beiden Contigs die Insertgröße den Erwartungen entsprach. Bei der Auswahl geeigneter Cosmide für *Walkings* wurde zum einen darauf geachtet, dass sämtliche Gegenläufe der ausgewählten Cosmide in demselben Contig lagen. Darüber hinaus sollten sich die

entsprechenden *forward* oder *reverse* Läufe nicht an derselben Position innerhalb des Genoms befinden. Damit sollte die Gefahr, dass ein *Misassembly* vorlag, minimiert werden. War die Lücke nicht größer als 8 kb und die in Abb. 9 dargestellte Anordnung der Contigs korrekt, erhielt man also im besten Fall verschiedene PCR-Produkte, die alle die Lücke überspannen sollten. Eine Sequenzierung der PCR-Produkte von beiden Seiten lieferte als Optimum eine 8-fache Abdeckung der Lücke. Lagen große PCR-Produkte vor, mussten auf diesen Produkten noch *Walkings* durchgeführt werden, um die vollständige Sequenz zu erhalten.

3.3.3 Kombinatorische Ansätze

Mit Hilfe der unter 3.3.2 beschriebenen Methode konnten 15 Lücken geschlossen werden. Darüber hinaus konnten bereits vier Plasmide der Größe 2,7 kb, 13,2 kb, 14,5 kb und 163 kb innerhalb des Genoms identifiziert werden. Das Auffinden und die Analyse der Plasmide in dem Genom von *G. oxydans* 621H wird im folgenden Kapitel (3.4) behandelt. Für die verbleibenden fünf Lücken wurden von sämtlichen Contigenden Primer abgeleitet und diese für kombinatorische PCR eingesetzt. Die Schwierigkeit dabei war, dass die vorhandenen Contigs fast ausschließlich mit repetitiven Sequenzen endeten. Somit konnten die Primer nicht direkt von den Enden abgeleitet werden, da gewährleistet sein musste, dass der jeweilige Primer spezifisch war, also nur ein einziges Mal im Genom band. Es wurden daher Sequenzen für die Primer gewählt, die sich weiter innerhalb der Contigs befanden.

Die abgeleiteten Primer wurden anschließend für eine kombinatorisch PCR eingesetzt (s. 2.12.3.2), wobei genomische DNA von *G. oxydans* als *Template* diente. Dabei wurden zwei PCR-Produkte der Größe 0,8 kb und 1 kb erhalten, mit denen zwei weitere Lücken geschlossen werden konnten.

Nachdem erneut die verbliebenen Contigenden mit der gesamten Datenbank abgeglichen worden waren, konnten als Resultat auch kleinere Contigs in den Supercontig integriert werden, so dass neue Anhaltspunkte erhalten wurden, um das Genom zu schließen. Darüber hinaus zeigte sich dabei, dass das Genom von *G. oxydans* 621H ein weiteres Plasmid der Größe 26,6 kb enthielt. Der erneute Abgleich der Contigenden mit der Gesamtdatenbank und die Assemblierung einzelner Sequenzen an

die Enden der Contigs ermöglichte es, auch mit Cosmiden als *Template* eine PCR durchzuführen.

3.3.4 Abschluss der Sequenzierung

Nachdem der Datensatz wie unter 3.3.1, 3.3.2 und 3.3.3 beschrieben, bearbeitet worden war, verblieb immer noch eine Lücke. Unter der Annahme, dass die Assemblierung der Datenbank richtig war, gab es demnach nur noch genau zwei Contigenden, die zusammengeführt werden mussten. Bei einem Blast-Abgleich der entsprechenden Sequenzen stellte sich heraus, dass es sich bei der verbliebenden Lücke um einen Bereich ribosomaler RNA-Gene handelte. Auffallend war, dass für einen so konservierten Bereich keine Sequenzen vorlagen, um die Lücke zu schließen. Die Lösung des Problems lag darin, dass die ribosomalen RNA-Gene auch zu *E. coli* homolog waren. Beim Screening innerhalb des Pregap4 (s. 2.11.2) auf *E. coli*-Sequenzen wurden diese Daten dadurch als Kontamination angesehen und entsprechend verworfen. Da die Rohsequenzierung und die erste automatische Assemblierung der Sequenz in Chicago durchgeführt wurde, war es unmöglich die verlorenen Daten nachträglich in die Datenbank wieder einzuspeisen. Eine erneute PCR der Contigenden lieferte jedoch ein eindeutiges Produkt und führte nach Sequenzierung zum Schluss der letzten Lücke. Anschließend wurden sämtliche repetitiven Bereiche des Genoms, wozu auch die rRNA-Gencluster gehörten, erneut eingehend überprüft. Innerhalb des Genoms traten vier nahezu identische Kopien von rRNA-Operons auf. Abb. 10 zeigt den typischen Aufbau eines rRNA-Operons aus *G. oxydans*. Ein Operon bestand jeweils aus einem Gen für die 16S, 23S und 5S rRNA. Zwischen den ersten beiden Genen lag jeweils eine Spacer-Region, die etwa 660 bp lang war. In allen vier rRNA Clustern befanden sich innerhalb dieser Spacer-Region jeweils zwei Gene für tRNAs. Um sicherzustellen, dass alle rRNA-Operone richtig assembliert waren, wurde jeweils das gesamte Operon mittels PCR amplifiziert. Wenn möglich wurden dabei spezifische Primer ausgewählt und genomische DNA von *G. oxydans* als *Template* eingesetzt. Teilweise waren die RNA-Operons jedoch von weiteren repetitiven Sequenzen flankiert (Transposons). In diesen Fällen mussten die Bereiche wieder mit Cosmiden als *Template* amplifiziert werden. Ziel war dabei vor allem die Überprüfung der Bereiche, welche die repetitiven Sequenzen flankieren. Bei einer falschen Assemblierung erhielt man mit spezifischen Primern zum einen kein PCR-Produkt, da dann die Orientierung

der Primer zueinander falsch war, zum anderen war es bei einer Fehlasmblierung nicht möglich, die erhaltenen Sequenzen in die entsprechende Stelle des Datensatzes einzufügen. Die Überprüfung lieferte jedoch das gewünschte Resultat und zeigte somit, dass die Assemblierung der vier rRNA Cluster richtig war.

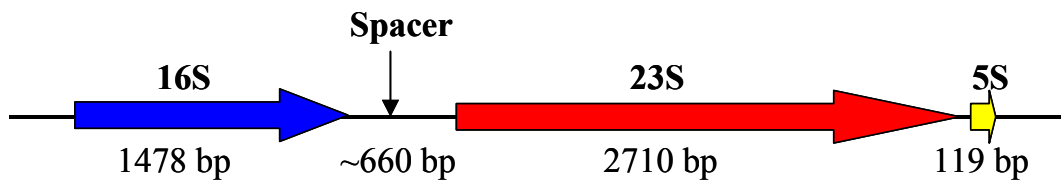


Abb. 10: Aufbau eines typischen rRNA-Operons in *G. oxydans*.

Darüber hinaus wurden sämtliche Bereiche des Genoms, die repetitive Sequenzen, wie z. B. Transposons enthielten, markiert. Die Überprüfung erfolgte zum einen anhand des *Template-Display*, indem kontrolliert wurde, ob die entsprechenden *forward* und *reverse* Sequenzläufe in der richtigen Orientierung zueinander waren, und ob es Inserts in Plasmiden oder Cosmiden gab, die diesen Bereich vollständig überspannten. Zur Veranschaulichung ist dazu in Abb. 11 ein Schema dargestellt. Gezeigt ist ein als schwarzer, horizontaler Strich dargestellter Contig; der repetitive Bereich ist rot markiert und von zwei vertikalen Linien umrahmt. Die entsprechenden Inserts sind nummeriert und weisen erwartungsgemäß eine Größe von 2 kb auf. Dennoch konnte nur anhand der Inserts 2, 4 und 7 eine Aussage darüber getroffen werden, ob die Assemblierung korrekt war. Bei den anderen dargestellten Inserts war zwar jeweils der *forward* oder *reverse Reading* in einer spezifischen Region des Contigs assembliert, sie ließen jedoch keine Aussage darüber zu, ob die flankierenden, spezifischen Bereiche des Repeats richtig assembliert waren. Der in der Abb. 11 rot markierte Bereich ließ sich innerhalb des Genoms an mindestens 50 Positionen assemblieren. Entscheidend war es deshalb, dass der repetitive Bereich von einem einzigen Insert überspannt wurde, wobei sowohl *forward*, als auch *reverse Reading* in einer spezifischen Region assembliert sein mussten. Diese Bedingungen waren jedoch nur bei den Insert 2, 4 und 7 gegeben (Abb. 11). Darüber hinaus wurden die entsprechenden Bereiche mittels PCR amplifiziert. Die Sequenzierung der PCR-Produkte und die Assemblierung dieser

Sequenzen in die Datenbank zeigte, das die Anordnung der repetitiven Bereiche richtig war.

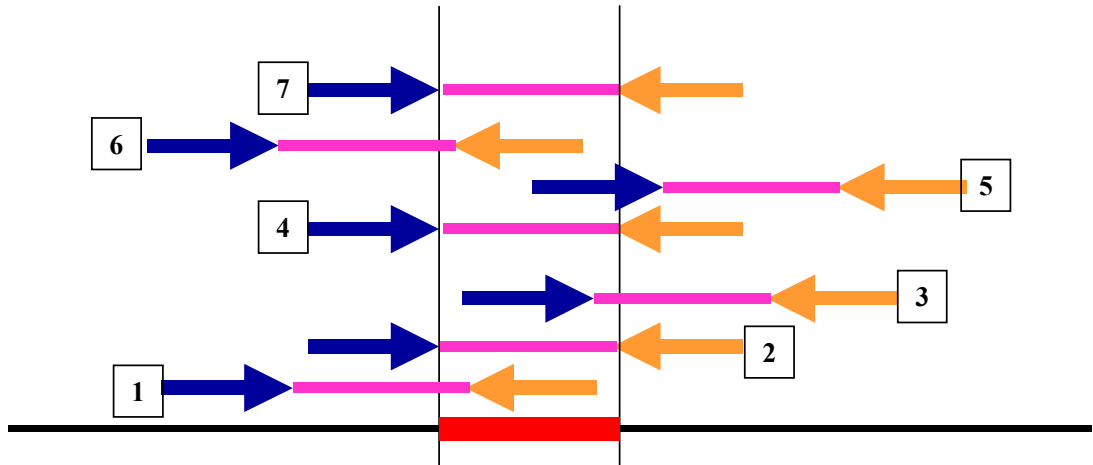


Abb. 11: Schematische Darstellung des Template Displays zur Überprüfung repetitiver Bereiche.

Die Pfeile symbolisieren jeweils die einzelnen Sequenzläufe, generiert mit universellen *forward* bzw. *reverse* Primern.

Mit dem Schließen der letzten Lücke und der eingehenden Prüfung, sowohl der rRNA Cluster als auch anderer repetitiver Bereiche wie IS-Elemente, war das Genom von *G. oxydans* 621H erfolgreich vollständig sequenziert. Das zirkuläre Chromosom dieses Organismus bestand aus 2,7 Mb. Darüber hinaus konnten fünf extrachromosomale Elemente identifiziert werden: Ein Megaplasmid der Größe 163 kb (pGOX1) sowie vier Plasmide der Größe 27 kb (pGOX2), 14,5 kb (pGOX3), 13,2 kb (pGOX4) und 2,7 kb (pGOX5). Es erfolgte noch das sogenannte *Polishing*, ein Aufbessern der Konsensussequenz. Nach Abschluss dieser Arbeiten hatte das Chromosom eine Fehlerrate von 1/289.582. Die Plasmide wiesen nach dem abschließenden *Polishing* eine Fehlerrate von 1/5.677.600 (pGOX1), 1/1.491.942 (pGOX2), 1/197.281 (pGOX3), 1/138.575 (pGOX4) und 1/1.082.316.574 (pGOX5) auf. Die Fehlerrate sollte einen Wert von 1/100.000 nicht unterschreiten, um die Korrektheit der Daten zu gewährleisten. In den folgenden Kapiteln der vorliegenden Arbeit werden sowohl die Plasmide als auch das Chromosom von *G. oxydans* vorgestellt.

3.4 Identifizierung von Plasmiden

Zu Beginn dieser Arbeit war bereits bekannt, dass *Gluconobacter*-Stämme häufig Plasmide enthalten. Der hier sequenzierte Stamm *G. oxydans* 621H wurde in dieser Hinsicht mittels Puls-Feld-Gelelektrophorese analysiert (VERMA et al., 1997), wobei ein Plasmid mit der Größe 21 kb identifiziert wurde. Dieser Befund konnte im Rahmen des Sequenzierungsprojekts nicht bestätigt werden. Bei der Bearbeitung des Genoms mit Hilfe des Gap4 Programms wurden jedoch insgesamt fünf Plasmide unterschiedlicher Größe identifiziert.

3.4.1 Auffinden der Plasmide mit Hilfe von Gap4

Von der BASF AG (Ludwigshafen) lagen Informationen vor, dass ein *Gluconobacter oxydans*-Stamm, der zur Produktion von Vitamin C eingesetzt wird, ein 2,7 kb großes Plasmid trägt. Von diesem Vektor lag der BASF AG auch bereits eine Sequenz vor. Ein BLAST-Abgleich mit der Plasmid-Sequenz der BASF AG zeigte, dass dieses Plasmid auch in dem sequenzierten Stamm 621H existiert. Die Sequenz lag bereits nach der ersten automatischen Assemblierung richtig arrangiert in einem Contig vor. Das Auffinden der weiteren Plasmide war dagegen aufwendiger.

In der Genomassemblierungssoftware Gap4 zeigte sich in der Ansicht der Plasmide ein charakteristisches Bild. Bei der linearisierten Darstellung des Contigs war eine zirkuläre Anordnung der einzige Weg, die Lücke zu schließen. Darüber hinaus zeigte sich im *Template Display* des Gap4 ein typisches Bild, das am Beispiel von pGOX5 in Abb. 12 dargestellt ist. In dieser Abb. sind die jeweiligen *forward* und *reverse Reading* an den Enden des Contigs durch einen schwarzen Strich verbunden. Dieses Schema symbolisiert, dass bei einer derartigen Assemblierung der Sequenzläufe die Anordnung der *Inserts* zueinander nicht richtig war. In Gap4 war ausschließlich eine lineare Darstellung der Contigs möglich. Dadurch liefen *forward* und *reverse Reading* nicht aufeinander zu, sondern entfernten sich voneinander, wenn ein zirkuläres Plasmid vorlag. Erst bei einer zirkulären Anordnung war die richtige Orientierung der *Readings* zueinander gewährleistet.

Insgesamt wurden in dem Genom von *G. oxydans* vier Plasmide der Größen 26,6 kb (pGOX2), 14,5 kb (pGOX3), 13,2 kb (pGOX4), 2,7 kb (pGOX5) und ein Megaplasmid (pGOX1) der Größe 163 kb identifiziert. Zur Verifizierung wurden die Plasmide anschließend isoliert und durch geeignete Restriktionsverdau überprüft (s. 3.4.2).

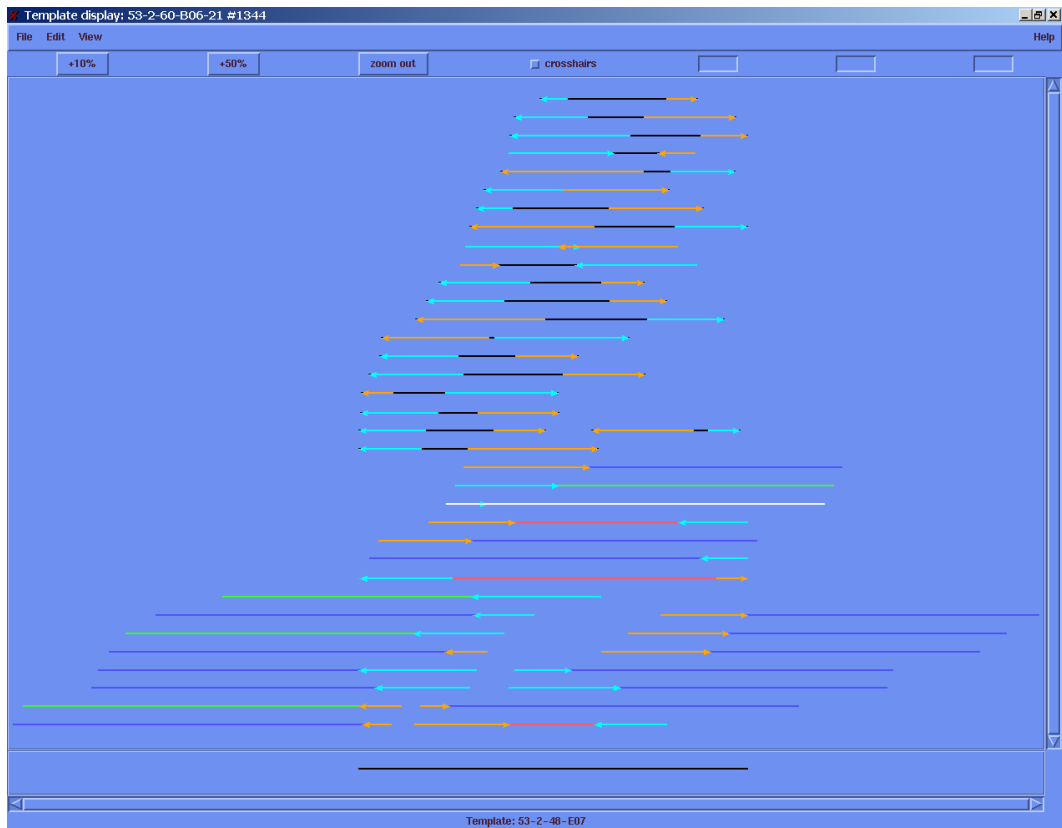


Abb. 12: Darstellung des *Template Display* von pGOX5

3.4.2 Isolierung der identifizierten Plasmide

Die anhand der Genomdaten identifizierten Plasmide sollten zur weiteren Untersuchung isoliert und im Agarosegel analysiert werden. Die Plasmide wurden mittels Minipräparation (s. 2.6.2) isoliert, wobei nach der Vorschrift für Cosmide und *Low Copy* Plasmide vorgegangen wurde. Zusätzlich wurde die isolierte Plasmid-DNA aufkonzentriert (s. 2.5.2.2). Abb. 13 A zeigt die unverdauten Plasmide. Hier ließ sich eine hochmolekulare Bande oberhalb von 10 kb nachweisen. Darüber hinaus waren drei weitere Banden bei ca. 9 kb, 2,9 kb und 1,9 kb zu erkennen. Da ungeschnittene

Plasmide in mehreren Erscheinungsformen auftreten, war eine Aussage über die Größe hier jedoch nicht möglich. Schon bei der ungeschnittenen Plasmid-DNA lag jedoch die Vermutung nahe, dass es in dem Genom mehrere Plasmide gab.

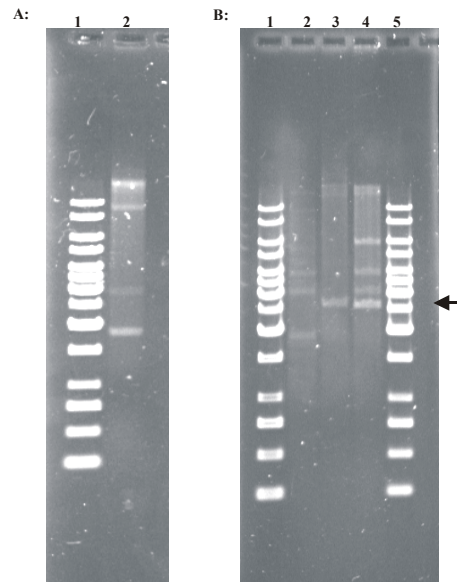


Abb. 13: Agarosegelelektrophorese der isolierten Plasmide aus *G. oxydans* 621H

A: Unverdaute Plasmide. Spur 1: 1 kb DNA-Ladder, Spur 2: Unverdaute Plasmide.
 B: Verdau der Plasmide. Spur 1: 1 kb DNA-Ladder, Spur 2: Verdau mit *Bam*HI, Spur 3: Verdau mit *Not*I, Spur 4: Verdau mit *Xho*I, Spur 5: 1 kb DNA-Ladder. Das linearisierte Plasmid pGOX5 ist durch einen Pfeil markiert.

Zur weiteren Untersuchung wurden mit den Plasmiden Restriktionsverdauungen mit *Bam*HI, *Not*I und *Xho*I durchgeführt. Das Ergebnis ist in Abb. 13 B gezeigt. Das 2,7 kb große Plasmid pGOX5 wurde durch *Not*I und *Xho*I linearisiert, *Bam*HI schnitt nicht in diesem Plasmid. Betrachtet man den *Bam*HI-Verdau (Abb. 13 B, Spur 2) der Plasmide erkennt man, dass sich das Bandenmuster im Vergleich zu der ungeschnittenen DNA deutlich änderte: Die hochmolekulare Bande oberhalb von 10 kb und die Bande bei ca. 9 kb waren verschwunden. Bei 4 kb, 3 kb und 2,9 kb waren deutliche Banden erkennbar. Der Verdau mit *Not*I (Abb. 13 B, Spur 3) und *Xho*I (Abb. X b, Spur 4) lieferte eine deutliche Bande bei 2,7 kb, was den Erwartungen für das linearisierte Plasmid pGOX5 entsprach (in Abb. 13 B mit einem Pfeil markiert). In Spur 4 (Verdau mit *Xho*I) waren darüber hinaus Banden auf der Höhe von 3 kb, 4 kb, 6 kb sowie eine hochmolekulare Bande zu erkennen. Durch die Verdauungen mit *Bam*HI und *Xho*I konnte das Plasmid pGOX5

eindeutig nachgewiesen werden. Durch die Linearisierung zeigte es die erwartete Größe von 2,7 kb. Die anderen Plasmide konnten hingegen nicht derartig eindeutig identifiziert werden, auf dem Gelfoto war jedoch zweifelsfrei zu erkennen, dass *G. oxydans* mehrere Plasmide trägt.

Zur weiteren Überprüfung der Plasmide pGOX1-4 wurden nachfolgend PCR-Reaktionen durchgeführt. Dafür wurde zunächst die folgende Annahme aufgestellt: Wie oben erwähnt, war in Gap4 nur eine lineare Anordnung von Plasmiden möglich. Dies wurde ausgenutzt, um an beiden Enden der linearisierten Plasmide Primer abzuleiten, die nach der PCR Reaktion ein Produkt definierter Größe ergeben sollten. Das Ergebnis dieses Vorgehens ist in Abb. 14 dargestellt.

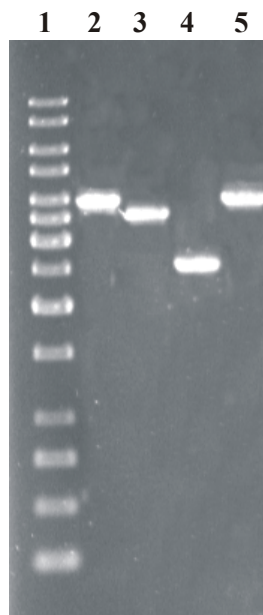


Abb. 14: Agarosegelelektrophorese der PCR-Produkte von pGOX1-4.

Spur 1: 1 kb DNA-Ladder, Spur 2: PCR-Produkt von pGOX1, Spur 3: PCR-Produkt von pGOX2, Spur 4: PCR-Produkt von pGOX3, Spur 5: PCR-Produkt von pGOX4.

Die durchgeführten PCR-Reaktionen lieferten deutliche Banden mit den erwarteten Größen: 3,8 kb für pGOX1, 3,6 kb für pGOX2, 2,6 kb für pGOX 3 und 4,0 kb für pGOX4. Auch wenn dieses kein eindeutiger Nachweis der Plasmide pGOX1-4 war, so ergaben die eindeutigen PCR-Produkte einen deutlichen Hinweis auf das Vorhandensein der im Sequenzierungsprojekt identifizierten Plasmide.

3.4.3 Analyse der Plasmide

Sämtliche in *G. oxydans* identifizierte Plasmide waren zirkulär. Die Ermittlung des möglichen Replikationsursprungs auf Plasmiden gestaltete sich im Vergleich zum Chromosom recht schwierig. Ein sogenannter kumulativer GC-Skew, wie er bei der Ermittlung des Replikationsursprungs auf dem Chromosom eingesetzt wurde (s. 3.5.2), ließ sich auf Plasmide nicht ohne weiteres übertragen. Der GC-Skew beruht auf einer unterschiedlichen GC-Verteilung im *leading* und *lagging*-Strang. Plasmide zeigen in der Regel jedoch keinen ausgeprägten Unterschied in der GC-Verteilung. Das traf auch für die Plasmide aus *G. oxydans* zu. Daher konnte zur Ermittlung des potentiellen Replikationsursprungs nur die Annotation herangezogen werden. Das ORF-Finding auf den Plasmiden erfolgte wie unter 2.14.1 beschrieben.

3.4.3.1 Analyse von pGOX1

Bei dem Plasmid pGOX1 handelte es sich um ein 163.187 bp großes, zirkuläres Plasmid, auf dem insgesamt 175 ORFs identifiziert werden konnten. Die Annotation zeigte jedoch, dass es sich bei der Mehrzahl der gefundenen ORFs um hypothetische Gene handelte. Insgesamt konnten nur 55 ORFs eine Funktion zugeordnet werden. Da Plasmide jedoch extrachromosomale Elemente darstellen, die keine essentiellen Informationen tragen, ist dieses Ergebnis nicht weiter erstaunlich. Die Funktionszuordnung der Gene auf dem Plasmid beruhte vor allem auf einem Abgleich mit öffentlich zugänglichen Datenbanken. *G. oxydans* ist jedoch der erste Vertreter der Essigsäurebakterien, dessen Genom vollständig sequenziert vorliegt und auch über Plasmide in *Gluconobacter*-Spezies bzw. deren Funktion ist bislang wenig bekannt. Daher zeigte die Mehrzahl der identifizierten ORFs keine Homologien zu bekannten Proteinen und wurden somit als hypothetisch annotiert. Tab. 4 gibt eine Übersicht über die ORFs auf pGOX1, denen eine Funktion bzw. eine mögliche Funktion zugeordnet werden konnte. Die ORFs 2862, 2863 und 2864 gaben aufgrund der Annotation einen Hinweis auf einen Bereich, der möglicherweise den Replikationsursprung enthält. Diese ORFs zeigten Ähnlichkeiten zu Plasmid-Replikationsproteinen aus *Rhizobium* sp. (ORF 2862), *S. melioli* (ORF 2863) bzw. *A. brasilense* (ORF 2864), die sich in den genannten Organismen im Bereich des Replikationsstartpunkts der jeweiligen Plasmide befinden (RAMIREZ-ROMERO et al., 1997; BARTOSIK et al., 2001).

Unter den ORFs auf pGOX1 mit Funktionszuordnung waren insgesamt 12 Gene, die für Transposasen kodieren. Zudem fanden sich viele ORFs, die für Proteine kodieren, die im konjugativen DNA-Transfer involviert sind. Dazu gehörten ORFs, die für sogenannte Dot-Proteine (dot = defect in organelle trafficking) kodieren. Dot- und auch Tra-Proteine sind Komponenten von Typ IV-Transportern, welche für den konjugativen Transfer von Plasmiden benötigt werden. In *Legionella pneumophila* kodiert eine Gruppe von 22 *dot/icm*-Genen (*icm* = intracellular multiplication) für dieses Typ IV-Sekretionssystem (MOROZOVA et al., 2004). Auf dem Plasmid pGOX1 fand sich ein Cluster, das insgesamt 29 zum Teil hypothetische ORFs umfasste, von denen 12 ORFs Ähnlichkeiten zu Genen zeigten, welche für den konjugativen DNA-Transfer notwendig sind. Darunter waren sechs ORFs, deren Aminosäuresequenz Ähnlichkeiten zu Dot-Proteinen aus *Legionella* zeigten. In direkter Nachbarschaft dieses Clusters, jedoch in entgegengesetzter Leserichtung befand sich ein ORF, dessen Aminosäuresequenz Ähnlichkeiten zu einem konjugativen Transferprotein TraA aus *Rhizobium loti* zeigte. Diese Befunde lieferten deutliche Hinweise, dass sich auf dem Plasmid pGOX1 ein System für den konjugativen DNA-Transfer befindet.

Tab. 4: Übersicht der ORFs auf pGOX1, denen eine Funktion zugeordnet werden konnte

ORF	Start	Stop	Länge in bp	Höchste Homologie zu Proteinen aus	Annotation
2862	43690	44901	1212	<i>Rhizobium</i> sp.	mögliches Replikationsprotein A
2863	45047	46099	1053	<i>Sinorhizobium meliloti</i>	mögliches Replikationsprotein B
2864	46102	47124	1023	<i>Azospirillum brasilense</i>	mögliches RepA Protein ähnlich zu ORF 3607 aus pGOX2
3524	47613	49058	1446	<i>Pseudomonas</i> sp.	DNA-Integrations-, Rekombinations-, Invertionsprotein
2870	51451	51876	426	<i>Acidithiobacillus ferrooxidans</i>	Plasmid-Stabilitäts-Protein StbB
3344	54191	54757	257	<i>G. oxydans</i> <i>Ralstonia solanacearum</i>	Transposase
2875	54760	54984	225	<i>G. oxydans</i> <i>Ralstonia solanacearum</i>	Transposase
2874	54981	57869	2889	<i>Legionella pneumophila</i>	Chemiosmotisches Efflux-System

3548	60338	60984	646	<i>Sphingomonas aromaticivorans</i> F199	Cobalt-Zink-Cadmium Resistenz-Protein CzcD
2879	61649	61975	327	<i>Escherichia coli</i> K12	PemK-ähnliches Protein
2881	62364	63098	735	<i>Erwinia amylovora</i>	Zink Metalloprotease (EC 3.4.24.-)
2882	63104	66397	3294	<i>Legionella pneumophila</i>	Typ I Restriktions-modifikations-System, Restriktions-Untereinheit
2883	66430	67635	1206	<i>Anabaena</i> sp. PCC7120	Typ I Restriktions-modifikations-System, spezifische Untereinheit
2884	67622	69145	1524	<i>Brucella melitensis</i>	Typ I Restriktions-modifikations-System, Methylierungsuntereinheit
3348	71396	72505	1110	<i>Ralstonia metallidurans</i>	konserviertes hypothetisches Protein
2889	73714	74535	822	<i>Gluconobacter oxydans</i> <i>Ralstonia solanacearum</i>	Transposase
2892	76075	77475	1401	<i>Burkholderia fungorum</i>	Membrangebundenes Transportprotein
2893	77527	78090	564	<i>Salmonella typhi</i>	mögliche Isochorismatase (EC 3.3.2.1)
2894	78199	78852	654	<i>Caulobacter crescentus</i>	Transkriptionsregulator, TetR-Familie
2900	83148	83981	834	<i>Xanthomonas axonopodis</i>	Transposase
2902	85161	86164	1003	<i>Xanthomonas axonopodis</i>	Transposase
2904	86648	87373	726	<i>Gluconobacter oxydans</i>	Transposase
2905	87446	88629	1183	<i>Rhizobium meliloti</i>	Transposase
2907	88697	89704	1008	<i>Rhizobium meliloti</i>	mögliche membrangebundene Oxidoreduktase
2910	91346	94048	2703	<i>Agrobacterium tumefaciens</i>	DNA Polymerase III, alpha Untereinheit (EC 2.7.7.7)
2929	111113	112036	924	<i>Legionella longbeachae</i>	DotH
2931	112697	113287	591	<i>Legionella pneumophila</i>	DotG
2935	116111	119266	3156	<i>Legionella pneumophila</i>	DotO
2936	119266	120231	966	<i>Ralstonia solanacearum</i>	Konjugatives Transfer-Protein TrbG
2938	121169	122881	1713	<i>Escherichia coli</i>	Hypothetisches Protein (PilN-ähnlich)
2941	124706	126388	1683	<i>Escherichia coli</i>	Nukleotid-Bindeprotein, ähnlich PilQ
2942	126391	127518	1128	<i>Escherichia coli</i>	PilR-ähnlich
2944	128565	129908	1344	<i>Escherichia coli</i>	PilV-ähnliches Protein
2946	130798	131394	597	<i>Enterobacter aerogenes</i>	Konjugatives Transfer-Protein TrbN

2948	131833	132765	933	<i>Legionella pneumophila</i>	DotC
2950	134257	135369	1113	<i>Legionella pneumophila</i>	DotB
2953	136430	138676	2247	<i>Legionella pneumophila</i>	DotL
2954	138724	141924	3201	<i>Rhizobium loti</i>	Konjugatives Transfer-Protein TraA
2959	142802	145756	2959	<i>Corynebacterium glutamicum</i>	Transposase
2960	145870	146502	633	<i>Corynebacterium glutamicum</i>	DNA-Integrations-, Rekombinations-, Inversionsprotein
2961	146641	147963	1323	<i>Brucella melitensis</i>	Hydroxyacylglutathion-Hydrolase (EC 3.1.2.6)
2963	148334	149227	894	<i>Brucella melitensis</i>	transkriptionelles Regulatorprotein
02964	149227	150345	1119	<i>Pyrococcus abyssi</i>	Transporter
2967	152530	152961	432	<i>Brucella melitensis</i>	Integrales Membran-Protein
2968	152978	153418	441	<i>Pseudomonas denitrificans</i>	Integrales Membran-Protein
2972	156209	157315	1107	<i>Acidianus ambivalens</i>	mögliche terminale Oxidase Untereinheit
2973	158025	158429	405	<i>Rhizobium loti</i>	Transposase
2974	158504	158800	297	<i>Brucella melitensis</i>	Transposase
2830	11229	12056	828	<i>Acetobacter xylinus</i>	Transposase
2832	19665	19970	306	<i>Rhizobium loti</i>	Transposase
2833	20144	22189	2046	<i>Salmonella typhi</i>	DNA-Helicase II (EC 3.6.1.-)
2840	24524	25354	831	<i>Ralstonia solanacearum</i>	Endonuclease (EC 3.1.30.-)
2852	31705	33564	1860	<i>Mesorhizobium loti</i>	Mögliches Plasmid-Stabilisierungsprotein
2857	35640	40778	5139	<i>Rhizobium loti</i>	mögliche DNA-Methylase
2860	41502	41996	495	<i>Caulobacter crescentus</i>	transkriptionelles Regulatorprotein

3.4.3.2 Analyse von pGOX2

Bei pGOX2 handelte es sich um ein 26.568 bp großes Plasmid. Insgesamt wurden 32 ORFs auf dem Plasmid identifiziert, wobei 13 eine Funktion bzw. eine mögliche Funktion zugeordnet werden konnte. Dies entsprach einem prozentualen Anteil von 41 % (Tab. 5). Neben vier Transposasen kodierte das Plasmid unter anderen auch für eine Glutathion-abhängige Formaldehyd-Dehydrogenase (ORF 3589) und eine Oxidoreduktase (ORF 3597), deren Aminosäuresequenzen Ähnlichkeiten zu einem Proteinen aus *Pseudomonas syringae* zeigten. Der Replikationsursprung dieses Plasmids könnte in der Region um ORF 3607 lokalisiert sein. Dieser ORF kodiert für

ein mögliches Replikationsprotein, dessen abgeleitete Aminosäuresequenz Ähnlichkeiten zu einem Replikationsprotein aus *A. brasilense* aufwies.

Tab. 5: Übersicht der ORFs auf pGOX2, denen eine Funktion zugeordnet werden konnte

ORF	Start	Stop	Länge in bp	Höchste Homologie zu Proteinen aus	Annotation
3607	23686	24600	915	<i>Azospirillum brasilense</i>	mögliches RepA Replikase-Protein
3608	24744	25421	678	<i>Zymomonas mobilis</i>	Mögliches Partitionierungsprotein
3581	2185	1691	495	<i>Rhizobium loti</i>	Transkriptionsregulator
3583	3107	4063	957	<i>Brucella melitensis</i>	Transposase
3587	5804	6196	393	<i>Caulobacter crescentus</i>	“Death ON curing”-Protein
3589	8319	7141	1179	<i>Methylobacter marinus</i>	Glutathion-abhängige Formaldehyd-Dehydrogenase (EC 1.2.1.1)
3590	9438	8482	957	<i>Brucella melitensis</i>	Transposase
3593	12352	11762	591	<i>Yersinia enterocolitica</i>	mögliche Resolvase
3597	14407	15477	1071	<i>Pseudomonas syringae</i>	Oxidoreduktase
3598	16378	15818	561	<i>Brucella melitensis</i>	Transposase
3601	18883	17807	1077	<i>Azotobacter vinelandii</i>	Transposase
3605	20913	22358	1446	<i>Pseudomonas</i> sp.	Integrase ähnliches Protein
3606	22549	23370	822	<i>Gluconobacter oxydans</i>	Transposase

3.4.3.3 Analyse von pGOX3

Das zirkuläre Plasmid pGOX3 wies eine Größe von 14.547 bp auf. Es konnten insgesamt 18 ORFs identifiziert werden, von denen sieben eine Funktion zugewiesen wurde (Tab. 6). Ein wenig überraschend war der Befund, dass auf pGOX3 offenbar keine Transposase lag. Diese mobilen Elemente traten im gesamten Genom gehäuft auf; pGOX3 war neben dem nur 2,7 kb großen Plasmid pGOX5 das einzige der fünf Plasmide in *G. oxydans*, das keine Transposase enthielt. Ein möglicher Replikationsursprung könnte die Region um ORF 3383 sein. Die von diesem ORF abgeleitete Aminosäuresequenz zeigte Ähnlichkeiten zu einem Replikations-Initiationsprotein aus *E. coli*.

Tab. 6: Übersicht der ORFs auf pGOX3, denen eine Funktion zugeordnet werden konnte

ORF	Start	Stop	Länge in bp	Höchste Homologie zu Proteinen aus	Annotation
2980	394	813	420	<i>Rhizobium</i> sp.	mögliches Plasmid-Stabilitätsprotein Y4JK
2985	4740	5411	672	<i>Klebsiella pneumoniae</i>	Resolvase
2995	10803	13886	3084	<i>Rhizobium loti</i>	Konjugatives Transfer-Protein TraA
3380	4041	4694	654	<i>Rhizobium</i> sp	DNA-Integrations-/Rekombinations- /Invertionsprotein
3381	5499	6173	675	<i>Pseudomonas alcaligenes</i>	mögliches Protein involviert in Partitionierung
3383	7610	6513	1098	<i>Escherichia coli</i>	Replikations-Initiationsprotein

3.4.3.4 Analyse von pGOX4

Das Plasmid pGOX4 wies eine Größe von 13.223 bp auf. Es konnten 17 ORFs identifiziert werden, von denen fünf eine mögliche Funktion zugeordnet wurde (Tab. 7). Auffällig war, dass es sich bei vier ORFs um Transposasen handelte. Die Annotation lieferte keinerlei Hinweise auf den möglichen Ursprung der Replikation.

Tab. 7: Übersicht der ORFs auf pGOX4, denen eine Funktion zugeordnet werden konnte

ORF	Start	Stop	Länge in bp	Höchste Homologie zu Proteinen aus	Annotation
3565	128	1099	972	<i>Brucella melitensis</i>	Transposase
3370	2694	3392	699	<i>Thiobacillus ferrooxidans</i>	Protein involviert in Partitionierung
3555	5174	4098	1077	<i>Azotobacter vinelandii</i>	Transposase
3566	6241	7101	861	<i>Pseudomonas putida</i>	Transposase
3017	13054	12227	828	<i>Acetobacter xylinus</i>	Transposase

3.4.3.5 Analyse von pGOX5

Das Plasmid pGOX5 war mit einer Größe von 2.687 bp das kleinste der in *G. oxydans* gefundenen Plasmide. Es kodiert für fünf ORFs, von denen nur einem eine Funktion zugeordnet werden konnte (Tab. 8). Dieses Gen (ORF 3415) kodiert vermutlich für ein

Replikationsprotein, basierend auf Ähnlichkeiten zu einem Replikationsprotein aus *A. europaeus*. Da dieses Plasmid relativ klein ist, könnte es als Grundlage für einen *Gluconobacter*-spezifischen Klonierungsvektor dienen.

Tab. 8: Übersicht der ORFs auf pGOX5, denen eine Funktion zugeordnet werden konnte

ORF	Start	Stop	Länge in bp	Höchste Homologie zu Proteinen aus	Annotation
3415	320	862	543	<i>Acetobacter europaeus</i>	Replikationsprotein

3.5. Analyse des Chromosoms von *G. oxydans*

Im Folgenden soll das zirkuläre Chromosom von *G. oxydans* 621H vorgestellt und analysiert werden. Aufgrund der Fülle von Daten, die ein Sequenzierungsprojekt liefert, können jedoch nur ausgewählte Eigenschaften betrachtet werden.

3.5.1 Allgemeine Merkmale des Chromosoms

G. oxydans ist das erste Bakterium aus der Familie der Essigsäurebakterien, dessen Sequenz vollständig vorliegt. Charakteristisch für Essigsäurebakterien ist die unvollständige Oxidation von vielen Zuckern, Alkoholen und Polyolen. Viele dieser Dehydrogenierungsreaktionen werden von membrangebundenen Enzymen katalysiert, deren aktives Zentrum in den periplasmatischen Raum ragt. Dadurch ist ein Transport der entsprechenden Substrate über die Cytoplasmamembran nicht notwendig. Außerdem werden die unvollständig oxidierten Produkte direkt in das Medium ausgeschieden (DEPPENMEIER et al., 2002). Diese Merkmale machen *G. oxydans* für die Biotechnologie sehr interessant. Von der Analyse der Sequenz erhofft man sich, neue biotechnologische Prozesse entwickeln zu können und bestehende zu verbessern.

In Tabelle 9 sind einige Merkmale des Genoms von *Gluconobacter* zusammengefasst. Das Gesamtgenom hatte eine Größe von 2,9 Mb, einschließlich der identifizierten Plasmide. Der G+C-Gehalt dieses Organismus war mit 60,8 % recht hoch. Auffallend waren in der Sequenzierungsphase vor allem die vielen repetitiven Sequenzen innerhalb

des Genoms, die bei der Editierung und Assemblierung erhebliche Schwierigkeiten bereiteten. Das Vorhandensein mobiler Elemente (Transposasen) lässt vermuten, dass viele Gene von anderen Organismen akquiriert wurden.

Tab. 9: Allgemeine Merkmale des Genoms von *G. oxydans* 621H

Größe des Chromosoms	2,7 Mbp
Größe der Plasmide	pGOX1: 163 kb pGOX2: 26,6 kb pGOX3: 14,5 kb pGOX4: 13,2 kb pGOX5: 2,7 kb
Kodierende Sequenz	90,1 %
G+C-Gehalt	60,8 %
Anzahl der offenen Leserahmen (Chromosom, pGOX1, pGOX2, pGOX3, pGOX4, pGOX5)	2492, 175, 32, 18, 21, 5
ORFs mit zugeordneter Funktion	70 %
Nicht charakterisierte Dehydrogenasen/Reduktasen	> 70
Nicht charakterisierte Alkohol-Dehydrogenasen	> 30
Nicht charakterisierte Zucker/Polyol-Dehydrogenasen	> 20
RNA-Gene	67
rRNA-Operons	4
tRNAs	55
Transposasen	98

Ein erster Hinweis auf lateralen Gentransfer in dem Organismus lieferte die G+C-Variation. Diese wurde im Chromosom von *G. oxydans* in Schritten von 1000 bp untersucht (Abb. 15). Betrachtet man Abb. 15 fällt auf, dass es auf dem Chromosom durchaus Bereiche gab, in denen der lokale G+C-Gehalt vom mittleren G+C-Gehalt mehr oder weniger stark abwich (Abb. 15, innere weiße Linie). Insgesamt fielen sieben Bereiche auf. Diese befanden sich, ausgehend von der in Abb. 15 dargestellten Skala (äußerer Rand), bei ca. 303 kb, 446 kb, 1.037 kb, 1.465 kb, 1.660 kb, 2.050 kb und 2.416 kb. Mit wenigen Ausnahmen kodierten die ORFs, die starke Abweichungen vom mittleren G+C-Gehalt aufwiesen, für hypothetische Proteine, also für Proteine, die keinerlei Ähnlichkeiten zu bekannten Proteinen zeigten. Eine Ausnahme bildete die

Region des Chromosoms bei ca. 446 kb. Hier kodierten die entsprechenden ORFs für Proteine, welche in die Biosynthese von Flagellen involviert sind. Darüber hinaus sind diese ORFs von einer Transposase flankiert. Dieser Befund könnte ein Hinweis darauf sein, dass die Gene, die für die Biosynthese von Flagellen kodieren, durch lateralen Gentransfer akquiriert wurden. Die Ausreißer bezüglich des mittleren G+C-Gehalts an den Positionen des Chromosoms 1.037 kb und 1.465 kb sind ebenfalls von Transposons flankiert, so dass auch hier die Vermutung nahe liegt, dass die betroffenen ORFs durch lateralen Gentransfer aufgenommen wurden. Wie schon erwähnt handelt es sich dabei jedoch um ORFs, denen keine Funktion zugeordnet werden konnte. Anzumerken bleibt, dass in dem Genom von *G. oxydans* 98 Transposasen identifiziert werden konnten. Viele dieser Gene zeigten jedoch keine signifikante Abweichungen vom mittleren G+C-Gehalt. Die zirkuläre Karte des Chromosoms von *G. oxydans* 621H gibt außerdem einen Überblick über sämtliche ORFs in Abhängigkeit von der Lokalisierung auf dem jeweiligen DNA-Strang, (dargestellt in hell- und dunkelblau). In gelb und rot sind sämtliche Dehydrogenasen, Reduktasen und Oxidoreduktasen auf dem Chromosom gezeigt. Hierbei sind auch solche potentiellen Gene aufgeführt, die anhand von Homologien zu Proteindomänen als mutmaßliche Oxidoreduktasen annotiert wurden. Es ist zu erkennen, dass diese relativ gleichmäßig über das gesamte Chromosom verteilt sind.

3.5.1.1 ORF-Vorhersage und Übersicht über die Annotation

Die ORF-Vorhersage wurde wie unter 2.14.1 beschrieben mit Hilfe des Computerprogramms YACOP (TECH & MERKL, 2003) durchgeführt. Auch wenn die ORF-Vorhersage mit diesem Programm exakter ist als mit Glimmer (SALZBERG et al., 1998), Critica (BADGER & OLSEN, 1999) oder ZCURVE (GUO et al., 2003) gab es dennoch überlappende ORFs. Diese mussten manuell auf ihre Richtigkeit geprüft und gegebenenfalls gelöscht werden.

Die 2492 ORFs des Chromosoms repräsentierten 90,1 % der gesamten Sequenz. Diese sind in Abb. 15 dargestellt. 62 % der gefundenen ORFs konnte eine Funktion zugeordnet werden. 24 % aller ORFs zeigen Homologien zu Proteinen aus anderen Bakterien mit unbekannter Funktion. 38 % aller ORFs konnte keine Funktion zugeordnet werden, wobei diese Zahl auch die hypothetischen ORFs auf den Plasmiden

beinhaltet. Neben sämtlichen ORFs des Chromosoms sind in Abb. 15 auch diejenigen ORFs dargestellt, die für Dehydrogenasen, Reduktasen oder Oxidoreduktasen kodieren. Hierbei sind auch solche potentiellen Gene aufgeführt, die anhand von Homologien zu Proteindomänen als mutmaßliche Dehydrogenasen, Reduktasen oder Oxidoreduktasen annotiert wurden.

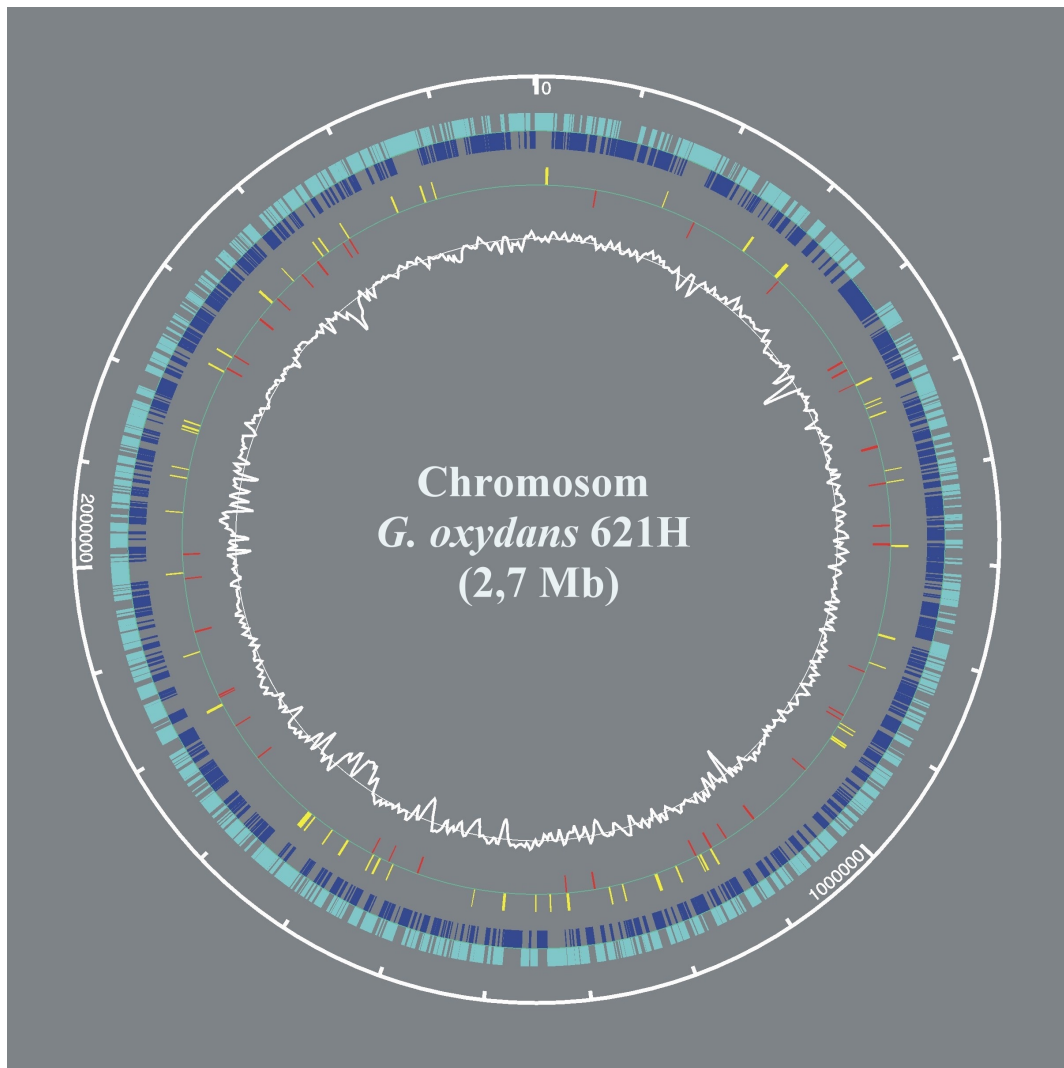


Abb. 15: Zirkuläre Karte des Chromosoms von *G. oxydans* 621H

Der äußere Rand stellt die Basen-Skala dar, der Nullpunkt entspricht der Position des ermittelten Replikationsursprungs. In hell- und dunkelblau sind sämtliche ORFs in Abhängigkeit der Lokalisierung auf dem jeweiligen DNA-Strang gezeigt. In gelb und rot sind sämtliche, auch potentielle, Dehydrogenasen, Reduktasen und Oxidoreduktasen des Chromosoms abgebildet. Die innere weiße Linie zeigt die Abweichung vom mittleren G+C-Gehalt.

3.5.2 Identifizierung des Replikationsursprungs

Erste Hinweise für die Bestimmung des Replikationsursprungs können schon mit bioinformatischen Methoden gewonnen werden. Eine Möglichkeit liefert der sogenannte GC-Skew (LOBRY, 1996). Dabei wird der prozentuale Gehalt von $(C-G)/(C+G)$ in einer definierten Fenstergröße berechnet und gegen das Genom (Anzahl der bp) aufgetragen. Eine Weiterentwicklung dieser Methode ist der kumulative GC-Skew (GRIGORIEV, 1998). Hierbei wird für den GC-Gehalt nicht mehr eine Fenstergröße definiert, sondern es wird jede einzelne Base einbezogen. Nach jeder Base, die das Programm gelesen hat, wird auf der X-Achse ein Schritt nach rechts gemacht. Wird ein G oder C gelesen, wird ein Schritt nach oben bzw. unten vorgenommen. Ein Vorzeichenwechsel in der daraus resultierenden Auftragung ist ein Hinweis auf den potentiellen Replikationsursprung. Hintergrund für diese Methode ist der Befund, dass die GC-Verteilung im *leading* und im *lagging* Strang der DNA unterschiedlich ist. Untersuchungen haben gezeigt, dass der *leading* Strang im Vergleich zum *lagging* Strang in der Regel mit den Basen G und T angereichert ist (TILLER & COLLINS, 2000). Erklärt wird dieses Phänomen damit, dass *leading* und *lagging* Strang der DNA während der Replikation unterschiedlichen Mutationsdrücken, bedingt durch die Dauer der Einzelsträngigkeit, ausgesetzt sind (LOBRY & SUEOKA, 2002).

Abb. 16 zeigt den kumulativen GC-Skew für das Chromosom von *G. oxydans*. Als Eingabe diente die Sequenz des linearisierten Chromosoms. In Abb. 16 befindet sich bei 1,3 Mb ein deutliches Maximum. Wie schon erwähnt deutet ein Vorzeichenwechsel bei dem GC-Skew darauf hin, dass sich hier der Replikationsursprung befindet. Zusätzlich wurde die identifizierte Position mit den sich dort befindenden ORFs verglichen. Die Region auf dem Chromosom, bei der sich in Abb. 16 das Maximum befindet, ist in Abb. 17 dargestellt.

In Übereinstimmung mit dem GC-Skew findet man in der gezeigten Region deutliche Hinweise darauf, dass es sich hier um den Replikationsursprung handelt. Neben dem Gen, das für das Initiatorprotein DnaA kodiert, findet man stromaufwärts davon noch weitere Gene, die charakteristischerweise in der Nähe des Replikationsstartpunkts

liegen. Dazu gehören die Gene, die für RecF, die DNA-Polymerase III und die DNA Gyrase kodieren.

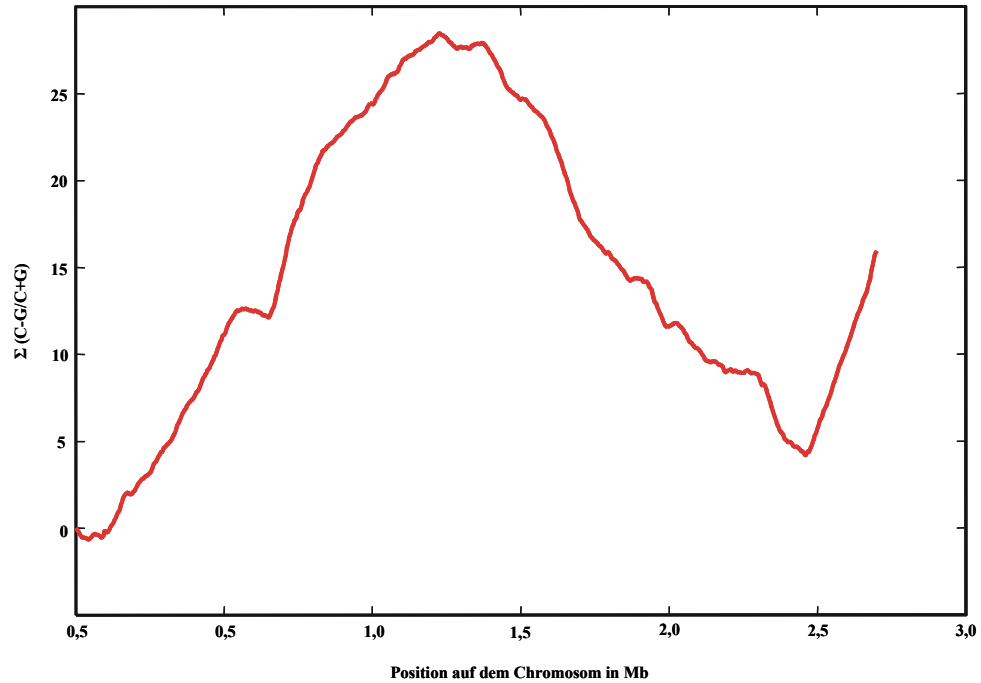


Abb. 16: Kumulativer GC-Skew des Chromosoms von *G. oxydans*

Auf der Y-Achse ist die lokale GC-Verteilung, berechnet nach der Formel $\Sigma(C-G/C+G)$ (GRIGORIEV, 1998), aufgetragen. Die X-Achse zeigt die Position auf dem linearisiert dargestellten Chromosom in Mb.

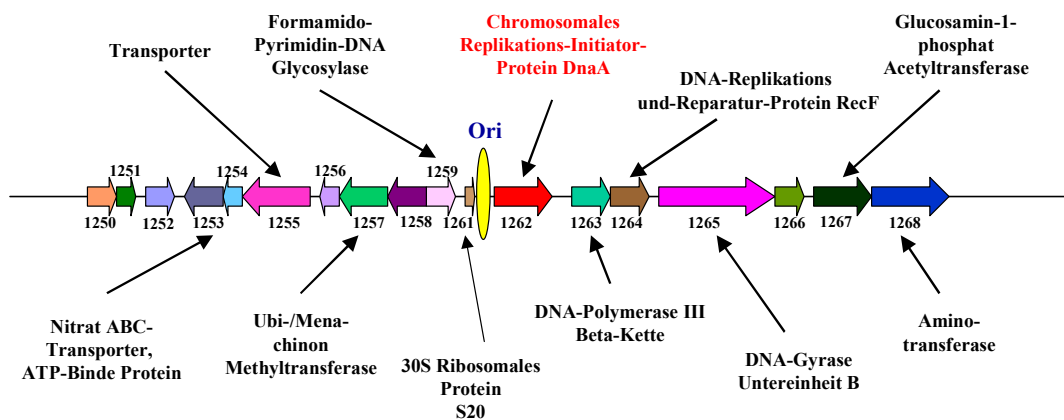


Abb. 17: DNA-Region um den potentiellen Replikationsursprung

Gezeigt sind jeweils 10 kb stromaufwärts und stromabwärts von *dnaA*, dem Gen, das für das Initiatorprotein der Replikation kodiert.

Zur weiteren Analyse wurde diese Region mit denen anderer Organismen verglichen, welche die höchsten Homologien zu DnaA aus *G. oxydans* aufweisen (Abb. 18).

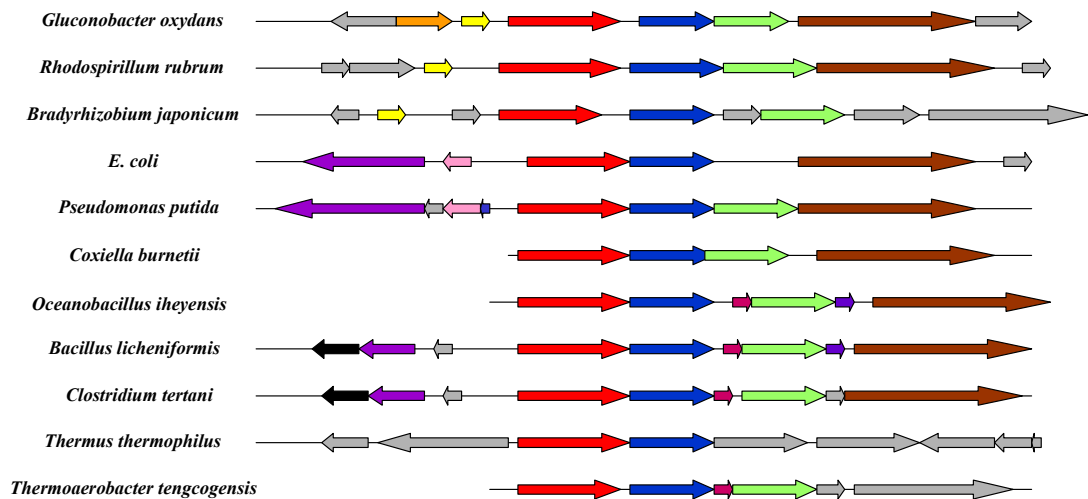


Abb. 18: Vergleich der Region des Replikationsursprungs von *G. oxydans* mit homologen Bereichen anderer Organismen

Dargestellt sind die in der Nachbarschaft von *dnaA* liegenden Gene. Farben der offenen Leserahmen: Blau: DNA-Polymerase III, Beta-Kette; rot: DnaA; pink: Ribonuklease P; grün: RecF; braun: DNA-Gyrase, Untereinheit B; gelb: ribosomales Protein; schwarz: Jag Protein; orange: Glycosylase; lila: Membranprotein. Den offenen Leserahmen, deren Farben hier nicht aufgeführt sind, konnte keine Funktion zugeordnet werden.

Da *Gluconobacter oxydans* der einzige vollständig sequenzierte Vertreter der Essigsäurebakterien ist, war es nicht möglich, diese Region mit den nächsten Verwandten bzw. mit der entsprechenden Region anderer Essigsäurebakterien zu vergleichen. Aber auch der Abgleich mit weniger nah verwandten Organismen macht die Konservierung der Gene, die sich in der Nähe des Replikationsursprungs befinden (DnaA, RecF, DNA-Polymerase III, DNA-Gyrase), deutlich.

3.6 Rekonstruktion des Stoffwechsels

Zur weiteren Analyse des Genoms von *G. oxydans* werden im folgenden auf der Basis der vorliegenden Sequenz die zentralen Stoffwechselwege vorgestellt. Darüber hinaus wird ein Einblick in den weiteren Metabolismus dieses Organismus gegeben. Aufgrund

der Fülle der erhaltenen Daten können im Rahmen dieser Arbeit jedoch nur ausgewählte Funktionen behandelt werden.

3.6.1 Transportsysteme in *G. oxydans*

Für die Aufnahme von verschiedensten Substraten findet man im Genom von *G. oxydans* unterschiedliche Aufnahmesysteme. Es wurden insgesamt etwa 200 Gene identifiziert, die für Transportproteine kodieren. Substrate, die in die Zelle gelangen, dienen jedoch überwiegend zum Aufbau von Zellsubstanz. Die Energiegewinnung findet durch membrangebundene Enzyme statt (MATSUSHITA et al., 1994). Abb. 19 gibt einen Überblick über Beispiele für unterschiedliche Aufnahmesysteme.

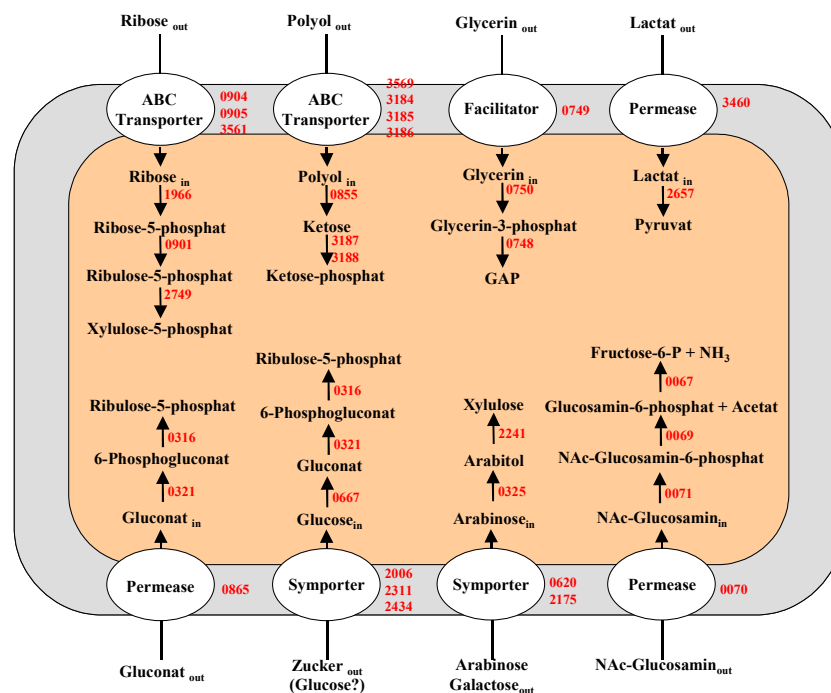


Abb. 19: Beispiele für die Aufnahme und Verwertung verschiedener Substrate von *G. oxydans*

Gezeigt sind Beispiele für Transportsysteme, die in *G. oxydans* gefunden wurden, sowie die Weiterverwertung der entsprechenden Substrate. Die Zahlen geben jeweils die entsprechenden ORF-Nr. an.

Auffallend war, dass in dem Genom keine spezifischen Aufnahmesysteme für Glucose gefunden wurden. ORFs, welche für ein funktionsfähiges Phosphotransferasesystem

kodieren, das den Transport durch Gruppentranslokation bewirken würde, konnten nicht identifiziert werden. Die „universellen“ Komponenten des PTS-Systems liegen zwar vor (HPr, EI, EII^A), jedoch fehlt die Komponente EII^B und das membranständige Zucker-spezifische EnzymII^C.

3.6.1.1 Erleichterte Diffusion

Bei der Betrachtung der Mechanismen für die Aufnahme verschiedener Substrate konnte nur ein Gen identifiziert werden, das für ein Transportprotein zur erleichterten Diffusion (Facilitator-Protein) kodiert. Es liegt in einem Operon, das für die Glycerinverwertung verantwortlich ist (Abb. 20). Neben dem Facilitator-Protein kodieren die Gene dieses Operons noch für ein Regulatorprotein, sowie für eine Kinase, die für die Phosphorylierung des Substrats sorgt und eine Glycerin-3-phosphat-Dehydrogenase.

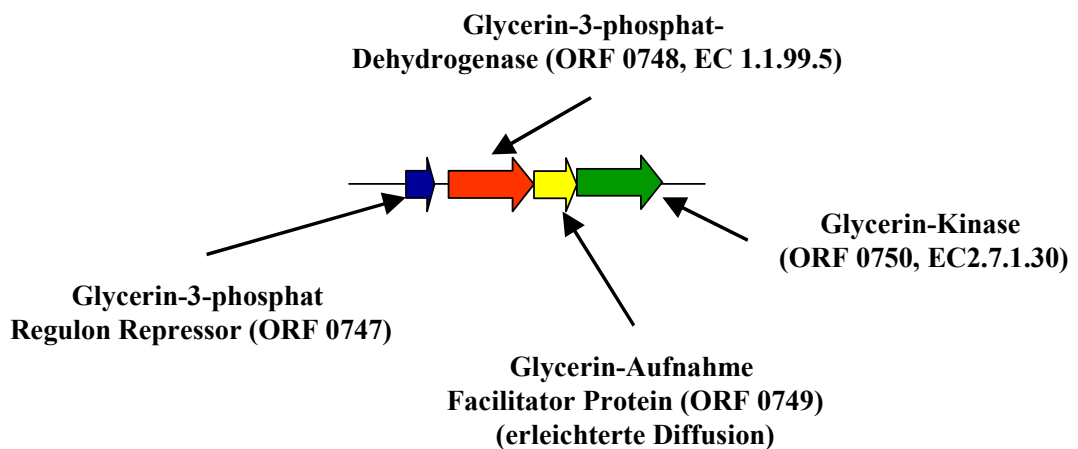


Abb. 20: Das Glycerin-Operon in *G. oxydans*

3.6.1.2 ABC-Transporter

ABC-Transporter sind sowohl in Eukaryoten als auch in Prokaryoten weit verbreitet. Grundsätzlich bestehen sie aus je zwei hydrophoben Domänen oder Untereinheiten, die in der Membran verankert sind und zwei hydrophilen Domänen oder Untereinheiten, welche die ATP-Bindekassette beinhalten und die notwendige Energie für den

Transport durch ATP Hydrolyse bereitstellen. Im Genom von *G. oxydans* wurden 33 mögliche ABC-Transporter identifiziert. Dabei konnte in einigen Fällen jedoch nur die ATP-Bindekassette manifestiert werden. Hierbei handelt es sich um eine sehr konservierte Domäne. Die hydrophobe Domäne oder Untereinheit hingegen, welche die Substratspezifität determiniert, muss nicht zwingend Sequenzähnlichkeiten zu bekannten Proteinen zeigen. Als Beispiel für einen ABC-Transporter, wie man ihn in *G. oxydans* findet, sind in Abb. 21 die Gene für die Aufnahme von Ribose dargestellt. Daneben wurden noch ABC Transporter für die Aufnahme von Metallionen, Aminosäuren, Phosphat, Di- und Oligopeptide, Zuckern und Polyole (z. B. Mannitol und Sorbitol) identifiziert (Abb. 22).

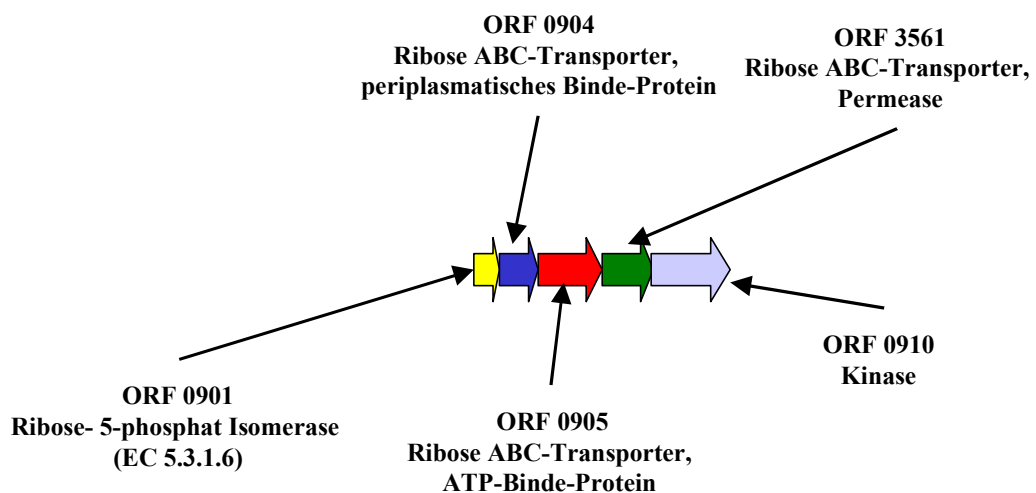


Abb. 21: Cluster des ABC-Transporters zur Aufnahme von Ribose

Neben den Genen des ABC-Transporters befinden sich in diesem Cluster auch die Gene für die ersten Schritte der Riboseverwertung (Kinase und eine Isomerase).

3.6.1.3 Symporter und Antiporter

Die Annotation des Genoms von *G. oxydans* zeigte das Vorhandensein von sieben Antiportern und sechs Symportern. Unter anderem erfolgt die Aufnahme von Galactose mit dem gleichzeitigen Einstrom von Protonen durch einen Symporter. Auch für die Aufnahme von Glucose liegt die Vermutung nahe, dass der Transport in die Zelle über einen Symporter erfolgt. Eindeutig konnte dieses anhand der Daten jedoch nicht geklärt

werden. Die Annotation zeigte jedoch, dass drei Gene vorliegen, die für einen Glucose/Protonen Symporter kodieren könnten (2006, 2311, 2434). Auch in der Literatur ist nichts über die Aufnahme von Glucose bei *Gluconobacter* beschrieben. Allerdings erfolgt die Glucose-Aufnahme nicht wie in *E. coli* durch das Phosphotransferase System. Bei der Annotation konnten zwar Gene für ein Phosphotransferase System identifiziert werden, jedoch fehlen ORFs, die für die zuckerspezifischen Enzyme (EII) kodieren. Dieser Befund war zunächst überraschend, doch ist es *de facto* gar nicht so ungewöhnlich, dass Glucose nicht mit Hilfe des Phosphotransferase Systems aufgenommen wird. Der Transport von Glucose durch Symporter findet man z. B. auch in *Pseudomonas aeruginosa* oder *Azotobacter vinelandii* (ROMANO et al., 1970). Da *G. oxydans* weder eine vollständige Glykolyse noch einen vollständigen Zitratzyklus (s. 3.6.2.1, 3.6.2.4) hat und die aufgenommenen Substrate nicht der Energiegewinnung, sondern dem Intermediärstoffwechsel dienen, ist die chemische Modifikation eines Zuckers bei dem Transport in die Zelle auch weniger entscheidend. Maßgebend sind vor allem die in der Membran stattfindenden energieliefernden Prozesse.

3.6.1.4 Permeasen

Unter dem Begriff Permeasen werden in dieser Arbeit solche Transportsysteme zusammengefasst, deren Mechanismus zur Bereitstellung von Energie für den Transportvorgang nicht bekannt ist. Mögliche energieliefernde Reaktionen sind die Hydrolyse von ATP (ATPasen) und der mit der Substrataufnahme einhergehende Einstrom (Symporter) oder Ausstrom (Antiporter) von Kationen unter Nutzung des elektrochemischen Ionengradienten über der Cytoplasmamembran. In dieser Gruppe von Transportsystemen finden sich solche, die Nukleotid-Derivate (Nucleoside, Purine, Xanthin/Uracil, Purin/Cytosin) und organische Verbindungen (Lactat, Glucarate, Gluconat, C4-Carboxylate und Aminosäuren) in die Zelle einschleusen. Auch die Aufnahme von N-Acetylglucosamin erfolgt über eine spezifische Permease (ORF 0070). Das aufgenommene Substrat kann direkt für den Aufbau der Mureinschicht genutzt werden. Alternativ kann dieser Organismus N-Acetylglucosamin auch in mehreren Reaktionsschritten zu Fructose-6-phosphat umwandeln, welches dann in den Pentosephosphat-Weg eingebracht werden kann (Abb. 19). Darüber hinaus wurden einige Ionen-translozierende Permeasen für die Aufnahme von Mg^{2+}/Co^{2+} , Mangan und

divalente Kationen identifiziert. Insgesamt besitzt *G. oxydans* mehr als 25 Permeasen. Diese Zahl beinhaltet jedoch auch mutmaßliche Permeasen, denen keine Substratspezifität zugeordnet werden konnte.

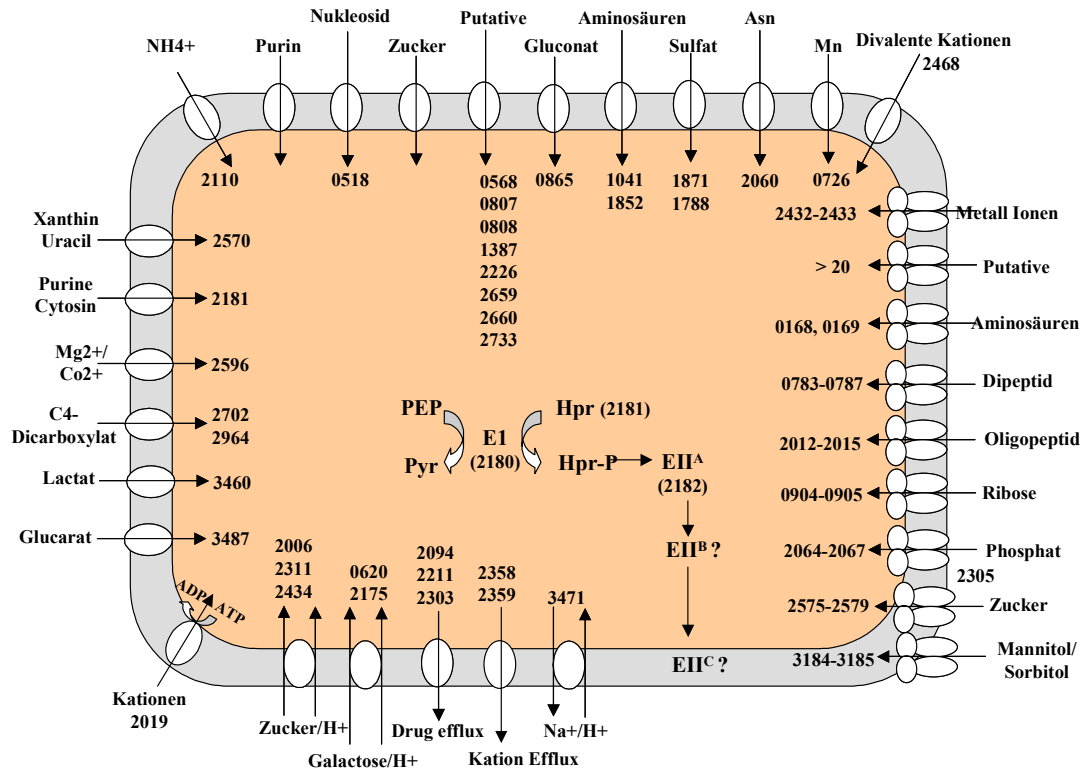


Abb. 22: Zusammenfassende Übersicht über die Transportsysteme in *G. oxydans*

3.6.2 Substratverwertung

Die Verwertung der unterschiedlichen Substrate (Polyole, Alkohole, Zucker) durch *G. oxydans* erfolgt vornehmlich durch membrangebundene Dehydrogenasen (MATSUSHITA et al., 1994). Dabei ragt das aktive Zentrum des jeweiligen Enzyms in den periplasmatischen Raum (DEPPENMEIER et al., 2002). Somit ist ein Transport des entsprechenden Substrats nicht notwendig. Die Substrate werden dabei jedoch nur unvollständig oxidiert. Die korrespondierenden Produkte, organische Säuren, Aldehyde und Ketone, werden fast vollständig ins Medium ausgeschieden. Gerade diese Merkmale sind für *Gluconobacter* charakteristisch und machen die oxidative

Fermentation aus, die vor allem für biotechnologische Prozesse interessant ist. Ein Transport von Substraten in die Zelle hinein und die anschließende Verstoffwechslung erfolgt lediglich für die Synthese von Intermediaten und den Aufbau von Zellsubstanz. Im Folgenden werden die Stoffwechselwege, die *G. oxydans* beschreiten kann, an ausgewählten Beispielen beschrieben.

3.6.2.1 Glykolyse und Gluconeogenese

Schon seit Jahrzehnten (STOUTHAMER, 1959) ist bekannt, dass *Gluconobacter*-Stämme keine Glykolyse durchführen. Die Gründe hierfür waren lange unbekannt, da es in früherer Zeit nicht möglich war die Aktivität der Phosphofruktokinase, dem Schlüsselenzym der Glykolyse, nachzuweisen. Auf der Basis der genomischen Sequenz zeigte sich jedoch, dass *G. oxydans* kein Gen aufweist, das für dieses Enzym kodiert (Abb. 23). Das Gen, welches für die Glucose-6-phosphat-Isomerase (ORF 0315) kodiert, weist eine bifunktionale Funktion auf. Zum einen katalysiert das Enzym die Isomerisierung von Glucose-6-phosphat zu Fructose-6-phosphat und umgekehrt, zum anderen weist es auch eine Transaldolase-Aktivität auf (SUGIYAMA et al., 2003b). Lokalisiert ist dieser ORF innerhalb eines Clusters, das für Enzyme des Pentosephosphat-Weges kodiert.

Ausgehend von Glucose-6-phosphat können Stoffwechselwege, wie der Pentosephosphat-Weg und der Entner–Doudoroff-Weg besprochen werden. In Abb. 23 sind außerdem diejenigen Reaktionen gekennzeichnet, bei denen ATP hydrolysiert oder gebildet wird. Die Reaktionen der Phosphoglycerat- und der Pyruvat-Kinase liefern ATP und stellen Reaktionen der Substratkettenphosphorylierung dar. Diese sind auf der Basis der genomischen Sequenz die einzigen Reaktionen der Substratkettenphosphorylierung, zu denen *G. oxydans* in der Lage ist. Allerdings kostet die Phosphorylierung der Glucose ATP, netto wird damit ein ATP gewonnen, wenn man berücksichtigt, dass *G. oxydans* die Reaktion der Phosphofruktokinase nicht katalysieren kann.

In Abb. 23 sind nur die Reaktionen der Glykolyse dargestellt, nicht aber die der Gluconeogenese. Ein Blick in die Informationen, welche die Genomdaten liefern, zeigte, dass auch hier das Schlüsselenzym, die PEP-Synthetase fehlt. Darüber hinaus

kodiert auch kein ORF dieses Organismus für die Pyruvat-Dikinase, oder andere Enzyme, die Phosphoenolpyruvat synthetisieren können. *G. oxydans* ist somit nicht in der Lage, aus Pyruvat oder Oxalacetat Phosphoenolpyruvat zu synthetisieren. Mit der PEP-Synthetase und der Phosphofruktokinase fehlen diesem Organismus entscheidende Enzyme des zentralen Stoffwechsels.

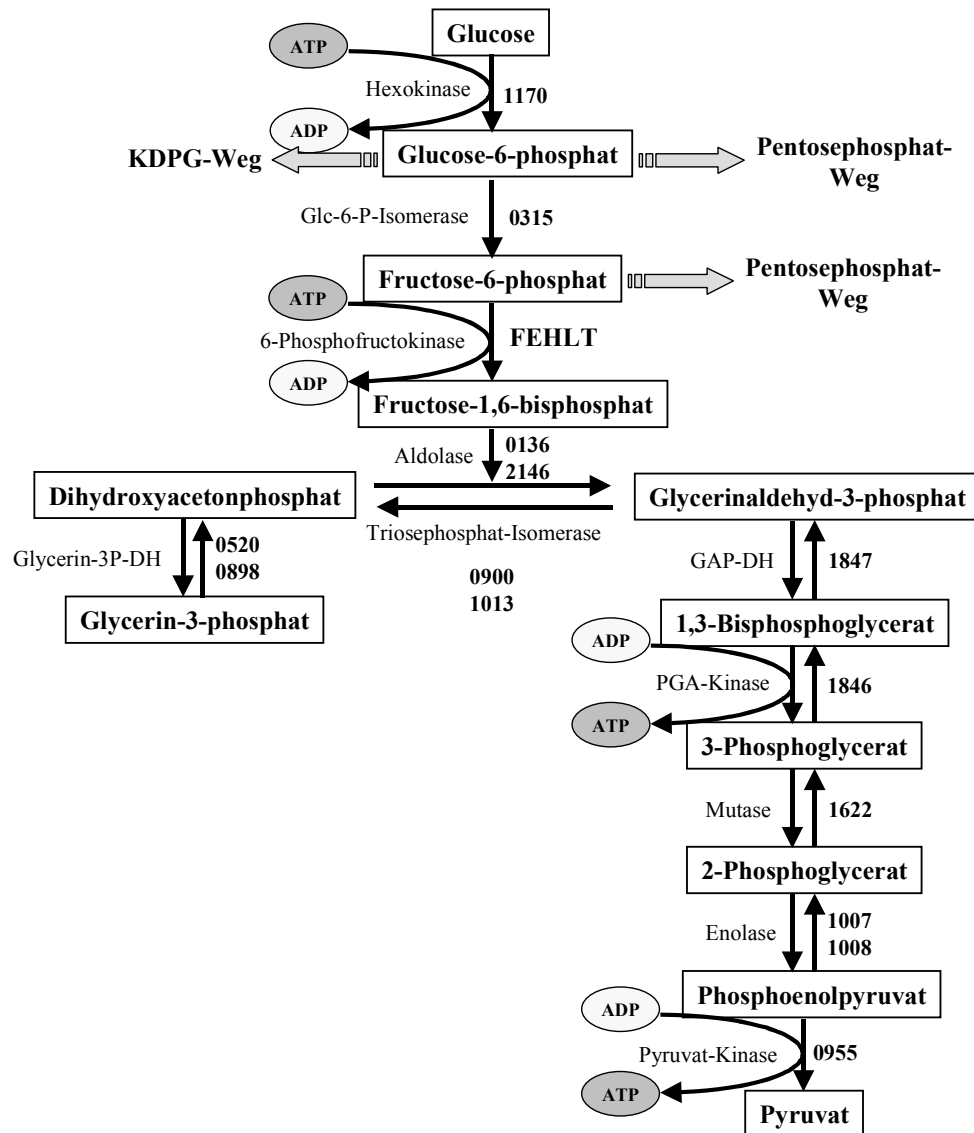


Abb. 23: Darstellung der Glykolyse

Die Nummern geben den entsprechenden ORF an. Das Schlüsselenzym der Glykolyse, die Phosphofruktokinase fehlt. Der ORF für die Glucose-6-phosphat-Isomerase kodiert gleichzeitig für die Transaldolase-Reaktion des Pentosephosphat-Weges. Abkürzungen: Glc-6-P: Glucose-6-phosphat, Glycerin-3P-DH: Glycerin-3-phosphat-Dehydrogenase, GAP-DH: Glycerinaldehydphosphat-Dehydrogenase, PGA: Phosphoglycerat.

3.6.2.2 Pentosephosphat-Weg

Der Pentosephosphat-Weg ist die wichtigste Quelle zur Bereitstellung von NADPH und zentralen Metaboliten des Intermediärstoffwechsels. NADPH wird vor allem für die Synthese von Proteinen und Lipiden benötigt. Die Reaktionsschritte des Pentosephosphat-Weges sind in Abb. 24 dargestellt. Der obere Teil der Abbildung zeigt die oxidativen Schritte, im unteren Teil sind die reversiblen Reaktionen der Transaldolase und Transketolase gezeigt. Durch die Transaldolase- und Transketolase-Reaktionen kann eine Isomerisierung von Aldosen in Ketosen und umgekehrt erfolgen, wodurch sich der Pentosephosphat-Weg zu einem Zyklus schließen lässt. In Abb. 25 sind die Gene, die für die entsprechenden Enzyme des Pentosephosphat-Weges kodieren dargestellt. Dabei fällt auf, dass ein Teil der Gene in einem Cluster vorliegen, andere hingegen sind einzeln in dem Genom lokalisiert.

In einem Cluster findet man die ORFs, die für die Transketolase (ORF 0314), Transaldolase (ORF 0315), 6-Phosphogluconat-Dehydrogenase (ORF 0316), Lactonase und Ribosephosphat-Isomerase (ORF 0320) kodieren. Ebenfalls in diesem Cluster befindet sich ein Gen, welches für eine Glucokinase (ORF 0321) kodiert. Darüber hinaus ist an anderer Position in dem Genom aber auch eine einzeln liegende Glucokinase (ORF 1170) kodiert. Außerdem sind außerhalb dieses Clusters die ORFs, die für die Glucose-6-phosphat-Dehydrogenase und die Ribulosephosphat-Epimerase kodieren, positioniert. Betrachtet man die betroffenen ORFs genauer (Abb. 25), fällt auf, dass einige für den Pentosephosphat-Weg spezifisch sind, andere kodieren für Proteine, die auch für andere Stoffwechselreaktionen relevant sind. Mit Ausnahme der Glucokinase (ORF 0321) sind genau diejenigen Gene in einem Cluster organisiert, die für den Pentosephosphat-Weg spezifisch sind.

Auffallend war, dass der ORF 0315, welcher für die Transaldolase-Reaktion im unteren Teil des Pentosephosphat-Weges kodiert, offensichtlich ein bifunktionales Enzym determiniert. Das entsprechende Protein ist nicht nur in den Pentosephosphat-Weg involviert, sondern weist zusätzlich eine Glucose-6-phosphat-Isomerase Funktion auf, wie sie im oberen Teil der Glykolyse benötigt wird. Da in der Glykolyse ein ORF, welcher für die Phosphofruktokinase kodiert fehlt, kann Fructose-6-phosphat durch diesen Weg nicht weiter verstoffwechselt werden. Das Vorhandensein einer

bifunktionalen Glucose-6-phosphat-Isomerase/Transaldolase impliziert, dass Fructose-6-phosphat direkt in den Pentosephosphat-Weg eingebracht werden kann. Das Vorhandensein eines Glucokinase-Gens innerhalb dieses Clusters ist sinnvoll, da *G. oxydans* kein funktionelles Phosphotransferasesystem zur Phosphorylierung von Glucose aufweist. Die über die Cytoplasmamembran eingeschleuste Glucose kann aber mittels der Glucokinase zu Glucose-6-phosphat aktiviert werden. Die Organisation der ORFs 0321 (Glucokinase), 0318 (Lactonase), 0316 (6-Phosphogluconat-Dehydrogenase), 0320 (Isomerase), 0314 (Transketolase) und 0315 (Transaldolase) in einem Cluster deuten weiterhin an, dass diese ORFs auch zusammen als Operon transkribiert und reguliert werden.

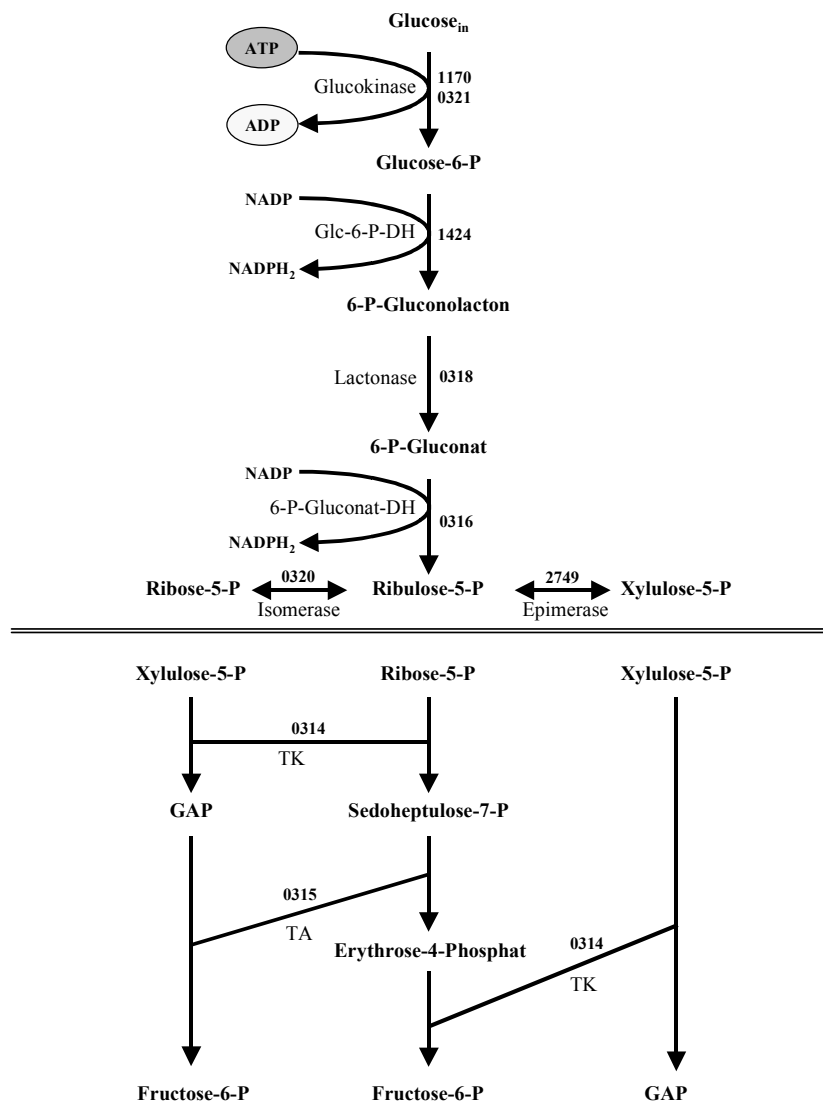


Abb. 24: Die Übersicht über den Pentosephosphat-Weg

Die Nummern geben jeweils den ORF an. Abkürzungen: -P: phosphat; GAP: Glycerinaldehydphosphat; TK: Transketolase; TA: Transaldolase.

Mit Hilfe des Pentosephosphat-Weges ist *G. oxydans* in der Lage, Intermediate des intrazellulären Stoffwechsels aufzubauen. Das durch die Reaktion der 6-Phosphogluconat-Dehydrogenase (ORF 0316) entstehende Ribulose-5-phosphat wird durch eine Isomerase (ORF 0320) in Ribose-5-phosphat umgesetzt. Dieses dient z. B. als Ausgangsprodukt für die Biosynthese von Histidin. Im Zuge der reversiblen Reaktionen der Transaldolase (ORF 0315) und Transketolase (ORF 0314) wird Erythrose-4-phosphat gebildet, welches Ausgangssubstrat für die Biosynthese für die aromatischen Aminosäuren ist. Ebenfalls durch die Reaktionen der Transaldolase und Transketolase wird Glycerinaldehydphosphat gebildet. Dieses kann in fünf Reaktionsschritten über Phosphoenolpyruvat zu Pyruvat umgesetzt werden. Die Reaktionen entsprechen denen, die im unteren Teil der Glykolyse ablaufen (s. 3.6.2.1). Ausgehend von Pyruvat erfolgt z. B. die Synthese von Alanin, Valin und Leucin. Darüber hinaus wird Pyruvat durch die Reaktion der Pyruvat-Dehydrogenase zu Acetyl-CoA umgesetzt, welches in den (unvollständigen) Zitratzyklus eingespeist werden kann. Der oxidative Pentosephosphat-Weg kann aber auch zur vollständigen Oxidation von Glucose genutzt werden, wenn genügend Synthesebausteine vorliegen und nur NADPH produziert werden soll. In diesem Fall werden drei Moleküle Glucose über 6-Phosphogluconat und die phosphorylierten Pentosen letztendlich zu zwei Molekülen Fructose-6-phosphat und einem Molekül Glycerinaldehydphosphat oxidiert, so dass insgesamt eine C3-Einheit vollständig zu CO₂ und NADPH umgesetzt wird. Die verbleibenden Intermediate können nach Umwandlung in Glucose-6-phosphat wieder in den Stoffwechselweg eingebracht werden.

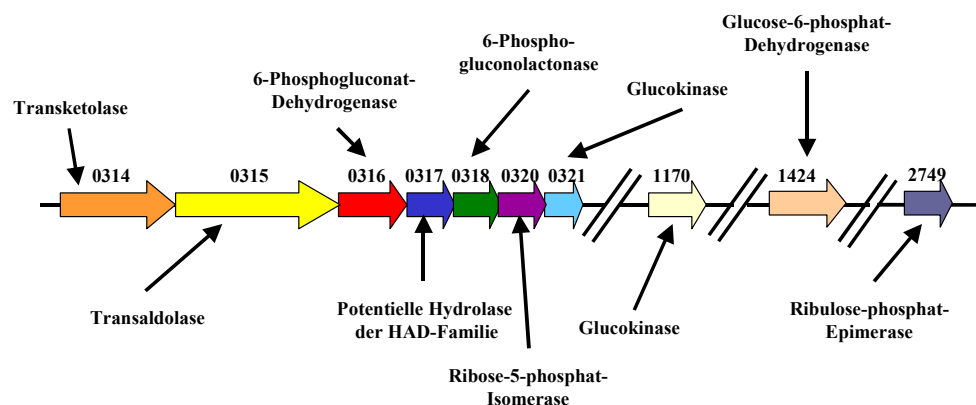


Abb. 25. Darstellung der ORFs, die für den Pentosephosphat-Weg notwendig sind

Die Zahlen geben die entsprechende ORF-Nummer an. Zu erkennen ist, dass Gene, die für den Pentosephosphat-Weg spezifisch sind, in einem Cluster vorliegen.

Im Zuge der Reaktionen der Glucose-6-phosphat-Dehydrogenase (ORF 1424) und der 6-Phosphogluconat-Dehydrogenase (ORF 0316) entsteht NADPH. Die Regeneration von NADP kann in *G. oxydans* möglicherweise durch Transhydrogenasen (s. 3.6.4) erfolgen. Dabei wird zum einen NADP regeneriert, zum anderen kann das gebildete NADH in die Atmungskette (s. 3.6.5.1) eingeschleust werden, wodurch letztendlich ein Energiegewinn erfolgt.

3.6.2.3 Entner-Doudoroff-Weg (KDPG-Weg)

Der Entner-Doudoroff-Weg ist innerhalb der Bakterien weit verbreitet. Er wird vor allem von Bakterien genutzt, denen die Phosphofruktokinase als Schlüsselenzym der Glykolyse fehlt. Bislang ist in der Literatur nichts darüber beschrieben, dass die Enzyme des Entner-Doudoroff-Weges in *G. oxydans* vorhanden sind. Die Analyse des Genoms zeigte jedoch, dass sämtliche ORFs, die für die entsprechenden Enzyme kodieren, in dem Genom zu finden sind. In Abb. 26 ist der Entner-Doudoroff-Weg schematisch dargestellt, wobei die Zahlen an den Reaktionspfeilen die entsprechenden ORF-Nummern wiedergeben. Die Reaktionen unterscheiden sich von denen der Glykolyse nur im oberen Teil. Ab der Stufe des Glycerinaldehyd-3-phosphats entsprechen die Reaktionsschritte denen der Glykolyse. Hiermit wird also die Reaktion der Phosphofruktokinase umgangen. Die Reaktionen im oberen Teil des Entner-Doudoroff-Weges entsprechen zunächst denen des Pentosephosphat-Weges. Im Gegensatz dazu wird 6-Phosphogluconat aber nicht zu Ribulose-5-phosphat umgesetzt, sondern über das für diesen Stoffwechselweg charakteristische Intermediat 2-Keto-3-Desoxy-Phosphogluconat zu Pyruvat und Glycerinaldehydphosphat umgewandelt. Die Gene, die für die spezifischen Enzyme dieses Stoffwechselwegs kodieren (Phosphogluconat-Dehydratase, ORF 1757 und KDPG-Aldolase, 1756) sind in dem Genom von *G. oxydans* in einem Cluster oder Operon organisiert. Es liegt daher der Schluss nahe, dass diese beiden Gene auch gemeinsam reguliert werden.

Betrachtet man diesen Stoffwechselweg in der Übersicht, so erscheint es für die Zelle durchaus von Vorteil zu sein nicht den Weg der Glykolyse, sondern den des KDPG-Wegs zu beschreiten. Hierbei kann in nur zwei Reaktionsschritten, nämlich dem der Phosphogluconat-Dehydratase- und der KDPG-Aldolase-Reaktion, aus Gluconat

Pyruvat hergestellt werden. Pyruvat dient in der Zelle als Vorläufer für viele weitere Metabolite (z. B. für Alanin und Valin).

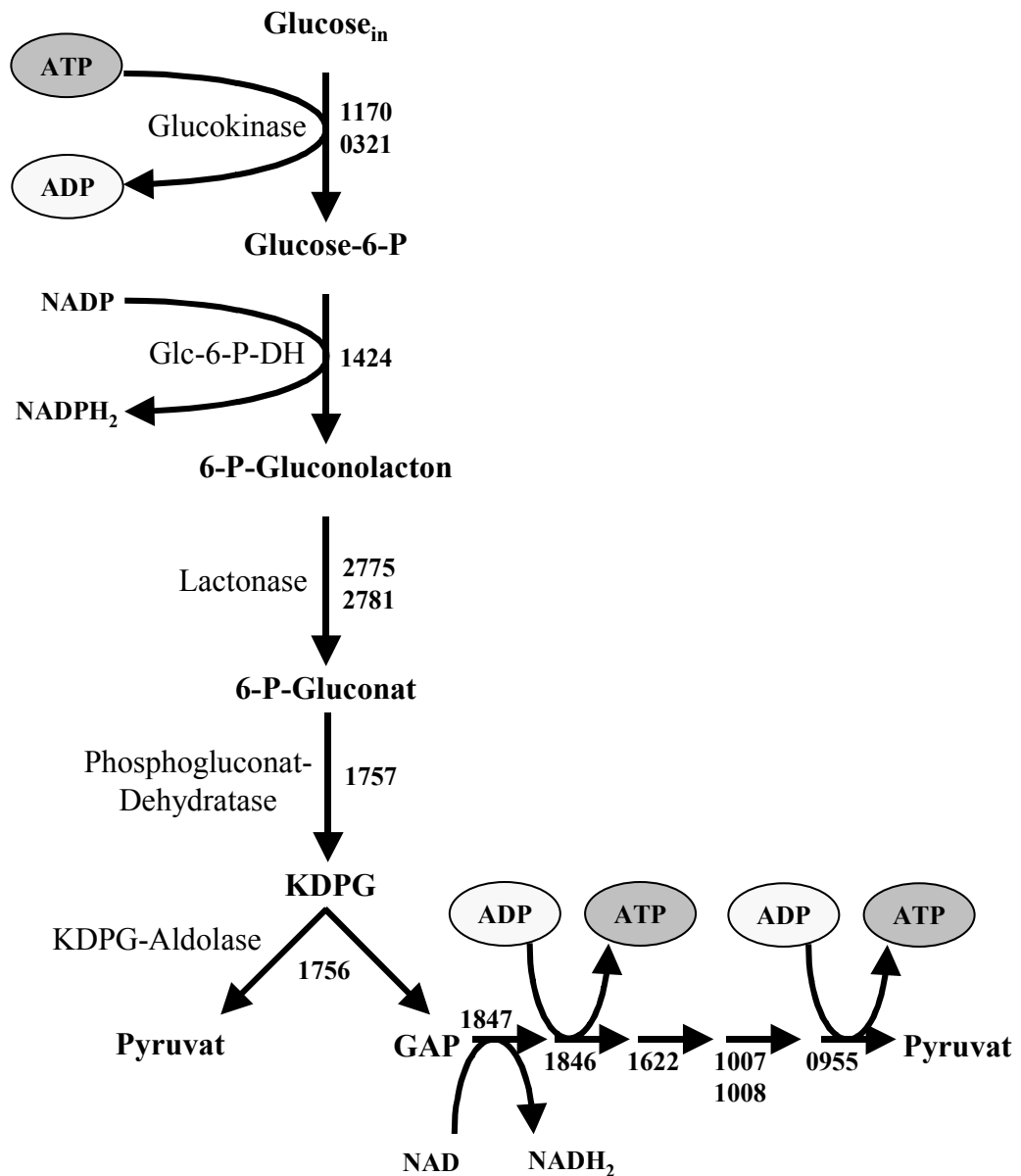


Abb. 26: Schematische Darstellung des Entner-Doudoroff-Wegs

Die Zahlen geben die entsprechende ORF-Nummer an. Im oberen Teil der Abb. sind an den Reaktionspfeilen die jeweiligen Enzyme angegeben. Die Reaktionen vom GAP bis zum Pyruvat sind analog zur Glykolyse. Abkürzungen: P: -phospho, Phosphat; Glc-6-P-DH: Glucose-6-phosphat-Dehydrogenase; KDPG: 2-Keto-3-Desoxy-6-Phosphogluconat; GAP: Glycerinaldehyd-3-phosphat.

3.6.2.4 Zitratzyklus

Schon vor Jahrzehnten (STOUTHAMER, 1959) wurde postuliert, dass *Gluconobacter*-Stämme nicht im Besitz eines vollständigen Zitronensäurezyklus sind. Zahlreiche biochemische Untersuchungen der letzten Jahrzehnte bestätigten dieses zumindest insofern, dass *G. oxydans* nicht alle Reaktionsschritte des Zitratzyklus durchführen kann. Ungeklärt war bislang jedoch die Frage nach dem Grund dafür. Die Analyse des Genoms ist ein sehr gutes Werkzeug, um genau solche Fragen zu klären. Eine Möglichkeit für einen unvollständigen Zitratzyklus wäre, dass im Genom zwar potentiell alle ORFs vorhanden sind, diese aber nicht exprimiert werden, was dazu führen würde, dass die entsprechenden Enzyme biochemisch nicht nachweisbar wären. Eine andere Möglichkeit besteht darin, dass nicht alle Gene im Genom vorhanden sind und somit die korrespondierenden Enzyme nicht gebildet werden. Dieses ist bei *G. oxydans* der Fall. Die Analyse des Chromosoms zeigte, dass sowohl ein ORF, der für die Succinat-Thiokinase, als auch ein ORF, der für die Succinat-Dehydrogenase kodiert im Genom nicht vorkommen. Dieser Mangel an essentiellen Genen hat drastische Auswirkungen auf den gesamten Stoffwechsel von *G. oxydans*. Ein nicht-kompletter Zitratzyklus impliziert, dass viele Substrate nicht vollständig oxidiert werden können, da Acetyl-CoA als Schlüsselintermediat der zuvor dargestellten Stoffwechselwege nicht zu CO₂ oxidiert werden kann (Abb. 27). Somit stellt der Zitratzyklus in *G. oxydans* nur eine Quelle für Intermediate des Anabolismus dar. Datenbankabgleiche zeigten, dass der Organismus über die Gene verfügt, die für die Enzyme des oxidativen Zweiges des Zitratzyklus kodieren. Somit kann *G. oxydans* mit Hilfe der Zitrat-Synthase (0650), der Aconitase (ORF 2730), der Isozitrat-Dehydrogenase (ORF 2731) und der 2-Ketoglutarat-Dehydrogenase (ORF 2260, 2460) Zitrat bis zum Succinyl-CoA umsetzen. Das letztgenannte Intermediat hat eine wichtige Funktion für die Biosynthese von Porphyrinen. Ein weiteres wichtiges Zwischenprodukt ist 2-Ketoglutarat, das im Stickstoffwechsel als Akzeptor von NH₄⁺ im Zuge der Reaktionen der Glutamat-Synthase und der Glutamin-Synthetase dient. Für den reduktiven Teil des Zitratzyklus finden sich in *G. oxydans* eine membrangebundene Malat-Dehydrogenase (ORF 3459) und eine Fumarase (ORF 0242), deren Aktivität zur Bildung von Fumarat führen. Die in diesem Teil der Reaktionsfolge gebildeten Intermediate sind ebenfalls wichtig für Biosynthesen. Ausgehend von Oxalacetat erfolgt die Biosynthese der Aminosäuren Aspartat, Asparagin, Lysin, Threonin und Methionin. Fumarat stellt ein weniger

wichtiges Intermediat für Biosynthesen dar. Es ist ein Produkt, das beim Abbau der aromatischen Aminosäuren Tyrosin und Phenylalanin entsteht. Da allerdings der Zyklus nicht geschlossen ist, muss Oxalacetat über anaplerotische Reaktionen neu gebildet werden. In der Regel erfolgt dieses durch die Reaktion der Pyruvat-Carboxylase, welche die Reaktion von Pyruvat zu Oxalacetat katalysiert oder durch eine Phosphoenolpyruvat-Carboxylase, die Phosphoenolpyruvat zu Oxalacetat umsetzt. In dem Genom von 621H kodiert der ORF 3479 für eine putative Phosphoenolpyruvat-Carboxylase, die eine derartige anaplerotische Reaktion bewerkstelligen könnte.

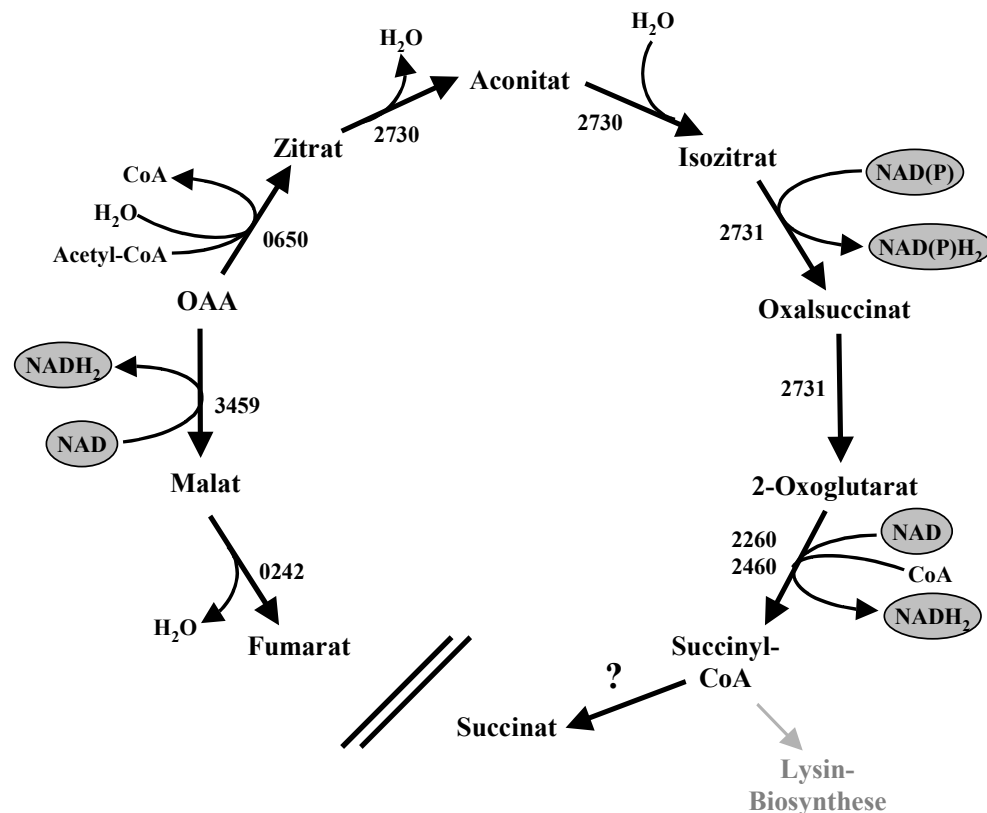


Abb. 27: Reaktionsschritte des Zitratzyklus

Dargestellt sind nur diejenigen Reaktionen, die *G. oxydans* durchföhren kann. Der Zitratzyklus ist in *Gluconobacter* kein Kreislauf, sondern stellt nur ein „Hufeisen“ dar. Im Genom fehlen ORFs, die für Gene kodieren, welche für die Succinat-Thiokinase und die Succinat-Dehydrogenase kodieren. Die Zahlen geben die entsprechenden ORF-Nummern an. Abkürzungen: OAA: Oxalacetat.

Die oben beschriebenen Reaktionen des Zitratzyklus sind in Abb. 27 schematisch dargestellt. Die entsprechenden ORFs, welche für die Enzyme des Zitratzyklus kodieren sind mit Ausnahme der Aconitat-Hydratase und der Isozitat-Dehydrogenase nicht vereint, sondern liegen einzeln in dem Genom. Bei der Analyse des Genoms konnten keine ORFs identifiziert werden, die für Proteine zur Hydrolyse von Succinyl-CoA oder zur Oxidation von Succinat kodieren. Das Fehlen eines ORFs, der für die Succinat-Dehydrogenase kodiert, steht im Einklang mit den Befunden von GREENFIELD & CLAUS (1972), die demonstrierten, dass in *G. oxydans* eine Aktivität der Succinat-Dehydrogenase fehlt. Dieses Enzym ist jedoch nicht nur Teil des Zitratzyklus, sondern es ist zudem ein membrangebundenes Enzym, das den Komplex II der Atmungskette darstellt. Das bedeutet, dass *G. oxydans* den Zitratzyklus nur dazu nutzen kann, Intermediate z. B. für den Aufbau von Aminosäuren zu synthetisieren. Er hat aber nicht die Funktion, den mit der Atmungskette verbundenen Ubichinonpool direkt mit Elektronen zu versorgen. Darüber hinaus kann Acetyl-CoA, wie oben beschrieben, nicht zu CO₂ aufoxidiert werden. Das bedeutet, dass Substrate nicht vollständig oxidiert werden können. Möglicherweise ist dies eine Erklärung für die unvollständigen Oxidationsreaktionen, die für *G. oxydans* charakteristisch sind.

3.6.3 Biosynthese der Aminosäuren

Die Biosynthese der Aminosäuren erfolgt über unterschiedliche Synthesewege. Jedoch stammt das Kohlenstoffgerüst immer aus Zwischenprodukten der Glykolyse, des Pentosephosphat-Weges oder des Zitratzyklus. In den vorangegangenen Kapiteln der vorliegenden Arbeit wurde gezeigt, dass *G. oxydans* entscheidende Enzyme des zentralen Stoffwechsels fehlen. Darunter sind vor allem Enzyme der Glykolyse und des Zitratzyklus. Dennoch ist dieser Organismus in der Lage, sämtliche Aminosäuren zu synthetisieren. In den folgenden Kapiteln werden die grundlegenden Synthesewege einschließlich der dazugehörigen ORFs der Aminosäurebiosynthese vorgestellt.

3.6.3.1 Biosynthese von Histidin

Die Biosynthese der Aminosäure Histidin zweigt vom Pentosephosphat-Weg ab. Ausgehend von Ribose-5-phosphat erfolgt durch die Ribose-phosphat-Pyrophosphokinase (ORF 2351) zunächst die Synthese von 5-Phosphoribosyl-1-

pyrophosphat (PRPP). Das letztgenannte Produkt dient anschließend zur Synthese von Histidin (Abb. 28).

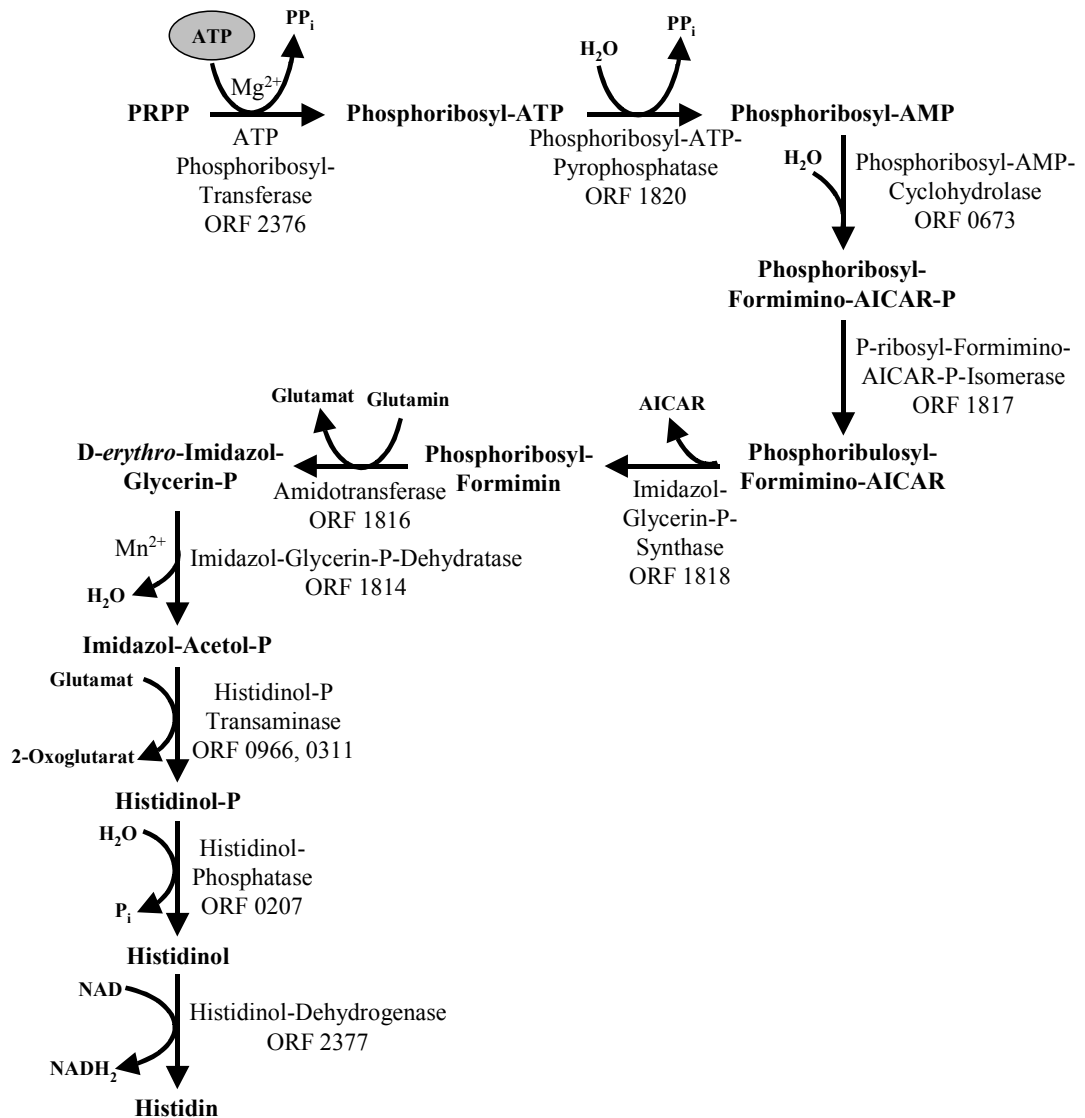


Abb. 28: Biosynthese von Histidin in *G. oxydans*

Mit Ausnahme der Reaktion der Phosphoribosyl-ATP-Pyrophosphatase (ORF 1820) handelt es sich bei den dargestellten Reaktionsschritten um reversible Reaktionen, die in beide Richtungen verlaufen können. Abkürzungen: PRPP: 5-Phosphoribosyl-1-pyrophosphat; AICAR: 5'-Phosphoribosyl-5-Amino-4-Imidazol-Carboxamid.

PRPP ist nicht nur ein Intermediat für die Biosynthese von Histidin, sondern stellt auch das Ausgangsprodukt für die Purinbiosynthese dar. Bei der Biosynthese von Histidin reagiert PRPP mit dem Purinring von ATP, der anschließend aufgebrochen wird. Nach

der Isomerisierung durch die Phosphoribosyl-Formimino-AICAR-phosphat-Isomerase (ORF 1817) entsteht das Produkt Phosphoribulosyl-Formimino-AICAR (AICAR: 5'-Phosphoribosyl-5-Amino-4-Imidazol-Carboxamid). Im nächsten Reaktionsschritt wird AICAR, das ebenfalls ein Zwischenprodukt der Purinbiosynthese ist, abgespalten. Nach der Übertragung einer Aminogruppe auf Phosphoribosyl-Formimin wird der Imidazolring, als Bestandteil von Histidin, geschlossen und es entsteht *D-erythro*-Imidazol-Glycerinphosphat. In vier weiteren Reaktionsschritten entsteht schließlich die Aminosäure Histidin (Abb. 28). Die ORFs, die für die Enzyme kodieren, welche in die Biosynthese von Histidin involviert sind, lagen zum Teil einzeln im Genom vor. Fünf ORFs, die für die entsprechenden Enzyme der Histidin-Biosynthese kodieren, befanden sich jedoch in einem Cluster (Abb. 29).

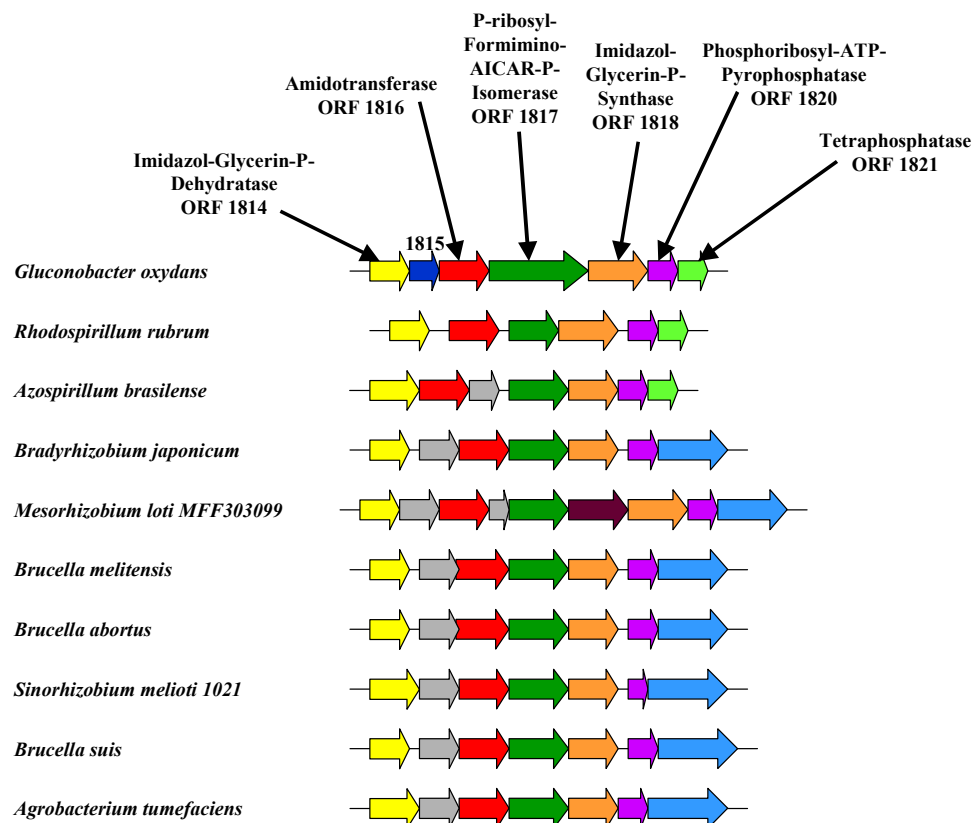


Abb. 29: Organisation des Histidin-Operons

Die Abb. zeigt die in einem Operon organisierten Gene der Histidin-Biosynthese. Gleiche Farben der Pfeile bedeuten gleiche Funktion. ORFs, die auch in *G. oxydans* Bestandteil des Clusters sind, sind in der Abb. beschriftet. Dem ORF 1815 konnte keine Funktion zugewiesen werden. Die Tetraphosphatase (ORF 1821) ist Bestandteil des Operons, ist jedoch nicht in die Biosynthese von Histidin involviert. Farben der nicht beschrifteten Pfeile: grau: hypothetischer ORF; hellblau: Panthotenat-Kinase; braun: Argininase.

Darüber hinaus beinhaltet dieses Cluster auch Gene, die nicht in die Synthese von Histidin involviert sind. ORF 1815, der Bestandteil dieses Clusters ist, zeigt keinerlei Ähnlichkeiten zu bekannten Proteinen. Ein Vergleich zur Organisation des Histidin-Operons in anderen Organismen zeigte jedoch, dass an dieser Position auch in einigen anderen Bakterien (z. B. *Bradyrhizobium japonicum*, KANEKO et al., 2002) ein zusätzlicher ORF ohne zugeordnete Funktion vorhanden ist (Abb. 29). Der ORF 1815 zeigte zu diesen ORFs jedoch keinerlei Ähnlichkeiten, wohingegen die hypothetischen Proteine in *Brucella sp.* und *Mesorhizobium loti* (KANEKO et al., 2000) untereinander Ähnlichkeiten aufweisen. Am Ende dieses Operons befindet sich außerdem noch ein ORF, der für eine Tetraphosphatase kodiert, die in den Purin- und Pyrimidin-Metabolismus involviert ist. Dieser Befund zeigte sich auch in *Rhodospirillum rubrum* und *Azospirillum brasilense* (FANI et al., 1989). Über die Funktion einer derartigen Operon-Organisation ist jedoch nichts bekannt. Bei *Brucella sp.* und einigen anderen Organismen (Abb. 29) kodiert der letzte ORF dieses Operons für eine Panthotenat-Kinase (PAULSEN et al., 2002). Dieses Enzym ist unter anderem in die Biosynthese von CoA involviert.

3.6.3.2 Biosynthese der aromatischen Aminosäuren Phenylalanin, Tyrosin und Tryptophan

Allgemein muss angemerkt werden, dass die Vorstufen zur Synthese der aromatischen Aminosäuren Phenylalanin, Tyrosin und Tryptophan aus dem Shikimat-Weg stammen (Abb. 30). Die Synthese beginnt mit der Kondensation von Phosphoenolpyruvat und Erythrose-4-phosphat, einem Intermediat aus dem Pentosephosphat-Weg. Das Produkt 7-Phospho-3-desoxyarabinoheptulose-7-phosphat wird durch die Dehydrochinat-Synthase zyklisiert und es entsteht 3-Dehydrochinat. Durch die Reaktionen der Dehydrochinat-Dehydratase (ORF 1763) und der 3-Dehydroshikimat-Dehydrogenase (ORF 0607) entsteht Shikimat. Dieses wird in zwei weiteren Reaktionen zu Chorismat umgesetzt. Chorismat stellt nicht nur ein Verzweigungsprodukt der aromatischen Aminosäuren dar, sondern ist unter anderem auch Zwischenprodukt bei der Synthese von Ubichinon. Für die Synthese von Phenylalanin und Tyrosin wird Chorismat durch eine Mutase (ORF 0652) zu Prephenat umgesetzt. Durch eine Dehydratase-Reaktion (ORF 1335) und anschließende Transaminierung wird Phenylalanin gebildet. Über eine

Dehydrogenierungs-Reaktion (ORF 0956) von Prephenat und nachfolgender Transaminierung erfolgt die Synthese von Tyrosin.

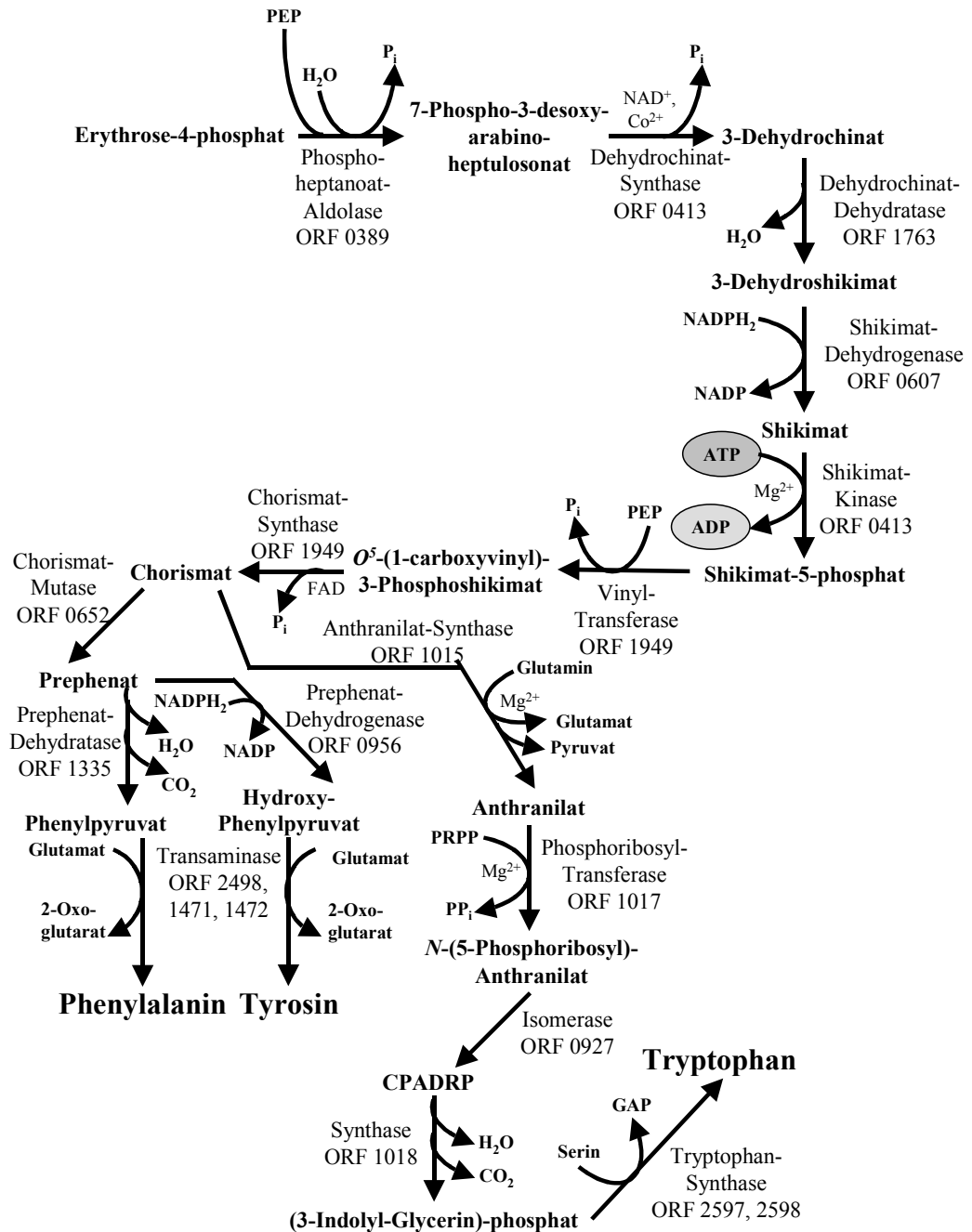


Abb. 30: Synthese der aromatischen Aminosäuren

Die Biosynthese der aromatischen Aminosäuren beginnt mit dem gemeinsamen Shikimat-Weg, dessen Endprodukt Chorismat ist. Davon ausgehend erfolgt die Biosynthese von Phenylalanin, Tyrosin und Tryptophan. Abkürzungen: PEP: Phosphoenolpyruvat; PRPP: 5-Phosphoribosyl-1-pyrophosphat; CPADRP: 1-(2-Carboxyphenylamino)-1'-Desoxyribulose-5-phosphat; GAP: Glycerinaldehyd-phosphat.

Bei der Synthese von Tryptophan wird Chorismat durch die Abspaltung von Pyruvat und gleichzeitiger Aminierung zu der aromatischen Verbindung Anthranilat umgesetzt. Katalysiert wird diese Reaktion durch eine Anthranilat-Synthase, die in dem Genom von *G. oxydans* von ORF 1015 kodiert wird. In vier weiteren Reaktionsschritten entsteht schließlich Tryptophan. Mit Ausnahme von drei ORFs, welche für Gene der Tryptophan-Biosynthese kodieren (ORF 1015, 1017, 1018), sind die Gene nicht in einem Operon oder Cluster organisiert, sondern liegen einzeln im Genom vor.

3.6.3.3 Biosynthese von Serin, Glycin und Cystein

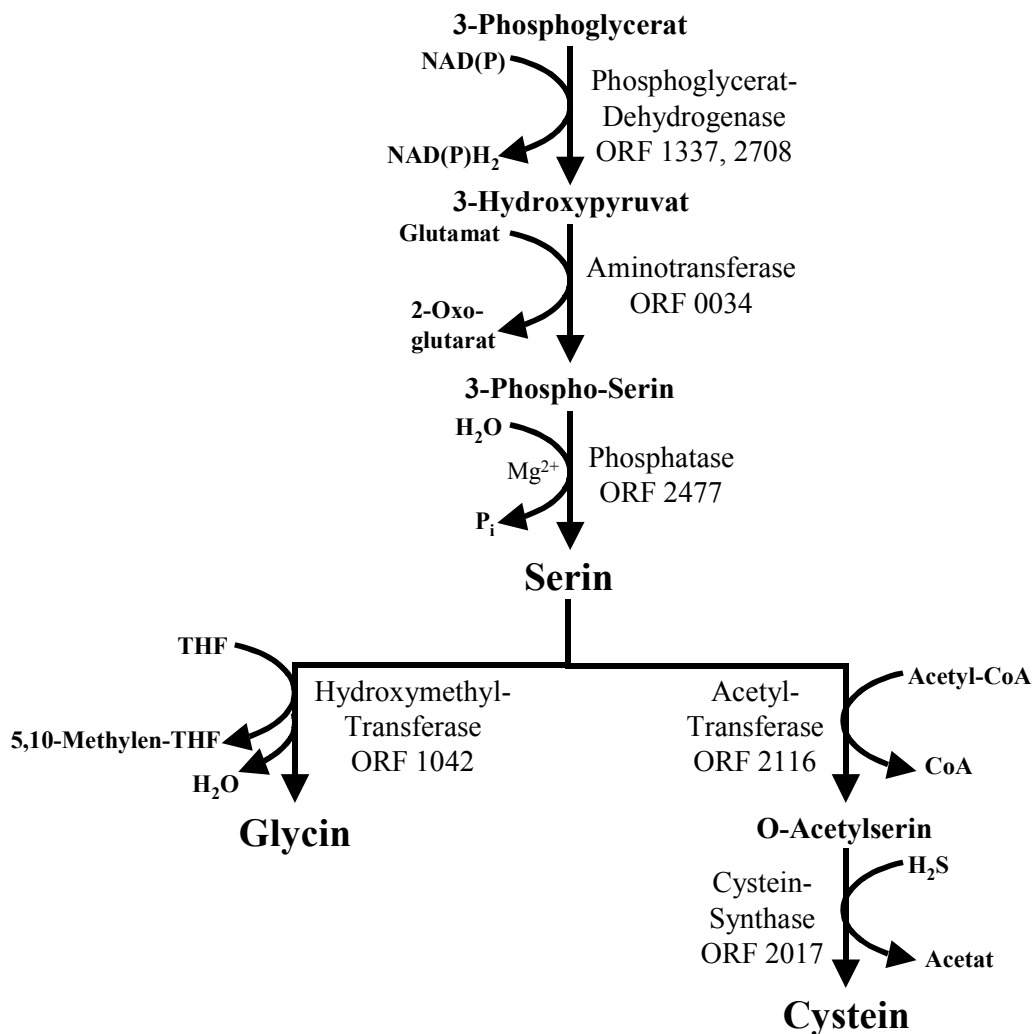


Abb. 31: Biosynthese von Serin, Glycin und Cystein

Die Biosynthese von Serin beginnt mit 3-Phosphoglycerat. Serin ist Vorläufer für die Synthese von Glycin und Cystein. Abkürzungen: THF: Tetrahydrofolat.

Die Biosynthese von Serin (Abb. 31) beginnt mit 3-Phosphoglycerat, welches in *G. oxydans* im Pentosephosphat-Weg entsteht. Serin kann durch Acetylierung und Austausch der Acetylgruppe gegen H₂S in Cystein umgewandelt werden. Die Thiolgruppe in Cystein spielt unter anderem eine Rolle bei der Ausbildung von Fe-S-Zentren in Elektronenüberträgern. Auch für die Bildung und Aufrechterhaltung von Sekundärstrukturen von Proteinen ist Cystein wichtig. Dabei wird Cystein zum Disulfid oxidiert (Cys-S-S-Cys). Die Bildung von Glycin erfolgt direkt aus Serin, wobei durch die Hydroxymethyl-Transferase (ORF 1042) eine Methylengruppe von Serin auf Tetrahydrofolat übertragen wird. Dabei entsteht 5,10-Methylen-Tetrahydrofolat. Die ORFs für die Biosynthese von Serin, Glycin und Cystein liegen jeweils einzeln in dem Genom von *G. oxydans* vor.

3.6.3.4 Biosynthese von Alanin, Leucin, Isoleucin und Valin

Alanin, Leucin, Isoleucin und Valin sind Aminosäuren mit aliphatischen Seitenketten. Diese nicht-reaktiven Gruppen tragen entscheidend zu hydrophoben Interaktionen in Proteinen bei. Alanin ist darüber hinaus Bestandteil der Mureinschicht. Das Kohlenstoffgerüst dieser Aminosäuren stammt von Pyruvat ab (Abb. 32). Alanin entsteht direkt durch Transaminierung aus Pyruvat (ORF 0113). Die Biosynthese von Leucin, Isoleucin und Valin ist eng miteinander verbunden. Im Fall von Valin und Isoleucin beginnt die Reaktionsfolge jeweils mit einer Kondensationsreaktion von zwei Molekülen Pyruvat bzw. von Pyruvat mit 2-Oxobutyrat, das aus einer Threonin Deaminierungsreaktion entstammt (Threonin-Dehydratase; ORF 1617). Das Enzym, das diese Reaktion katalysiert, wird als Acetolactat-Dehydrogenase bezeichnet und enthält als prosthetische Gruppe Thiaminpyrophosphat (ThPP; ORF 2480, 2481). Im Zuge der Synthese von Valin wird 2-Oxoisovalerat gebildet, das gleichzeitig auch als Vorstufe für die Produktion von Isoleucin verwendet wird. Bei dieser Reaktionsfolge wird das erwähnte Intermediat wiederum mit Acetyl-CoA kondensiert. Diese Reaktion wird durch eine Isopropylmalat-Synthase (ORF 0346) katalysiert und es entsteht 2-Isopropylmalat. Durch die katalytische Aktivität einer Dehydratase (ORF 1774, 1775) und einer Dehydrogenase (ORF 1473) wird 2-Oxo-Isocapronat gebildet. Die Transaminierung dieser Oxosäure liefert die Aminosäure Leucin. Im Zuge der Synthese von Isoleucin entsteht 2-Oxo-3-Methylvalerat. Die Synthese von Isoleucin und Valin erfolgt durch eine Transaminierungsreaktion der entstandenen Oxosäuren 2-

Oxoisovalerat bzw. Oxo-3-Methylvalerat. Der letzte Schritt der Synthese dieser drei Aminosäuren wird durch eine „branched chain aminotransferase“ katalysiert. Hierbei handelt es sich in *G. oxydans* immer um dasselbe Enzym, welches von ORF 1778 kodiert wird. Diese Tatsache zeigt deutlich, wie eng verwandt die Synthesewege dieser Aminosäuren sind. Die einzelnen Enzyme zur Biosynthese von Leucin und Valin sind identisch (Abb. 32), nur mit dem Unterschied, dass die für die Kondensationsreaktion notwendige 2-Oxosäure jeweils unterschiedlich ist.

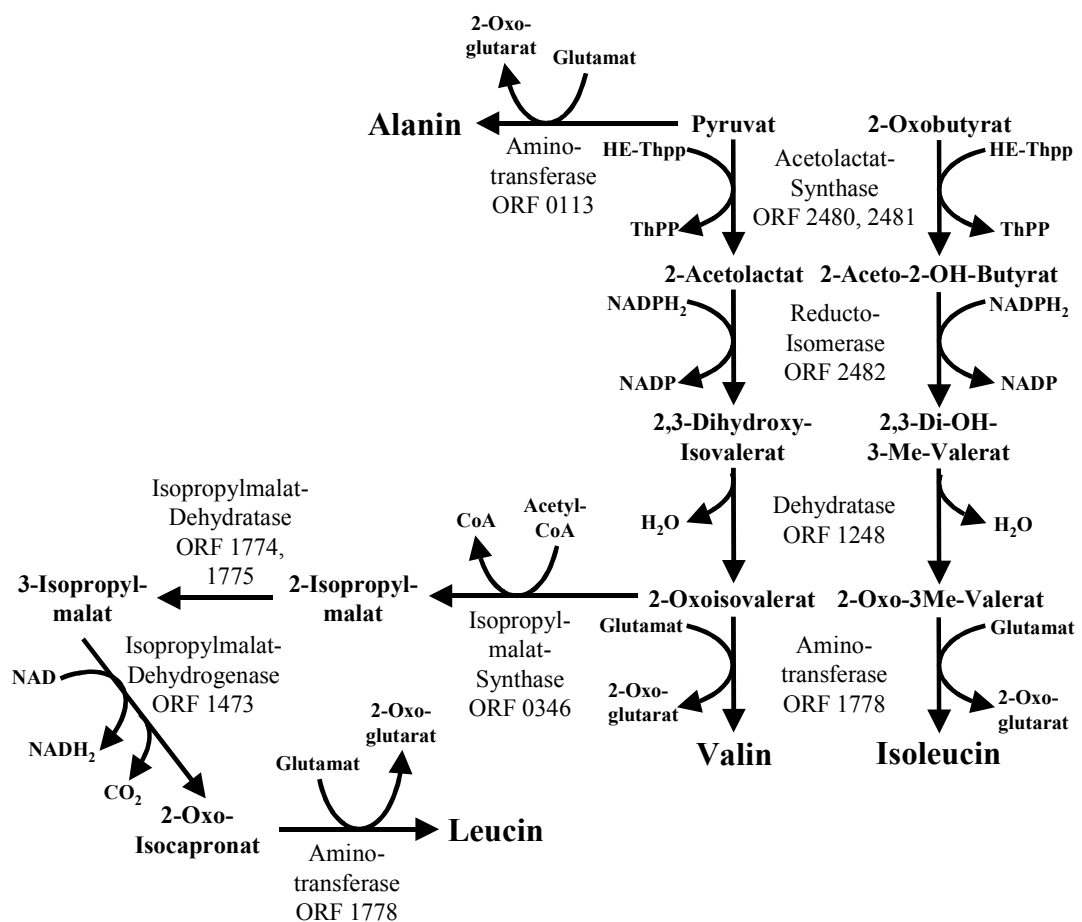


Abb. 32: Biosynthese der Aminosäuren Alanin, Valin, Leucin und Isoleucin in *G. oxydans*

3.6.3.5 Biosynthese von Aspartat, Asparagin, Lysin, Threonin und Methionin in *G. oxydans*

Grundsätzlich handelt es sich bei den Aminosäuren Aspartat, Asparagin, Lysin, Threonin und Methionin um völlig unterschiedliche Aminosäuren: Aspartat gehört mit seiner endständigen Carboxygruppe zu den „sauren“ Aminosäuren, die Amid-Gruppe von Asparagin trägt zu Wasserstoffbrückenbindungen bei, ebenso wie die Hydroxy-Gruppe in Threonin. Lysin ist aufgrund seiner Amino-Seitengruppe den „basischen“ Aminosäuren zuzuordnen und Methionin ist die universelle Start-Aminosäure (ATG). Gemeinsam ist diesen Aminosäuren jedoch, dass ihre Synthese von Oxalacetat, einem Zwischenprodukt des Zitratzyklus, ausgeht (Abb. 33). Aspartat entsteht direkt aus Oxalacetat durch eine Transaminasereaktion (ORF 1471, 1472). Durch die Reaktion der Asparagin-Synthetase (ORF 0613) wird Asparagin gebildet. Glutamin dient dabei als Aminogruppendonor.

Die Biosynthese von Lysin, Threonin und Methionin erfolgt durch die Aspartat-Kinase (ORF 1305), wobei Aspartat phosphoryliert wird. In einem weiteren Reaktionsschritt entsteht Aspartat-4-Semialdehyd, katalysiert durch eine Dehydrogenase (ORF 2255). Ausgehend von diesem Intermediat erfolgt die Biosynthese von Lysin, Threonin und Methionin. Im Falle der Synthese von Lysin erfolgt zunächst eine Kondensation von Aspartat-4-Semialdehyd mit Pyruvat und eine anschließende Zyklisierung, wobei Dihydrodipicolinat entsteht (ORF 1334). Durch die Reaktion einer Reduktase (ORF 2240) und einer Succinyl-Transferase (ORF 0465) wird *N*-Succinyl-2-amino-6-keto-Pimelat gebildet. Durch eine Transaminierungsreaktion erfolgt anschließend die Übertragung einer Aminogruppe auf dieses Intermediat. Im Zuge von drei weiteren Reaktionen, welche die Abspaltung von Succinat (ORF 0466, 0467), die Epimerisierung (ORF 0048) und Decarboxylierung (ORF 0591) der entsprechenden Intermediate einschließt, wird schließlich Lysin gebildet.

Im Zuge der Biosynthese von Threonin und Methionin erfolgt zunächst durch die Homoserin-Dehydrogenase (ORF 0112) eine Dehydrogenierung von Aspartat-4-Semialdehyd zu Homoserin. Durch eine Kinase- (ORF 1459) und eine Synthase-Reaktion (ORF 0508) wird Threonin gebildet. Für die Biosynthese von Methionin erfolgt zunächst die Übertragung einer Acetylgruppe auf Homoserin, wodurch Acetyl-

Homoserin entsteht. Anschließend wird durch eine Sulfhydrylase (ORF 0276) eine -SH-Gruppe eingefügt, wobei Cystein in der Regel als Donor fungiert, und Acetat wird abgespalten. Im nächsten Reaktionsschritt erfolgt die Synthese von Methionin durch eine Methyltransferase (ORF 0883) unter Abspaltung von Tetrahydropteroyl-tri-Glutamat.

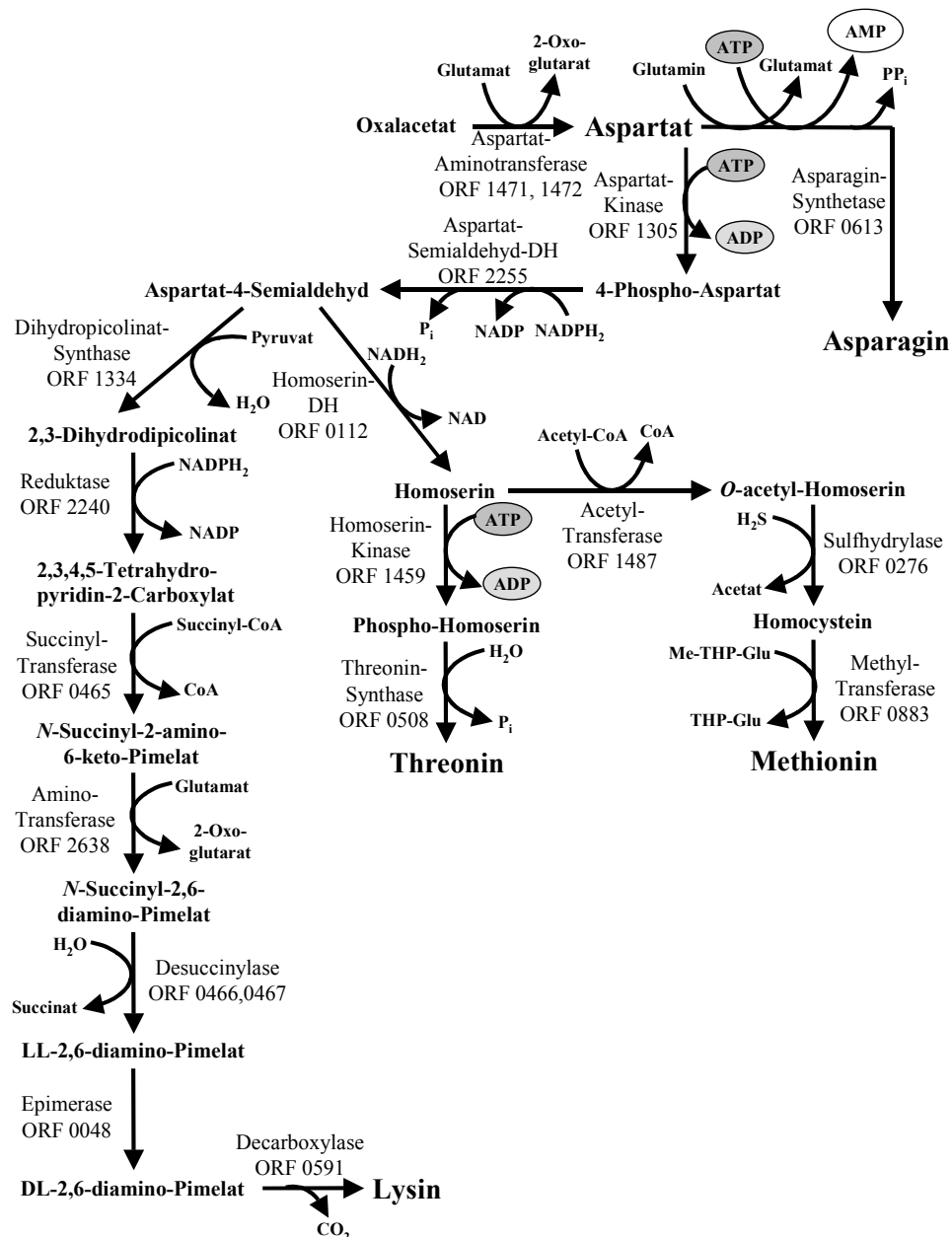


Abb. 33: Biosynthese von Aspartat, Asparagin, Lysin, Threonin und Methionin in *G. oxydans*

Abkürzungen: DH: Dehydrogenase; Me-THP-Glu: 5-Methyltetrahydropteroyl-tri-Glutamat; THP-Glu: Tetrahydropteroyl-tri-Glutamat.

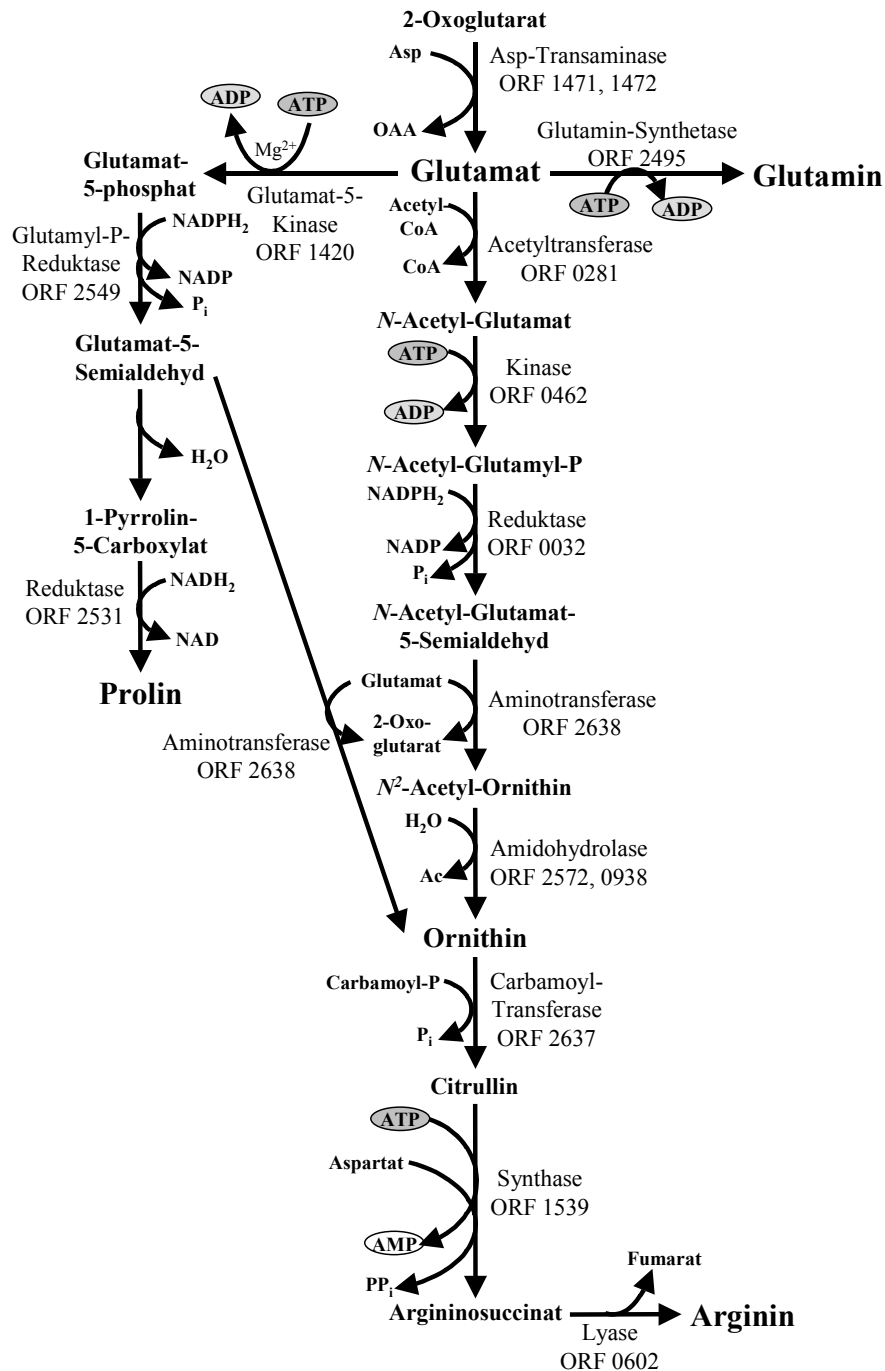
3.6.3.6 Biosynthese von Glutamin, Glutamat, Prolin und Arginin in *G. oxydans*

Abb. 34: Biosynthese von Glutamin, Glutamat, Prolin und Arginin in *G. oxydans*.

Die Biosynthese von Glutamin, Glutamat, Prolin und Arginin geht vom 2-Oxoglutarat (Abb. 34), einem Zwischenprodukt des Zitratzyklus, aus. Glutamat entsteht dabei durch

direkte Aminierung von 2-Oxoglutarat. Ausgehend von Glutamat werden die Aminosäuren Glutamin, Prolin und Arginin gebildet. Durch die Reaktion der Glutamat-Synthase (ORF 2495) erfolgt die Synthese von Glutamin. Für die Biosynthese von Prolin erfolgt eine Phosphorylierung von Glutamat durch die Glutamat-Kinase (ORF1420), wodurch Glutamat-5-phosphat entsteht. Dieses wird durch eine Reduktase (ORF 2549) zu Glutamat-5-Semialdehyd reduziert. Aus diesem Zwischenprodukt wird in einem nicht-enzymatischen Schritt 1-Pyrrolin-5-Carboxylat gebildet. Durch anschließende Reduktion (ORF 2531) dieses Intermediats erfolgt die Synthese von Prolin.

Das Intermediat Glutamat-5-Semialdehyd, das im Zuge der Biosynthese von Prolin entsteht, kann durch eine Transaminase-Reaktion zu Ornithin, einem Zwischenprodukt der Arginin Biosynthese umgesetzt werden. Ausgehend von Glutamat erfolgt für die Biosynthese von Arginin zunächst eine Kondensation mit Acetyl-CoA (ORF 0281), wodurch *N*-Acetyl-Glutamat entsteht. Durch anschließende Reduktion (ORF 0032) und Übertragung einer $-NH_2$ -Gruppe (ORF 2638) wird *N*²-Acetyl-Ornithin gebildet. Dieses Zwischenprodukt wird durch eine Amidohydrolase (ORF 2572, 0938) zu Ornithin umgesetzt. Durch eine Carbamoyl-Transferase (ORF 2637) wird dieses in Citrullin umgewandelt und über Argininosuccinat entsteht schließlich Arginin.

3.6.4 Transhydrogenasen

Viele der intrazellulären Reaktionen in *G. oxydans* sind NAD(P)-abhängig. Diese Abhängigkeit beinhaltet nicht nur die zentralen Stoffwechselwege, sondern z. B. auch die Reaktion der löslichen Glucose-Dehydrogenase. Schon lange wird über die Wege der Regeneration/Bereitstellung von NADP in *Gluconobacter* diskutiert. Auf der Grundlage der Genomdaten kann dieser Diskussionspunkt nunmehr im Detail betrachtet werden. Eine Möglichkeit zur Bildung von NADP aus NADPH sind Transhydrogenasen. Diese Enzyme katalysieren die reversible Oxidation von NADPH und übertragen dabei Reduktionsäquivalente auf NAD. Das so generierte NADH kann in die Atmungskette eingeschleust und zur Energiekonservierung genutzt werden. Bereits 1963 wurde von EAGON postuliert, dass *Gluconobacter* ein derartiges Enzym fehlt.

Im Gegensatz zu den oben genannten Hypothesen zeigte die Analyse des Chromosoms von *G. oxydans* jedoch, dass im Genom sehr wohl die Information für eine Transhydrogenase enthalten ist. Ein BLAST-Abgleich des Genoms ließ deutlich erkennen, dass die ORFs 1600, 1601 und 1602 Homologien zu bekannten Transhydrogenasen aufweisen.

3.6.4.1 Analyse der Transhydrogenase aus *G. oxydans*

Im Genom von *G. oxydans* wurden drei ORFs identifiziert, die Homologien zu bekannten Transhydrogenasen aufweisen, was dafür spricht, dass dieses Enzym aus drei Untereinheiten zusammengesetzt ist. Im Gegensatz dazu besteht die Transhydrogenase aus *E. coli* nur aus zwei Untereinheiten. In anderen Organismen, unter anderem in *Rhodospirillum rubrum* existiert jedoch eine Transhydrogenase, die ebenfalls aus drei Polypeptidketten zusammengesetzt ist (WILLIAMS et al., 1994). Es handelt sich jeweils um membrangebundene Enzyme, die protonentranslozierende Eigenschaften aufweisen. In Abb. 35 sind die in einem Cluster organisierten ORFs aus *G. oxydans*, gezeigt, die für die Transhydrogenase kodieren.

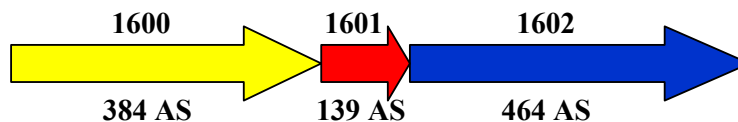


Abb. 35: ORFs, die für die Transhydrogenase in *G. oxydans* kodieren

Die Zahlen oberhalb der Pfeile geben jeweils die entsprechende ORF-Nummer an, die Zahlen unterhalb der Pfeile zeigen die Anzahl der aus den Basen resultierenden Aminosäuren.

Ein BLAST-Abgleich der Proteinsequenzen, die sich aus den ORFs ergeben, zeigte für ORF 1600 57 % Identität zu einer Transhydrogenase-Untereinheit aus *Burkholderia fungorum* und 42 % Identität zu einer entsprechenden Untereinheit aus *Rhodospirillum rubrum*. ORF 1601 wies kaum Ähnlichkeiten zu Proteinen aus *R. rubrum* auf, jedoch eine Identität von 75 % zu einer Transhydrogenase-Untereinheit aus *B. fungorum*. ORF 1602 war zu 57 % identisch zu einer Transhydrogenase-Untereinheit aus *B. fungorum* und wies 43 % Identität zu einer entsprechenden Untereinheit aus *R. rubrum* auf.

In der Literatur wird beschrieben, dass die Domäne I der Transhydrogenase (ORF 1600) die Bindestelle für NAD(H) beinhaltet, während die Bindestelle für NADP(H) in der Domäne III (ORF 1602) zu finden ist. Die Domäne II wird als hydrophobes, membrandurchspannendes Segment beschrieben (BIZOUARN et al., 1996). Zur Untersuchung einer möglichen membranassoziierten Lokalisation der Transhydrogenase aus *G. oxydans* wurde für alle drei ORFs eine Vorhersage der transmembranen Helices durchgeführt (Abb. 36).

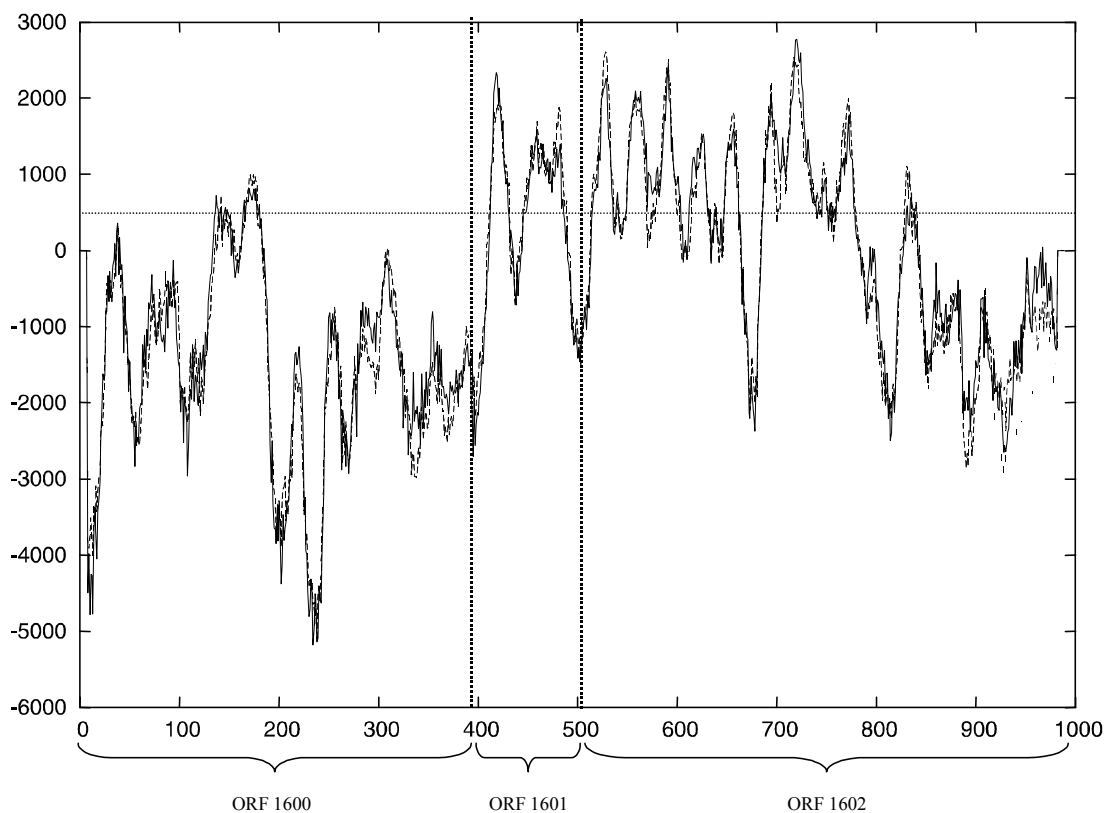


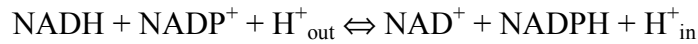
Abb. 36: Vorhersage der transmembranen Helices der aus ORF 1600-1602 resultierenden Proteine

Die Ordinate gibt einen Score für mögliche transmembrane Helices wieder. Scores, die größer als 500 sind werden als signifikant bezeichnet. Die Abszisse zeigt die Position der Aminosäuren, beginnend vom N-Terminus. Die gepunkteten vertikalen Linien trennen die einzelnen ORFs voneinander, die gepunktete horizontale Linie markiert den Cut-off (500) für die Wahrscheinlichkeit, dass es sich um eine transmembrane Helix handelt.

Bei der verwendeten Software -TMbase- handelt es sich um ein frei zugängliches Programm (http://www.ch.embnet.org/software/TMPRED_form.html; HOFMAN & STOFFEL, 1993), welches auch auf dem lokalen ERGO-Server (s.2.14.2) integriert ist. Als Eingabesequenz für den Plot diente die aus den ORFs 1600-1602 abgeleitete Aminosäuresequenz. Die gestrichelten vertikalen Linien in Abb. 36 trennen die einzelnen ORFs voneinander. Die gepunktete horizontale Linie markiert den Cut-off für die Wahrscheinlichkeit, dass es sich um eine transmembrane Helices handelt. Scores, die größer als 500 sind, die sich also oberhalb der gepunkteten horizontalen Linie befinden, werden als signifikant bezeichnet. Das aus ORF 1601 resultierende Protein stellt laut Literatur die membranassoziierte Domäne dar. Betrachtet man den Plot, so ist denkbar, dass auch das aus ORF 1601 resultierende Protein in der Membran assoziiert ist, da mindestens zwei transmembrane Helices vorhergesagt werden.

Im Gegensatz dazu zeigte ORF 1600 keine signifikanten transmembranen Helices. Etwas anders verhält sich der Plot für ORF 1602. Hier wäre es durchaus möglich, dass der N-terminale und der zentrale Teil des aus dem Gen resultierenden Proteins in der Membran verankert ist; möglich wäre aber auch, dass sich die hydrophoben Bereiche bei der Faltung im Innern des Proteins anordnen. Es sei jedoch noch einmal darauf hingewiesen, dass weder die Bildung einer aktiven Transhydrogenase, noch die Lokalisation dieser in der Zelle für *G. oxydans* biochemisch nachgewiesen wurde. Zusammen mit den Ergebnissen des BLAST-Abgleichs ist es aber wahrscheinlich, dass es sich um eine membrangebundene Transhydrogenase handelt. Bei der Annotation wurden sowohl ORF 1600 als auch 1601 als alpha-Untereinheit der Transhydrogenase bezeichnet, basierend auf Ähnlichkeiten zu anderen Organismen. Dieses kann auch ein Hinweis auf einen Frameshift sein, was bedeuten würde, dass eine erneute Editierung der Rohdaten in diesem Bereich notwendig wäre. Eine Überprüfung der Rohdaten ergab jedoch keinen Hinweis darauf, dass an dieser Stelle der Sequenz ein Fehler vorliegt. Eine andere Möglichkeit besteht darin, dass die hier beschriebene Transhydrogenase aufgrund einer Mutation nicht funktionstüchtig ist. Dieses kann aber nur durch biochemische Experimente gezeigt werden. Eine dritte Möglichkeit besteht jedoch darin, dass die Bindestelle für NAD(H) als separates Polypeptid vorliegt. Ein solcher Aufbau wurde bislang jedoch nur für das entsprechende Protein aus *R. rubrum* beschrieben (BIZOUARN et al., 1996). Die Ergebnisse deuten darauf hin, dass bei der Transhydrogenase aus *G. oxydans* eine ähnliche Zusammensetzung vorliegen könnte.

In der Regel sind membrangebundene Transhydrogenasen nicht am Aufbau eines Membranpotentials Δp beteiligt, sondern verbrauchen dieses für die Bereitstellung von NADPH (JACKSON, 2003):



Im Gegensatz zu vielen anderen Organismen kann *G. oxydans* NADPH nicht nur durch den Pentosephosphat-Weg sondern auch durch im Cytoplasma stattfindende unvollständige Oxidationen generieren. Als Beispiel sei hier die Produktion von 5-Ketogluconat durch die NADP-abhängige Gluconat-Dehydrogenase aufgeführt (Abb. 37). Aufgrund derartiger Reaktionen der unvollständigen Oxidation liegt der Schluss nahe, dass *G. oxydans* weniger ein Problem damit hat, NADPH für anabolische Reaktionen zur Verfügung zu stellen, als NADP zu regenerieren. Die Genomdaten liefern Hinweise darauf, dass die Regeneration von NADP durch eine Transhydrogenase erfolgen könnte, die gleichzeitig am Aufbau des Membranpotentials beteiligt ist.

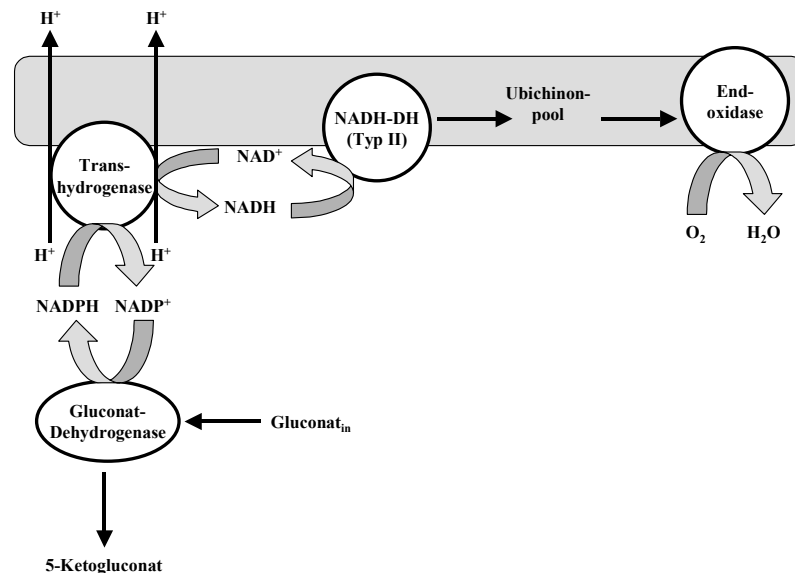


Abb. 37: Reaktion der cytoplasmatischen Gluconat-Dehydrogenase

Bei der Oxidation von Gluconat im Cytoplasma wird NADPH generiert. Durch die Reaktion der Transhydrogenase entsteht NADH. Die Oxidation von NADH erfolgt durch eine NADH Dehydrogenase von Typ II, wobei keine Protonentranslokation erfolgt. Die Elektronen fließen vom Ubichinonpool auf die entsprechenden Endoxidasen.

3.6.5 Membrangebundene Prozesse

Die bei *G. oxydans* fehlenden Enzyme der zentralen Stoffwechselwege, die z. B. zu dem unvollständigen Zitratzyklus führen, zwingen diesen Organismus zur unvollständigen Oxidation der entsprechenden Substrate. Diese Reaktionen werden in der Regel von membrangebundenen Dehydrogenasen katalysiert. Dabei ragt das aktive Zentrum der Enzyme in den periplasmatischen Raum, so dass ein Transport der entsprechenden Substrate über die Cytoplasmamembran nicht notwendig ist. Die bei den unterschiedlichen Dehydrogenierungsreaktionen frei werdenden Elektronen fließen über Cytochrome oder den Ubichinonpool in die Atmungskette ein.

3.6.5.1 Rekonstruktion der Atmungskette

Die zentrale Atmungskette von *G. oxydans* besteht aus einer NADH-Dehydrogenase von Typ II, von der Elektronen aus cytoplasmatischen Dehydrogenierungsreaktionen auf den Ubichinonpool fließen und aus zwei Chinol-Oxidasen von Typ *bo₃* und *bd*. Darüber hinaus zeigte das Genom das Vorhandensein eines *bc₁*-Komplexes (Abb. 38). Im Folgenden werden die einzelnen Proteinkomplexe kurz vorgestellt.

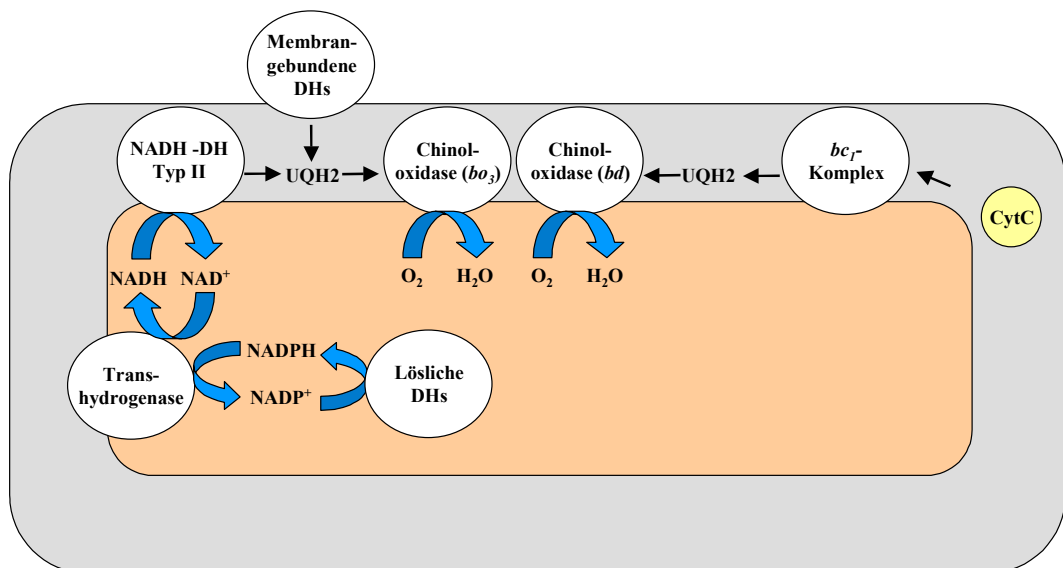


Abb. 38: Rekonstruktion der Atmungskette von *G. oxydans*

Die Oxidation von NADH erfolgt durch eine NADH-Dehydrogenase Typ II, bei der keine Protonentranslokation erfolgt. Die Elektronen werden auf den Ubichinonpool (UQH₂) übertragen und fließen auf die entsprechenden Chinol-Oxidasen vom Typ *bo₃* oder *bd*.

Grundsätzlich werden die Elektronen vom Ubichinonpool auf die entsprechenden Endoxidasen übertragen, die molekularen Sauerstoff als terminalen Elektronenakzeptor verwenden. Die Sequenzdaten von *G. oxydans* deuten an, dass dieser Organismus Chinol-Oxidasen vom Typ *bo₃* und *bd* besitzt. Für die *bo₃*-Typ Oxidase liegen bereits biochemische Daten vor, die belegen, dass das Enzym aus *G. suboxydans* eine hohe Aktivität aufweist (MATSUSHITA et al., 1994). Die Cytochrom *bo₃*-Oxidase dient in vielen Bakterien als terminales Enzym in der Atmungskette, wenn die Sauerstoffkonzentration im Medium hoch ist. Das membrangebundene Protein besteht aus vier Untereinheiten, wobei die Untereinheit 1 ein low-spin Häm (Häm b) und ein high-spin Häm (Häm O₃) enthält, das mit dem mononuklearen Kupfer Zentrum (CuB) das aktive Zentrum zur Sauerstoffreduktion bildet. Die Untereinheit 2 beinhaltet ein reaktives Ubichinonmolekül, das in Kontakt mit Häm b steht. Dieses Polypeptid bindet auch Ubichinol als Substrat. Die Oxidation von Ubichinol gekoppelt mit der Reduktion von $\frac{1}{2}$ O₂ führt zur Translokation von vier Protonen (SCHULTZ & CHAN, 1998; MATSUSHITA et al., 1987). In dem Genom von *G. oxydans* kodieren die ORFs 0554, 0555, 0557 und 0558 für die vier Cytochrom *bo₃*: Ubichinol-Oxidase-Untereinheiten. Der ORF 0555 kodiert für die Untereinheit I. Die kalkulierte molekulare Masse des aus der Aminosäuresequenz resultierenden Proteins betrug 74 kDa. Die molekulare Masse des korrespondierenden Proteins aus *G. suboxydans* beträgt 76 kDa (MATSUSHITA et al., 1987). Für die Untereinheit II des Enzyms kodiert der ORF 0554. Die geschätzte molekulare Masse des aus dem Gen resultierenden Proteins betrug 35 kDa, was dem Wert für das korrespondierende Protein aus *G. suboxydans* entspricht (MATSUSHITA et al., 1987). Die ORFs 0557 und 0558 kodieren für die Untereinheiten III und IV der Chinol-Oxidase, die abgeleiteten molekularen Massen betragen 23 bzw. 12 kDa. Die aus *G. suboxydans* aufgereinigten Untereinheiten III und IV weisen eine molekulare Masse von 21 bzw. 15 kDa auf (MATSUSHITA et al., 1987). Über das Vorhandensein von Cytochrom *bd*-Oxidasen in *Gluconobacter*-Stämmen wurde bislang nicht berichtet. Für Spezies aus der Gattung *Acetobacter* konnte jedoch eine Cytochrom *bd*-Oxidase nachgewiesen werden (MATSUSHITA et al., 1994). In dem Genom von *G. oxydans* 621H kodieren die ORFs 1563 und 1564 für die Untereinheiten I und II des erwähnten Enzyms. Die aus den Aminosäuresequenzen resultierenden molekularen Massen betragen 53 bzw. 38 kDa.

Die Elektronen fließen vom Ubichinonpool auf die entsprechenden Endoxidasen. Der Ubichinonpool kann von der NADH-Dehydrogenase Typ II gespeist werden, wobei NADH aus cytoplasmatischen Reaktionen stammt. Es ist jedoch anzunehmen, dass die meisten Elektronen von den membrangebundenen Dehydrogenasen entweder über Cytochrome oder direkt über den Ubichinonpool in die Atmungskette eingebracht und anschließend auf die entsprechenden Endoxidasen übertragen werden. Für *G. suboxydans* wurde gezeigt, dass fast alle der primären Dehydrogenasen, die in Zucker- oder Alkohol-oxidierende Systeme involviert sind, in der Cytoplasmamembran vorliegen (MATSUSHITA et al., 1994). Dabei handelt es sich entweder um Chinoproteine, die Pyrrolochinolinchinon (PQQ) als Cofaktor aufweisen oder aber um Flavoproteine, welche FAD binden (MATSUSHITA et al., 1994). Als Beispiel für eine Flavoprotein-Dehydrogenase sei die 2-Ketogluconat-Dehydrogenase aus dem Genom von *G. oxydans* genannt. Kodiert wird dieses Protein von den ORFs 2631, 2632, 2633. Aus *G. melanogenus* wurde eine 2-Ketogluconat-Dehydrogenase aufgereinigt, die ebenfalls aus drei Untereinheiten besteht (61, 47, 25 kDa, SHINAGAWA et al., 1981). Die kalkulierten molekularen Massen der aus den ORFs 2631, 2632, 2633 resultierenden Proteinen betragen 47, 64 und 26 kDa, was im Einklang mit den ermittelten Größen des aufgereinigten Proteins aus *G. melanogenus* steht. Als Beispiele für Chinoproteine aus dem Genom von *G. oxydans* seien die Alkohol-Dehydrogenase und die membrangebundene Glucose-Dehydrogenase genannt. Die Alkohol-Dehydrogenase wird von den ORFs 2454 und 2455 kodiert, was ein Hinweis darauf sein könnte, dass das Protein aus zwei Untereinheiten besteht. Aus *G. suboxydans* wurde eine Alkohol-Dehydrogenase aufgereinigt, die aus drei Untereinheiten (80, 48, 14 kDa) zusammengesetzt ist (MATSUSHITA et al., 1994). Die kalkulierten Massen der aus der Aminosäuresequenz resultierenden Proteine betragen für ORF 2454 und 2455 51 und 83 kDa. Hier stellt sich die Frage, ob das Protein aus *G. oxydans* im Gegensatz zu dem aus *G. suboxydans* nur aus zwei Untereinheiten zusammengesetzt ist. Dieses lässt sich nur mit biochemischen Methoden nachweisen. An anderer Position des Genoms findet sich jedoch ein ORF (2124), der für eine 15 kDa-Untereinheit einer Alkohol-Dehydrogenase kodiert. Dieses Protein entspricht der 14 kDa Untereinheit des Enzyms aus *G. suboxydans*. Die PQQ-abhängige, membrangebundene Glucose-Dehydrogenase aus dem Genom von *G. oxydans* wird von ORF 1549 kodiert. Das korrespondierende Protein entspricht der in der Literatur beschriebenen Glucose-Dehydrogenase aus *G. suboxydans* (AMEYAMA et al., 1981).

Zusammenfassend kann man sagen, dass die Atmungskette von *G. oxydans* eine Vielzahl membrangebundener Dehydrogenasen aufweist, die ihre Elektronen direkt in den Ubichinonpool einspeisen.

Gene, die für ein wichtiges Enzym der Atmungskette, die NADH Oxidoreduktase oder Komplex I kodieren, konnten im Genom von *G. oxydans* nicht identifiziert werden. Lediglich eine NADH-Dehydrogenase von Typ II wurde in den Genomdaten gefunden. Hierbei handelt es sich jedoch um ein Protein, welches nicht an der Protonentranslokation und damit am Aufbau eines Membranpotentials beteiligt ist. Desweiteren konnte im Einklang mit biochemischen Untersuchungen zum Aufbau der Atmungskette keine Cytochrom c-Oxidase identifiziert werden (MATSUSHITA et al., 1994). Um so erstaunlicher ist es, dass die Genomdaten das Vorhandensein eines *bc₁*-Komplexes vorhersagen. Über die biologische Funktion dieses *bc₁*-Komplexes kann man ohne biochemische Untersuchungen jedoch nur spekulieren. Auch ein Komplex II der Atmungskette konnte in dem Genom von *G. oxydans* nicht identifiziert werden.

Die auf den Genomdaten basierende Rekonstruktion der Atmungskette spiegelt die für *G. oxydans* charakteristischen Eigenschaften wider: Dieser Organismus zeichnet sich vor allem durch hohe Oxidationsraten einhergehend mit geringen Wachstumserträgen aus. Der Prozess der unvollständigen Oxidation, bei dem energiereiche Produkte ausgeschieden werden und die spärlichen Möglichkeiten zum Aufbau eines Membranpotentials lassen keine hohen Wachstumserträge zu. Bei den membrangebundenen Dehydrogenasen, durch deren Dehydrogenierungsreaktionen die Elektronen direkt in die Atmungskette eingespeist werden, handelt es sich vor allem um PQQ-abhängige Enzyme. Proteine, welche PQQ als Cofaktor aufweisen, treten vor allem bei Organismen auf, deren Atmungskette eher schlicht aufgebaut ist, wie es auch bei *G. oxydans* der Fall ist (MATSUSHITA et al., 2002). Einerseits stellt die unvollständige Oxidation von Zuckern, Polyolen und Alkoholen für diesen Organismus eine Verschwendung von Substraten dar, andererseits macht ihn aber auch gerade diese Eigenschaft zu einem interessanten Organismus für die Biotechnologie.

3.6.5.2 Die ATP-Synthase

Der terminale Schritt der oxidativen Phosphorylierung besteht in der Synthese von ATP. Wie bei fast allen Pro- und auch Eukaryoten erfolgt diese Reaktion in *G. oxydans* durch eine F_1F_0 -ATP-Synthase. Die Gene, die für den Multienzym-Komplex der ATP-Synthase kodieren, liegen in der Regel in einem Operon vor und umfassen insgesamt neun Gene, die für die Untereinheiten α , β , γ , δ , ϵ , a, b, c und i kodieren (SCHNEIDER & ALTENDORF, 1985; WALKER et al. 1982).

In dem Genom von *G. oxydans* konnten drei Operons identifiziert werden, die möglicherweise für ATP-Synthasen vom F_1F_0 -Typ kodieren. Das erste Operon ist in Abb. 39 dargestellt und umfasst erwartungsgemäß neun ORFs. Allerdings konnten bei der Annotation nur acht ORFs eine Funktion zugewiesen werden. Bei dem dargestellten Operon konnten ORFs, die für die β - (ORF 3463), ϵ - (ORF 0839), i- (ORF 0840), a- (ORF 3464), c- (ORF 0844), b- (ORF 0845) und α -Untereinheit (ORF 0847) kodieren identifiziert werden. Ein ORF, der für die δ -Kette kodiert wurde nicht gefunden. Innerhalb des Operons befindet sich zwar noch ein zusätzlicher ORF (0840), dieser zeigte jedoch keine Ähnlichkeiten zu bekannten Proteinen.

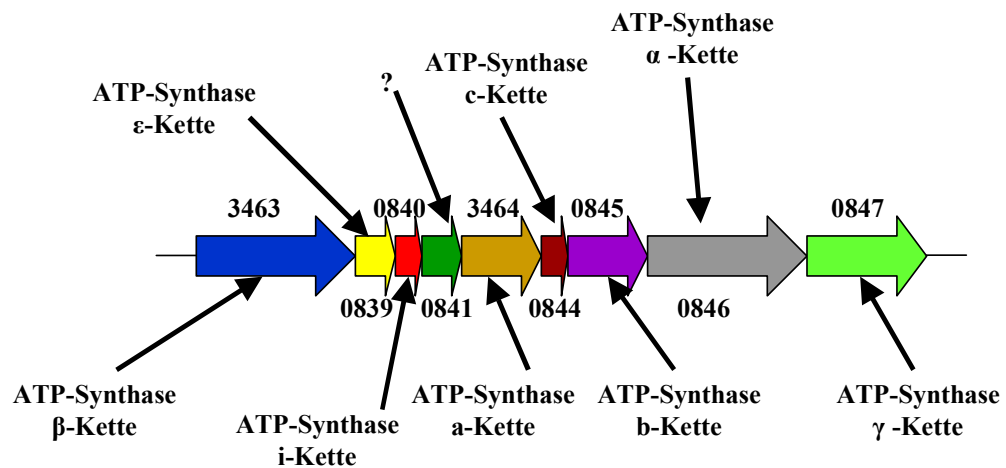


Abb. 39: Operon der ATP-Synthase

Das Operon umfasst neun Gene. Bei der Annotation konnte dem ORF 0841 jedoch keine Funktion zugeordnet werden.

An anderer Position des Genoms konnten zwei weitere Operons identifiziert werden, die ebenfalls für eine F_1F_o -Typ ATP-Synthase kodieren (Abb. 40). Das erste Operon umfasst die Gene für die hydrophobe b- (ORF 2504, 2506), c- (ORF 2508) und a-Kette (ORF 2510) der ATP-Synthase. Das zweite Operon enthält die Gene der hydrophilen δ - (ORF 2717), α - (ORF 3520), γ - (ORF 2720), β - (ORF 2721) und ϵ - (ORF 2723) Untereinheiten. Diese beiden Operons umfassen insgesamt neun ORFs, jedoch konnte kein ORF identifiziert werden, der für die i-Kette der ATP-Synthase kodiert. Eine Anordnung der Gene in zwei Operons, die für die ATP-Synthase kodieren, wurde bislang in der Literatur noch nicht beschrieben. Ein derartiges Resultat gab Anlass dazu, die Rohdaten in den entsprechenden Bereichen erneut zu überprüfen. Die Überprüfung zeigte jedoch keinerlei Hinweise auf Fehler in den Sequenz-Rohdaten. Eine eindeutige Identifizierung der ATP-Synthase kann daher erst durch biochemische Experimente erfolgen. Dennoch ist es wahrscheinlich, dass diese Gencluster die notwendige Information zur Bildung der ATP-Synthase enthalten.

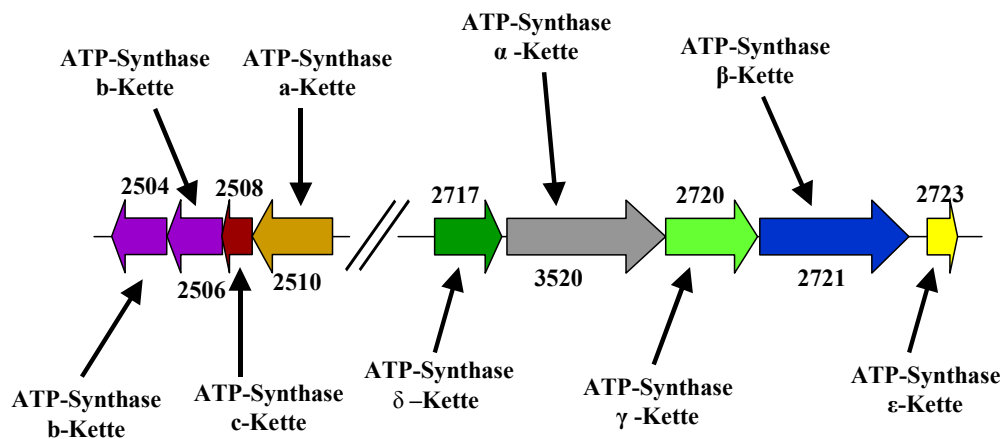


Abb. 40: Operons, die für die ATP-Synthase kodieren.

Die ORFs, die für die ATP-Synthase kodieren befinden sich auf zwei verschiedenen Operons.

3.6.6 Dehydrogenasen

Charakteristisch für *Gluconobacter*-Stämme ist die unvollständige Oxidation von vielen Zuckern, Alkoholen und Polyolen. Die Oxidation der Substrate erfolgt dabei stereo- und regioselektiv. Die meisten dieser Dehydrogenierungsreaktionen werden von

membrangebundenen Dehydrogenasen katalysiert. Dabei ragt das aktive Zentrum der Enzyme in den periplasmatischen Raum, so dass ein Transport des entsprechenden Substrats in die Zelle nicht notwendig ist. Die unvollständig oxidierten Produkte werden oft direkt ins Medium ausgeschieden. Diese Eigenschaften machen *G. oxydans* vor allem für die Biotechnologie interessant, da zum einen eine aufwendige und auch kostspielige Schutzgruppenchemie für Oxidationsreaktionen an Zuckern nicht notwendig ist, zum anderen ist die Aufreinigung der gewünschten Produkte einfach. *G. oxydans* wird z. B. für die Synthese von Miglitol, einem Antidiabetis-Medikament eingesetzt. Die Analyse des Genoms zeigte, dass neben bekannten Dehydrogenasen auch viele noch nicht charakterisierte Dehydrogenase-Gene im Chromosom vorliegen.

In der Literatur wird immer wieder auch die Produktion von Di-Ketogluconaten durch *Gluconobacter* beschrieben (SHINAGAWA et al., 1981). Die Analyse des Genoms des Stammes 621H zeigte die Anwesenheit von Genen zur Produktion einer membrangebundenen und einer cytoplasmatischen Gluconat-Dehydrogenase (ORF 2631, 2632, 2633 und ORF 0864). Demgegenüber wurden jedoch keine Gene gefunden, die für Ketogluconat-Dehydrogenasen kodieren, so dass dieser Stamm offensichtlich nicht in der Lage ist, Diketogluconat herzustellen. Die Produktion von Ketogluconaten durch *Gluconobacter*-Stämme wird nachfolgend in der Diskussion eingehend erläutert.

Erwähnt sei an dieser Stelle noch die Cytochrom c abhängige Sorbitol-Dehydrogenase aus *G. oxydans*, von der bereits eine Sequenz vorlag (Accession No. ABO039821.1). Diesbezüglich ergab die Analyse des Genoms, dass dieses Enzym in dem hier sequenzierten Stamm 621H vermutlich nicht funktionell ist. Die ORFs 0755 und 0756 zeigen beide Ähnlichkeiten zu den bereits sequenzierten Untereinheiten des Proteins, jedoch fehlen bei ORF 0755 100 Aminosäuren. Diese werden laut der Datenanalyse von ORF 0756 kodiert. Eine derartige Konstellation deutet stark auf einen Frameshift hin, also auf einen Fehler in den Rohdaten. Eine erneute Überprüfung zeigte jedoch keinerlei Anhaltspunkte dafür, so dass zunächst davon ausgegangen werden muss, dass dieses Gen durch eine natürliche Mutation nicht mehr funktionell ist.

In Tab. 10 sind alle nicht-charakterisierten Dehydrogenasen und Oxidoreduktasen aus dem Genom von *G. oxydans* aufgelistet. In der Tabelle ist jeweils die entsprechende ORF-Nr. angegeben, sowie die Annotation. Auf der Basis der aus den ORFs

resultierenden Aminosäuresequenz wurden die identifizierten Dehydrogenasen entsprechenden Familien bzw. Klassen zugeordnet. Diese sind ebenfalls in Tab. 10 aufgelistet. Insgesamt wurden im Genom Gene für 73 Dehydrogenasen/Oxidoreduktasen identifiziert, die bislang nicht charakterisiert sind. Darunter sind 16 ORFs, die für Proteine kodieren, die der Familie der kurzkettigen Dehydrogenasen/Reduktasen zugeordnet wurden. Proteine dieser Familie zeigen ein breites Substratspektrum. So ist zu erwarten, dass die entsprechenden Proteine Aktivitäten für die Oxidation von Alkoholen, aromatischen Verbindungen und verschiedenen Zuckern besitzen. Ebenfalls in der Tab. 10 ist ein Protein aus der Familie der kurzkettigen Dehydrogenasen/Reduktasen aufgeführt, das kürzlich aus *G. oxydans* 621H gereinigt und charakterisiert wurde. Es handelt sich dabei um eine cytoplasmatische Xylitol-Dehydrogenase (in Tab. 10 grau unterlegt). In dem Genom von *G. oxydans* wird diese Enzym von ORF 2241 kodiert. Die Xylitol-Dehydrogenase katalysiert die Umsetzung von D-Xylulose zu Xylitol (SUGIYAMA et al., 2003a). Xylitol wird unter anderem als alternativer Süßstoff eingesetzt. Weiterhin konnten sechs ORFs identifiziert werden, die für noch nicht charakterisierte Dehydrogenase kodieren, die der Familie der Aldo/Keto-Reduktasen zugeordnet wurden (ORF 1579, 2001, 3006, 0065, 0194, 0212). Vier der identifizierten ORFs kodieren für Enzyme aus der Familie der Aldehyd-Dehydrogenasen (ORF 1107, 1838, 0325, 0774). Weiterhin wurden 14 ORFs gefunden, die vermutlich für FAD-abhängige Enzyme kodieren. Vier ORFs wurden identifiziert, welche für Proteine kodieren, die in die Familie der Chinoproteine einzuordnen sind. Flavo- und Chinoproteine sind in der Regel in der Membran lokalisiert und übertragen in der Atmungskette Elektronen aus den jeweiligen Dehydrogenierungsreaktionen auf den Ubichinonpool. Es ist zu erwarten, dass eine biochemische Charakterisierung der hier auf der Basis der genomischen Sequenz identifizierten Dehydrogenasen neue Wege für die Biotechnologie eröffnet. Die Ermittlung des Substratspektrums der einzelnen Dehydrogenasen kann z. B. dazu führen, dass neue biotechnologische Verfahren entwickelt und bestehende verbessert werden können.

Darüber hinaus sind in der Tabelle zwei weitere ORFs grau hervorgehoben. Diese ORFs kodieren für Enzyme, die kürzlich identifiziert und charakterisiert wurden und eine wichtige Rolle bei der Zucker-Oxidation spielen. Einer dieser ORFs (2223) kodiert für eine NADP-abhängige Sorbose-Reduktase und ist in die Familie der Mannitol-

Dehydrogenasen einzuordnen. Die Sorbose-Reduktase katalysiert die Reaktion zur Umsetzung von L-Sorbose zu D-Sorbitol (SHINJOH et al., 2002). Durch die cytoplasmatische Sorbitol-Dehydrogenase (ORF 0010, in Tab. 10 grau unterlegt) wird D-Sorbitol zu Fructose umgesetzt und in den Pentosephosphat-Weg eingeschleust (SHIBATA et al., 2000). Sorbose ist ein wichtiges Intermediat bei der Produktion von Vitamin C und ist ein Produkt, welches bei der Umsetzung von D-Sorbitol durch die membrangebundene Sorbitol-Dehydrogenase entsteht.

Tab. 10: Aufstellung von Genen im Chromosom von *G. oxydans* 621H, die für bislang nicht-charakterisierte Dehydrogenasen und Oxidoreduktasen kodieren

ORF	Annotation	Familie oder Domäne
1784	Putative transmembrane Oxidoreduktase	3-beta Hydroxysteroid Dehydrogenase/Isomerase Familie IPR002225
1107	Putative Aldehyde-Dehydrogenase	Aldehyd-Dehydrogenase Familie IPR002086
1838	NAD-abhängige Aldehyd-Dehydrogenase	Aldehyd-Dehydrogenase Familie IPR002086
1579	Putative Oxidoreduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395
2001	wahrscheinliche 2,5-Diketo-D-Gluconat-Reduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395
3006	Putative Oxidoreduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395
618	Nicht-charakterisierte PQQ-enthaltende Dehydrogenase 2	Bakterielle Chinoprotein-Dehydrogenase Familie IPR002372
1859	Nicht-charakterisierte PQQ-enthaltende Dehydrogenase 4	Bakterielle Chinoprotein-Dehydrogenase Familie IPR002372
1609	Putative Oxidoreduktase (EC 1.1.1.-)	Delta 1-Pyrrolin-5-Carboxylat-Reduktase Familie IPR000304
2245	Elektronen-Transfer Flavoprotein-Ubichinon Oxidoreduktase (EC 1.5.5.1)	Elektronen-Transfer Flavoprotein-Ubichinon Oxidoreduktase Familie IPR001308
2247	Elektronen-Transfer Flavoprotein beta-Untereinheit	Elektronen-Transfer Flavoprotein-Ubichinon Oxidoreduktase Familie IPR001308
2248	Elektronen-Transfer Flavoprotein alpha-Untereinheit	Elektronen-Transfer Flavoprotein-Ubichinon Oxidoreduktase Familie IPR001308
1988	wahrscheinliche D-Aminosäure-Dehydrogenase (EC 1.4.99.1)	FAD-abhängige Oxidoreduktase Familie 927
1269	Putative Oxidoreduktase (EC 1.1.1.-)	FAD-abhängige Oxidase Familie IPR004113
970	Putative Oxidoreduktase (EC 1.1.1.-)	Flavin enthaltende Amin Oxidoreduktase Familie IPR002937
1999	Putative Oxidoreduktase (EC 1.1.1.-)	Flavin-Reduktase ähnliche Domäne Familie IPR002563
0771	wahrscheinliche NADPH-spezifische Chinon-Oxidoreduktase („drug modulator“)	Flavodoxin 2 Familie IPR003690
2458	Putative Oxidoreduktase (EC 1.1.1.-)	GMC Oxidoreduktase Familie PF05199
1718	NAD(FAD)-abhängige Dehydrogenase	HI0933-ähnliche Protein-Familie (Dinukleotid-Binde- Motiv Proteine einschl. DHs) IPR004792)
1918	Aldehyde-Dehydrogenase ähnliches Protein	HTH_3 IPR001387
2223	wahrscheinliche NADPH-abhängige L-Sorbose-Reduktase	Mannitol-Dehydrogenase Familie IPR000669
0090	Putative Oxidoreduktase (EC 1.1.1.-)	Molybdopterin-Dehydrogenase, FAD-Binde-Familie IPR002346
2143	Putative Oxidoreduktase (NADH Dehydrogenase/NAD(P)H	Nitroreduktase Familie IPR000415

Nitroreduktase)				
2206	Putative Oxidoreduktase (EC 1.1.1.-)	Nitroreduktase Familie IPR000415		
1138	Putative Oxidoreduktase (NAD-Binde-Motiv)	Oxidoreduktase Familie, NAD-bindend Rossmann fold IPR000683		
2800	Putative Oxidoreduktase (EC 1.1.1.-)	Pyridin Nucleotide-Disulfid	Oxidoreduktase IPR001327	
0828	Putative Dehydrogenase (NAD ⁺ -Binde-Motiv)	Saccharopin-Dehydrogenase Familie IPR005097		
0058	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
0134	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
0542	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
0690	Putative Oxidoreduktase (NAD(P)-abhängig)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
0855	Putative Polyol-Dehydrogenase	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
1110	kurzkettige Alkohol-Dehydrogenase	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
1167	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
1787	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige IPR002198	Dehydrogenase/Reduktase	Familie
1603	Propanol- Alkohol Dehydrogenase (EC 1.1.1.1)	Zink-bindende Dehydrogenase Familie IPR002085		
1604	Putative Alkohol Dehydrogenase	Zink-bindende Dehydrogenase Familie IPR002085		
1742	wahrscheinliche 2-Hydroxysäure-Dehydrogenase (EC 1.1.1.-)	2-Hydroxysäure-Dehydrogenase Familie IPR002162		
0959	Putative Oxidoreduktase (EC 1.1.1.-)	3-beta Hydroxysteroid	Dehydrogenase/Isomerase Familie IPR002225	
0325	Aldehyd-Dehydrogenase (geringe Ähnlichkeit zu L-Sorboson-Dehydrogenase)	Aldehyd-Dehydrogenase Familie IPR002086		
0774	Putative Aldehyd-Dehydrogenase	Aldehyd-Dehydrogenase Familie IPR002086		
0065	Putative Oxidoreduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395		
0194	Putative Oxidoreduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395		
0212	Putative Oxidoreduktase (EC 1.1.1.-)	Aldo/Keto-Reduktase Familie IPR001395		
0027	Nicht-charakterisierte PQQ-abhängige Dehydrogenase 3	Bakterielle IPR002372	Chinoprotein-Dehydrogenase	Familie
0494	Nicht-charakterisierte PQQ-enthaltende Dehydrogenase 1	Bakterielle IPR002372	Chinoprotein-Dehydrogenase	Familie
0427	Putative Oxidoreduktase (EC 1.1.1.-)	Delta 1-Pyrroline-5-Carboxylat	Reduktase	Familie IPR000304
2657	D-Lactat-Dehydrogenase (NAD-abhängig) (EC 1.1.1.28)	FAD Binde-Domäne Familie IPR001575		
0731	Putative D-Lactat-Dehydrogenase (Cytochrom c-abhängig) (EC 1.1.1.28)	FAD Oxidase Familie IPR004113		
2568	Lactat-Dehydrogenase (Cytochrom c-abhängig)	FAD Oxidase Familie IPR004113		
0228	Putative Oxidoreduktase (EC 1.1.1.-)	FAD-abhängige Oxidoreduktase	Pyridin Nucleotid-Disulfid	Familie IPR001327
2565	membrangebundene Dehydrogenase	FMN-abhängige Dehydrogenase Familie IPR000262		
0330	Putative Oxidoreduktase (EC 1.1.1.-)	HI0933-ähnliche Protein Familie (Dinukleotid-Binde Motiv f. Proteine einschl. DHs) IPR004792)		
0772	NAD ⁺ -abhängige Eisen-enthaltende Alkohol-Dehydrogenase	Eisen-enthaltende Alkohol-Dehydrogenase Familie IPR001670		
0010	Putative NADP-D-Sorbitol-Dehydrogenase	Mannitol-Dehydrogenase Familie IPR000669		
1841	wahrscheinliche NAD(P)H-abhängige 2-Cyclohexen-1 -Reduktase	NADH:Flavin-Oxidoreduktase IPR001155		

2676	Putative Oxidoreduktase (EC 1.1.1.-)	NADP Oxidoreduktase Coenzyme F420-abhängig	IPR004455
1485	NAD(P)H-Flavin Oxidoreduktase	Nitroreduktase Familie	IPR000415
1813	Putative Oxidoreduktase (EC 1.1.1.-)	Nitroreduktase Familie	IPR000415
0956	Putative Cyclohexadienyl Dehydrogenase	Prephenat Dehydrogenase Familie	IPR003099
2286	Putative Oxidoreduktase (EC 1.1.1.-)	Pyridin Nukleotid-Disulphid Oxidoreduktase	IPR001327
1869	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
1952	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2003	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2080	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2241	wahrscheinliche Polyol-Dehydrogenase	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2337	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2422	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
2539	Putative Oxidoreduktase (EC 1.1.1.-)	kurzkettige Dehydrogenase/Reduktase	Familie IPR002198
1810	Putative Oxidoreduktase (EC 1.1.1.-)	Zink-bindende Dehydrogenase Familie	IPR002085
2907	Putative transmembrane Oxidoreduktase	Zink-bindende Dehydrogenase Familie	IPR002085

3.6.6.1 Pyrrolochinolin Chinon (PQQ)

Viele der membrangebundenen Dehydrogenasen in *G. oxydans* sind PQQ abhängig. Dieser Redox-Faktor wurde erstmals 1979 in Bakterien entdeckt (SALISBURY et al., 1979). Enzyme, welche PQQ als prosthetische Gruppe enthalten werden als Chinoproteine bezeichnet. Besonders charakteristisch sind Chinoproteine für Organismen mit einer zwar ausgeprägten, aber doch eher schlicht gestalteten Atmungskette (MATSUSHITA et al., 2001), wie es auch bei *G. oxydans* der Fall ist. Beispiele für bereits bekannte Chinoproteine sind die membrangebundene Alkohol-Dehydrogenase (ORF 2454, 2455) oder die Sorbitol-Dehydrogenase (kodiert von *slmA* und *sldB*; ORF 2230, 2231). Die ORFs 0027, 0494, 0618, 1859 kodieren für bislang nicht charakterisierte Chinoproteine. Die Biosynthese des Redox-Faktors PQQ wird von einem Cluster, das fünf Gene umfasst, kodiert (FELDER et al., 2000). Dieses Cluster findet man in dem Genom des Stammes 621H in einfacher Kopie vor (ORFs 5001 (*ppqA*), 2373 (*ppqB*), 2372 (*ppqC*), 3312 (*ppqD*) und 2370 (*ppqE*)). Dieses Resultat steht im Einklang mit den Ergebnissen von FELDER et al. (2000), die diese Anordnung im Stamm IFO 3293 identifiziert hatten. Das Genprodukt von *ppqA* ist mit nur 26

Aminosäuren ein sehr kleines Protein. Die Genprodukte von *ppqB-E* weisen Größen von 304, 239, 96 und 358 Aminosäuren auf.

Es deutet sich jedoch an, dass in dem in dieser Arbeit sequenzierten Stamm nur die ORFs, welche für *ppqB-E* kodieren, in einem Operon organisiert sind. Bei diesen ORFs überlappen jeweils Start und Stopcodon. Stromaufwärts von ORF 2373 (*ppqB*) findet man eine potentielle σ^{70} -abhängige Promotorstruktur (Abb 41). Der ORF, welcher für *ppqA* kodiert, befindet sich 119 bp stromaufwärts von ORF 2373 und weist auch eine potentielle Struktur σ^{70} -abhängiger Promotoren auf (Abb. 41).

	-35-Region	-10-Region	
<i>ppqBCDE</i>	<u>TTC</u> ATG	<u>TG</u> AAAA	3 bp*
<i>ppqA</i>	GGA <u>ACA</u>	<u>AAAA</u> AT	
Konsensus aus <i>E. coli</i>	TTGACA	TATAAT	3 bp*

Abb. 41: Mögliche σ^{70} -abhängige Promotorstrukturen der ORFs, die für *ppqA* und *ppqBCDE* kodieren.
*: Startcodon. Die mit der Konsensussequenz aus *E. coli* (HAWLEY & Mc CLURE, 1983) übereinstimmenden Basen sind unterstrichen.

3.7 Abschließende Bemerkungen

Die Analyse des Genoms zeigte, dass *G. oxydans* zwar entscheidende Stoffwechselwege nicht beschreiten kann, da wichtige Schlüsselenzyme fehlen, aber dennoch ist dieser Organismus in der Lage, die für den intrazellulären Stoffwechsel notwendigen Produkte selbst zu synthetisieren. Dargestellt wurde dieses an ausgewählten Beispielen, eine vollständige Analyse des Stoffwechsels würde den Rahmen dieser Arbeit sprengen. Die Synthese der Aminosäuren ist recht eng mit den zentralen Stoffwechselwegen verknüpft, das Kohlenstoffgerüst stammt meistens von Pyruvat oder Zwischenprodukten des Zitratzyklus. Angemerkt sei noch, dass vor allem die Rekonstruktion der Biosynthese der Aminosäuren auf den Annotationsdaten beruhte. In der Literatur ist bislang nichts über biochemische Untersuchungen bezüglich der

Synthese und Regulation der Aminosäuren in *G. oxydans* bekannt. Die Analyse der Daten hat auch deutlich gemacht, dass aufgrund des Fehlens zentraler Stoffwechsellzyme die Energiegewinnung zwangsläufig in der Membran stattfinden muss. Dieses führt dann zu der unvollständigen Oxidation verschiedenster Substrate, wobei energiereiche Produkte ausgeschieden werden. Hintergrund für die in der Biotechnologie so wertvollen unvollständigen Oxidationsreaktionen sind also schlicht Mängel in der Enzym-Ausstattung dieses Organismus, die dazu führen, dass in die Zelle aufgenommene Substrate nur zum Aufbau intrazellulärer Intermediate dienen, nicht aber zur Energiegewinnung. Die Transkription und Translation der entsprechenden Gene, die für die membrangebundenen Enzyme kodieren, erfolgt jedoch im Innern der Zelle. Dies bedeutet, dass ein Protein, welches im periplasmatischen Raum oder in der äußeren Membran lokalisiert ist, erst die cytoplasmatische Membran passieren muss, was als Translokation bezeichnet wird. Hierfür ist in dem Genom von *G. oxydans* ein vollständiges Sec-System vorhanden.

4 DISKUSSION

In den vorangehenden Kapiteln wurden die Sequenzdaten des Genoms von *G. oxydans* vorgestellt und die wichtigsten Stoffwechselwege rekonstruiert. In der nun folgenden Diskussion sollen charakteristische Eigenschaften des Genoms analysiert werden. Der Schwerpunkt liegt dabei in der Betrachtung des biotechnologischen Potentials dieses Organismus in Bezug auf die erhaltenen Sequenzdaten.

4.1 Mobile Elemente

Eine besondere Schwierigkeit beim Schließen des Genoms stellten die vielen repetitiven Elemente in Form von Transposons und Insertionselementen dar. Dabei handelt es sich um mobile Elemente, die in der Regel im Lauf der Evolution durch Rekombinationsereignisse akquiriert wurden.

Das Vorhandensein der vielen mobilen Elemente, legt einerseits die Vermutung nahe, dass viele der darauf liegenden Gene ursprünglich nicht aus *G. oxydans* stammen, sondern, dass es sich um Gene aus anderen Bakterien handelt, die im Folgenden als „Alien-Gene“ bezeichnet werden. Andererseits kann man schließen, dass Insertionselemente eine genetische Instabilität bedingen. „Alien-Gene“ sind meistens ORFs unbekannter Funktion, Transposasen, Gene von Prophagen und Restriktions/Modifikations-Systeme. Da „Alien-Gene“ oft in Clustern auftreten kann geschlossen werden, dass mehrere Gene in das Transferereignis involviert waren (MRAZEK & KARLIN, 1999).

FRANK et al. (2002) postulierten, dass es eine Korrelation zwischen der Größe eines Genoms, der Anzahl an repetitiven Elementen und dem Lebensraum des Organismus gibt. Aufgrund der wechselnden Umweltbedingungen wird ein hoher Gehalt an repetitiven Sequenzen vor allem bei frei-lebenden und fakultativ intrazellulären Parasiten erwartet (FRANK et al., 2002). Mit anderen Worten bedeutet dies, dass intrazellulär-lebende Mikroorganismen aufgrund der relativ konstanten Umweltbedingungen weniger Insertionselemente enthalten sollen, woraus man einen höheren Grad an genetischer Stabilität erwarten könnte.

Auf der Basis statistischer Untersuchungen postulierten FRANK et al., 2002, dass die Korrelation zwischen der Genom-Größe und dem Gehalt an repetitiven Sequenzen dann „perfekt“ ist, wenn in einem mikrobiellen Genom die Dichte der repetitiven Sequenzen bei 1,7 % liegt. Im Fall von *G. oxydans* 621H lag dieser Wert bezogen auf das Chromosom bei 2,2 %. Die Theorie von FRANK et al. (2000) besagt, dass frei-lebende Organismen einen höheren Gehalt an repetitiven Sequenzen aufweisen, als es bei intrazellulären Organismen der Fall ist. In der Natur findet man *Gluconobacter*-Spezies vor allem auf Früchten und in alkoholischen Getränken (ASAI, 1968). Das bedeutet, dass ein durchschnittlich höherer Gehalt an repetitiven Sequenzen, wie es in *G. oxydans* der Fall ist, im Einklang mit den Ergebnissen von FRANK et al (2000) steht. Darüber hinaus wird angenommen, dass der G+C-Gehalt in einem Organismus nur mit geringen Abweichungen um einen Mittelwert schwankt. Größere Abweichungen vom durchschnittlichen G+C-Gehalt werden in der Regel dadurch erklärt, dass Transpositionereignisse stattgefunden haben (OCHMAN & SANTOS, 2003).

Im Gegensatz dazu wurde postuliert, dass der G+C-Gehalt bei aeroben Prokaryoten höher ist als bei anaeroben (NAYA et al., 2002). Wenn man davon ausgeht, dass ein genetischer Austausch eher in Mikroorganismen stattfindet, die in einem ähnlichen Habitat leben, würde dies einer signifikanten Abweichung vom mittleren G+C-Gehalt bei Transpositionereignissen widersprechen. Dann müsste man erwarten, dass sich die Aufnahme fremder DNA weniger deutlich in Abweichungen des G+C-Gehalts widerspiegelt.

In *G. suboxydans* wurde beobachtet, dass ein Insertionselement (IS12528) mit der Inaktivierung der membrangebundenen Alkohol-Dehydrogenase assoziiert war. Von dem Insertionselement IS12528 und anderen IS-Elementen wird angenommen, dass diese zu einer genetischen Instabilität bei Essigsäurebakterien beitragen (KONDO & HORINOUCI, 1997b). Für *A. pasteurianus* wurde eine ähnliche Beobachtung gemacht. In diesem Organismus wurde eine inaktive Alkohol-Dehydrogenase in Zusammenhang mit dem Insertionselement IS1452 gebracht (KONDO & HORINOUCI, 1997c). Die Autoren zogen die Schlussfolgerung, dass multiple Insertionssequenzen eine maßgebliche Rolle bei spontanen Mutationen in

Essigsäurebakterien spielen. Diese führen dann zu dem Verlust verschiedener physiologischer Eigenschaften.

Im Göttinger Institut für Mikrobiologie und Genetik wurde ein Software-Paket entwickelt (SIGI: Score-based identification of genomic islands, MERKL, 2004) mit dem es möglich ist, ein Genom auf das Vorhandensein von „Alien-Genen“ zu untersuchen. Außerdem sollte mit Hilfe dieser Software eine Aussage über möglicherweise hoch exprimierte Gene getroffen werden. Der dazugehörige Algorithmus basiert darauf, dass die „Codon-usage“ bei akquirierten und hoch exprimierten Genen unterschiedlich ist (MERKL, 2004; Merkl, 2003). Mit dieser Software wurde auf das Chromosom von *G. oxydans* 621H untersucht, das Resultat ist in Abb. 42 gezeigt.

Auffallend war, dass ORFs, welche von dem Programm als „Alien-Gen“ ermittelt wurden, meistens in einen Cluster auftraten. Dabei handelte es sich in der Regel um Cluster, welche für hypothetische Proteine kodieren, oft gefolgt von einer Transposase. Als Beispiele seien hier die in einem Cluster organisierten ORFs 1823, 1824, 1825, 1826, 1827, und 3499 genannt. Die ORFs 1823 und 1824 wurden als Integrations-/Rekombinations-/Inversions-Protein annotiert; die ORFs 1825, 1826, und 1827 kodieren für hypothetische Proteine. Der ORF 3499 kodiert für eine Transposase, die allerdings von dem Computerprogramm nicht als „Alien-Gen“ eingestuft wurde. Ein weiteres Beispiel sind die ebenfalls in einem Cluster organisierten ORFs 3111, 3057, 3058 und 2725. Der ORF 3111 kodiert für ein konserviertes, hypothetisches Protein, die ORFs 3057 und 3058 kodieren für hypothetische Proteine. ORF 2725 wurde als Transposase annotiert, die von dem Computerprogramm als „Alien-Gen“ erkannt wurde.

Eine Ausnahme von der oben genannten Regel stellten die Glutamat-Synthetase und die NADP-abhängige Gluconat-Dehydrogenase dar. Diese ORFs wurden von dem Programm auch als „Alien-Gen“ erkannt. Als hoch exprimierte ORFs schlug die Software mit wenigen Ausnahmen ribosomale Proteine vor. Eine der Ausnahmen war die Chinoloxidase von Typ *bo*₃. *In vivo* ist vor allem zu erwarten, dass unter unterschiedlichen Wachstumsbedingungen verschiedene Gene hoch exprimiert oder aber auch reprimiert werden. Derartige Beobachtungen lassen sich jedoch nur mit geeigneten Experimenten, z. B. DNA-Microarray-Analysen nachweisen.

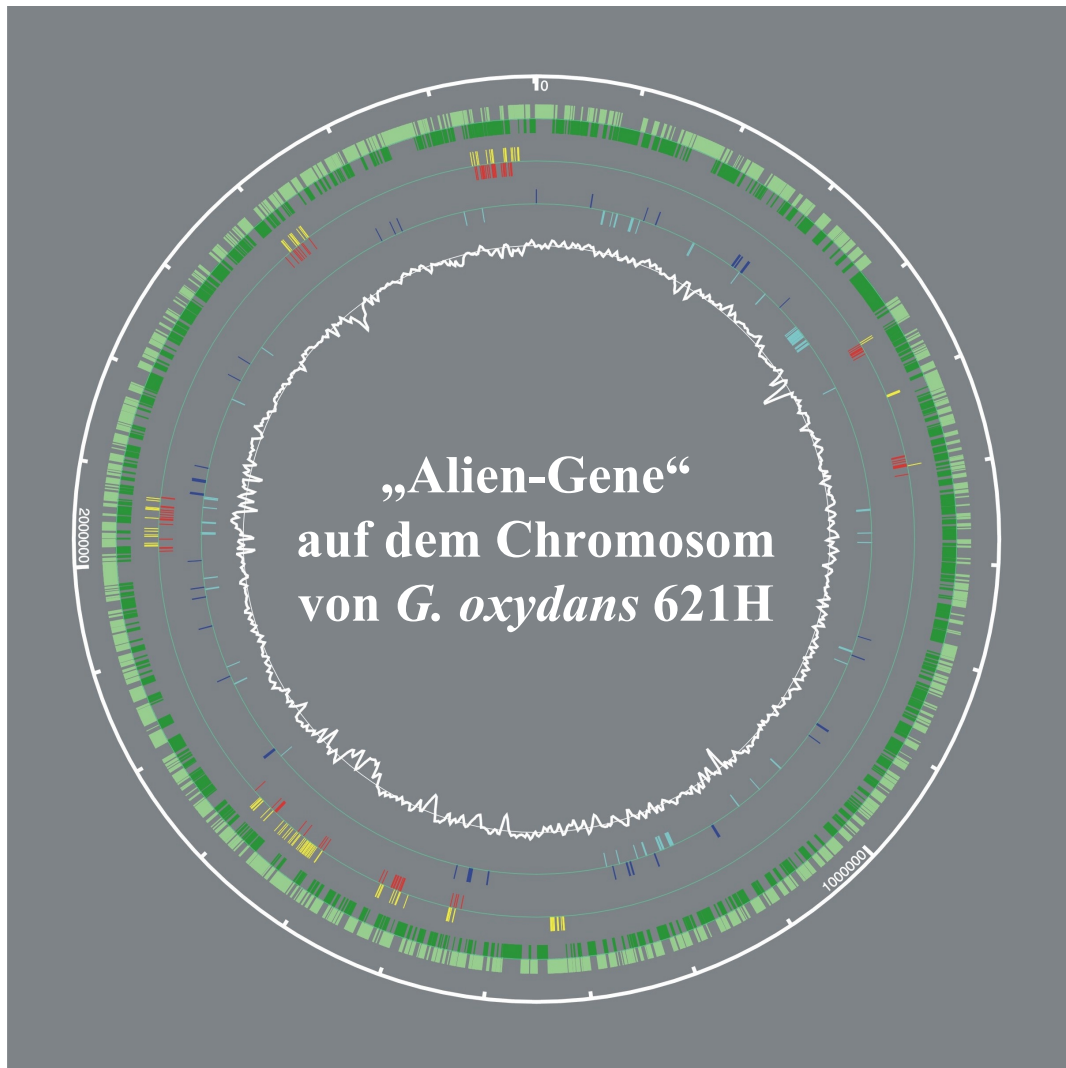


Abb. 42: Darstellung mutmaßlicher „Alien-Gene“ sowie hoch exprimierter Gene auf dem Chromosom von *G. oxydans* 621H.

Der äußere Rand stellt eine Skala dar; der Nullpunkt entspricht der Position des ermittelten Replikationsursprungs. In hell- und dunkelgrün sind sämtliche ORFs in Abhängigkeit der Lokalisierung auf dem jeweiligen DNA-Strang dargestellt. In gelb und rot sind diejenigen ORFs gezeigt, welche von der Software SIGI als „Alien“- oder potentielle „Alien-Gene“ identifiziert wurden. In hell- und dunkelblau sind diejenigen ORFs gezeigt, die als hoch oder potentiell hoch exprimierte ORFs eingestuft wurden. Die innere weiße Linie zeigt die Abweichung vom mittleren G+C-Gehalt.

4.2 Betrachtungen zur Energiekonservierung

Eine charakteristische Eigenschaft von *G. oxydans* sind die geringen Wachstumserträge, also die geringe Produktion von Biomasse. Dieses ist unter anderem eine der Eigenschaften, die *Gluconobacter* für biotechnologische Anwendungen interessant macht. Wie ist dieser Befund vor dem Hintergrund des entschlüsselten Genoms zu erklären? BAUCHHOP und ELSDEN postulierten schon 1960, dass die Energiekosten für die Biomasse-Produktion bei verschiedenen bakteriellen Spezies ähnlich sind. Obwohl genaue experimentelle Daten bis heute fehlen, wird doch allgemein angenommen, dass die Energiekosten für alle Organismen annähernd gleich sind. Etwas vereinfacht kann man auch sagen, je weniger effizient die Energiekonservierung, desto geringer sind zwangsläufig auch die Wachstumserträge pro Mol Substrat.

Die Grundvoraussetzung für die Energiekonservierung ist der Aufbau eines elektrochemischen Membranpotentials. Benötigt werden dafür Enzymsysteme, welche in der Lage sind, ein solches Membranpotential aufzubauen. Die Genomanalyse von *G. oxydans* 621H ergab, dass die Möglichkeiten zur Protonentranslokation für diesen Organismus gering sind. Die Komponenten zum Aufbau einer Proton-motorischen Kraft in vielen Pro- und Eukaryoten sind der Komplex I der Atmungskette, die Cytochrom c-Reduktase sowie Cytochrom c-Oxidase und einige Chinol-Oxidase. Die Analyse des Genoms von *G. oxydans* hat gezeigt, dass der Komplex I der Atmungskette nicht vorhanden ist. Diese NADH-Chinon:Oxidoreduktase oxidiert NADH, wobei gleichzeitig eine Protonentranslokation durch die Membran erfolgt. Das Genom von *G. oxydans* weist lediglich eine NADH-Dehydrogenase vom Typ II (Ndh II) auf. Kodiert wird dieses Enzym in *G. oxydans* von ORF 0280. Dieser ORF umfasst 1227 bp und somit enthält das zugehörige Protein 409 Aminosäuren. In *E. coli* weist die Ndh II 434 Aminosäuren auf und stellt ein Enzym dar, welches aus nur einer Untereinheit aufgebaut ist (JAWOROWSKI et al., 1981). Dieses Protein ist zwar in der Lage NADH zu oxidieren, jedoch werden dabei keine Protonen gepumpt. Trotz dieser Befunde scheint auch bei Organismen, die sowohl mit einem Komplex I als auch mit einer Ndh II ausgestattet sind, die NADH-Dehydrogenase von Typ II einen erheblichen Beitrag zur Atmungsaktivität beizutragen. In *Azotobacter vinelandii*, einem strikt aeroben Bakterium, welches in der Lage ist molekularen Stickstoff zu fixieren, wurde beobachtet, dass Ndh II-Mutanten einen merklichen Rückgang in der Atmungsaktivität

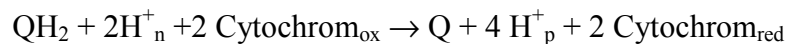
aufweisen (BERTSOVA et al., 2001). In *E. coli* wird Ndh II durch das *fnr*-Genprodukt reprimiert. Eine optimale Ndh II-Expression findet bei hohen Sauerstoffkonzentrationen statt (SHARROCKS et al., 1991; SPIRO et al., 1989).

Die Succinat-Dehydrogenase (Komplex II der Atmungskette) ist eines der Enzyme des zentralen Stoffwechsels, die bei *G. oxydans* nicht vorhanden sind (s. 3.6.2.4). Damit fehlt diesem Organismus jedoch nicht nur ein Enzym des Zitratzyklus, sondern auch ein weiterer Bestandteil der Atmungskette.

Als Endoxidasen wurden in *G. oxydans* zwei Chinol-Oxidasen identifiziert, eine Chinol-Oxidase vom Typ *bd* und eine Chinol-Oxidase vom Typ *bo₃*. Damit liegt, wie auch bei *E. coli*, eine Atmungskette vor, die auf dem Level von Ubichinon verzweigt. In *E. coli* ist die *bd*-Typ Oxidase weniger effizient als die Chinol-Oxidase vom Typ *bo* (CALHOUN et al., 1993). Interessanterweise wurde gezeigt, dass bei hohen Sauerstoffkonzentrationen und diazotrophen Wachstum in *Azotobacter vinelandii* die Endoxidase vom Typ *bd* genutzt wird, bei gleichzeitig hoher Expression der Ndh II (MOSHIRI et al., 1991; WU et al., 1997). Erklärt wurde diese Beobachtung damit, dass Ndh II und die *bd*-Typ-Oxidase möglicherweise Komponenten des Schutzmechanismus gegen molekularen Sauerstoff für die Nitrogenase darstellen (BERTSOVA et al., 2001). Die Cytochrome *bo₃* Chinol-Oxidase ist in *E. coli* hauptsächlich bei hohen Sauerstoffkonzentrationen aktiv. Sie wird in Abhängigkeit von der Sauerstoffkonzentration sowohl durch das *fnr*-, als auch durch das *arc*-Genprodukt reguliert (ANRAKU & GENNIS, 1988; KRANZ & GENNIS, 1983; RICE & HEMPFLING, 1978). Bezüglich der Regulation der Chinol-Oxidasen *bd* und *bo₃* in *G. oxydans* ist zur Zeit noch nichts bekannt. Im Institut für Mikrobiologie und Genetik der Universität Göttingen werden jedoch derzeit Tanskriptionsstudien durchgeführt. Diese beinhalten die Erstellung eines DNA-Microarrays, dessen Grundlage die genomische Sequenz von *G. oxydans* bildet. Es ist davon auszugehen, dass die Ergebnisse dieser Untersuchungen bereits Hinweise auf die Expression der Endoxidasen in *G. oxydans* liefern.

Die oben beschriebenen Chinol-Oxidasen sind die einzigen Endoxidasen, die im Genom von *Gluconobacter* identifiziert wurden. Eine Cytochrom c-Oxidase konnte weder durch das Sequenzierungsprojekt noch durch biochemische Untersuchungen gefunden

werden. Bereits 1994 beschrieben MATSUSHITA et al., dass *Gluconobacter*-Stämmen eine Cytochrom c-Oxidase-Aktivität fehlt. Um so erstaunlicher ist es, dass die Analyse des Genoms eindeutig das Vorhandensein eines bc_1 -Komplexes, also einer Ubichinol: Cytochrom c-Oxidoreduktase, ergab. Der bc_1 -Komplex besteht aus drei Proteinen, die Elektronen transportieren. Dabei handelt es sich um ein Cytochrom b, an das zwei Häm b-Gruppen nicht-kovalent gebunden sind, ein Cytochrom c1, welches ein kovalent gebundenes Häm c enthält und ein Rieske Eisen-Schwefel-Protein. Dieses Enzym oxidiert Ubichinol und reduziert ein lösliches c-Typ-Cytochrom. Die Redox-Reaktion ist an eine Protonentranslokation über die Membran gekoppelt (HUNTE et al., 2003). Der zugrunde liegende Mechanismus ist der sogenannten Q-Zyklus, der durch folgende Gleichung beschrieben werden kann (TRUMPOWER, 1990):

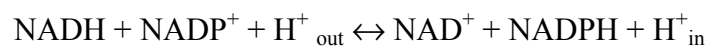


Die Buchstaben n und p bezeichnen die negative, bzw. positive Seite der Membran. In Mitochondrien und vielen Bakterien werden die Elektronen nachfolgend über Cytochrom c auf die Cytochrom c-Oxidase (Komplex IV) übertragen. Der Komplex IV der Atmungskette stellt die Endoxidase dar und transferiert die Elektronen auf Sauerstoff als terminalen Elektronenakzeptor.

Der in *G. oxydans* identifizierte bc_1 -Komplex wird von drei ORFs kodiert, die in einem Cluster organisiert sind (ORFs 1914-1916). ORF 1914 kodiert für das Rieske Eisen-Schwefel-Protein, ORF 1915 für die Cytochrom b-Untereinheit der Ubichinol-Cytochrom c-Reduktase und ORF 1916 für Cytochrom c1. Wie in Kapitel 3.6.5.1 und oben bereits beschrieben, besitzt *G. oxydans* jedoch keine Cytochrom c-Oxidase. Daher kann über die Funktion dieses Komplexes in *G. oxydans* nur spekuliert werden. Eine Möglichkeit könnte sein, dass der Komplex in der stationären Wachstumsphase von Bedeutung ist, wenn Substrate weitgehend aufoxidiert sind und aufgrund von Säurebildung (Acetat oder Gluconat) der pH-Wert stark abgefallen ist (DE LEY et al., 1984). Es besteht dann die Möglichkeit, dass der bc_1 -Komplex reduziertes Ubichinol nutzt um Cytochrom c zu reduzieren. Diese Reaktion hat zwei Effekte: 1. Der elektrochemische Protonengradient wird aufrechterhalten, da im Zuge des Elektronentransports Protonen nach außen gepumpt werden. 2. Reduziertes Cytochrom c könnte als Elektronendonator für Cytochrom-abhängige Dehydrogenasen dienen. Diese

Enzyme würden dann nicht in der oxidativen Richtung arbeiten sondern eine rückwertige Reaktion katalysieren, wobei Säuren wieder in Aldehyde bzw. Alkohole umgewandelt werden. Beispiele hierfür könnten die Gluconat-Dehydrogenase, die Alkohol- und die Acetaldehyd-Dehydrogenase sein, die alle eine Cytochrom c-Untereinheit aufweisen (SHINAGAWA et al., 1981; MATSUSHITA et al., 1994). Durch die Reduktion von Gluconat zu Glucose bzw. von Acetat zu Ethanol könnte somit Energie gewonnen und der pH-Wert erhöht werden. Ob und unter welchen Bedingungen dieser *bc₁*-Komplex gebildet wird, muss durch geeignete Versuche, z. B. durch Transkriptionsanalysen, in der Zukunft erst noch gezeigt werden. Ebenso die energetischen Aspekte, also die Frage ob mit Hilfe des *bc₁*-Komplexes in *G. oxydans* eine Protonentranslokation erfolgt, kann erst durch biochemische Experimente geklärt werden.

Neben den Chinol-Oxidase der Atmungskette konnte in *G. oxydans* ein weiteres Enzym identifiziert werden, das möglicherweise zur Protonentranslokation in der Lage ist. Es handelt sich dabei um die in 3.6.4 beschriebene Transhydrogenase. Die von Transhydrogenasen katalysierte Reaktion,



kann entweder durch eine membrangebundene, an der Protonentranslokation beteiligte Transhydrogenase, oder aber durch eine lösliche, energie-unabhängige Transhydrogenase katalysiert werden (HOEK & RYDSTROM, 1988; BIZOUARN et al., 2002). In der Regel ist jedoch nur eine der beiden Formen vorhanden. Eine Ausnahme bildet *E. coli*, da in diesem Organismus sowohl eine lösliche als auch eine membrangebundene Transhydrogenase identifiziert wurde (CLARKE et al., 1986; BOONTRA et al., 1999).

Aus den Informationen, welche die Analyse des Chromosoms von *G. oxydans* lieferte, lässt sich das Vorhandensein einer membrangebundenen Transhydrogenase voraussagen. Unter den meisten physiologischen Bedingungen ist dieses Enzym jedoch ein Konsument des Membranpotentials, um NADPH bereitzustellen (JACKSON, 2003). Dabei wird jedoch davon ausgegangen, dass ein Substrat über die Glykolyse, den Zitratzyklus und letztendlich über die Atmungskette abgebaut wird. Da *Gluconobacter*

das Schlüsselenzym der Glykolyse, die Phosphofruktokinase fehlt, muss z. B. die Glucose über den Pentosephosphat-Weg verstoffwechselt werden, wobei NADPH entsteht. Auch bei der cytoplasmatischen Umsetzung von Glucose zu Gluconat durch die Glucose-Dehydrogenase entsteht NADPH. Daher stellt sich bei *Gluconobacter* eher die Frage, wie NADP regeneriert wird.

Für *E. coli* wurde von SAUER et al. (2004) gezeigt, dass PntAB eine Quelle für NADPH ist. Bei PntAB handelt es sich um die membrangebundene, energieabhängige Transhydrogenase. Die lösliche, energieunabhängige Transhydrogenase UdhA aus *E. coli* ist dann essentiell, wenn NADPH in ausreichenden Mengen vorhanden ist. Dieses ist z. B. beim Wachstum auf Acetat oder in Mutanten der Fall, die die Glykolyse nicht mehr durchlaufen können, so dass Glucose über den Pentosephosphat-Weg abgebaut wird. Diese Befunde würden eher dafür sprechen, dass bei *G. oxydans* eine lösliche Transhydrogenase vorliegt. Die Ähnlichkeiten der ORFs 1600, 1601 und 1602 zu dem membrangebundenen Enzym aus *R. rubrum* (BIZOUARN et al., 1996) und der Aufbau aus mehreren Untereinheiten deuten jedoch darauf hin, dass es sich um eine membrangebundene Transhydrogenase handelt. Das Protein aus *R. rubrum* umfasst drei Untereinheiten PntAA, PntAB and PntB, die an der Bindung von NAD (PntAA) bzw. NADP (PntB) und an der Verankerung in der Membran (PntAB) beteiligt sind (WILLIAMS et al., 1994). Wie bereits oben erwähnt sind die abgeleiteten Aminosäuresequenzen der ORF 1600-1602 aus *G. oxydans* hoch homolog zu den Untereinheiten des protonentranslozierenden Enzyms aus *R. rubrum*. Dieser Befund spricht eindeutig dafür, dass *G. oxydans* ebenfalls über ein derartiges Enzym verfügt. Eine Bestätigung dieser Befunde würde bedeuten, dass zum einen ein weiteres Enzym zum Aufbau des Membranpotentials vorhanden wäre, zum anderen wäre es eine Erklärung für die schon lange bestehenden Frage nach der Regenerierung von NADP in *Gluconobacter*-Stämmen, für die es schon unterschiedlichste Ansätze gab:

Basierend auf Enzymmessungen mit zellfreien Extrakten und der Kinetik der Glucose-Oxidation und der Glucose-Aufnahme, wurde vorgeschlagen, dass Glucose hauptsächlich durch die cytoplasmatische, NADP-abhängige Glucose-Dehydrogenase zu Gluconat oxidiert wird (OLIJVE & KOK, 1979). Dieses würde jedoch einen erhöhten Spiegel an NADPH nach sich ziehen. EAGON berichtete 1963, dass *G. oxydans* nicht im Besitz einer Transhydrogenase für die Regeneration von NADP ist.

Daher wurden andere Modelle vorgeschlagen, welche die Anhäufung von NADPH vermeiden, bzw. die Regeneration von NADP bewirken. Ein Modell postulierte das sogenannte Gluconat-Recycling. Dabei wurde angenommen, dass das bei der Glucose-Oxidation entstandene Gluconat hauptsächlich durch die membrangebundene Gluconat-Dehydrogenase zu Ketogluconat umgesetzt und anschließend durch die lösliche Ketogluconat-Reduktase zu Gluconat reduziert wird (OLIJVE & KOK, 1979). Ein alternativer Vorschlag zur Regeneration von NADP bestand in der Aktivität der NADPH-Oxidase („old yellow enzyme“, ADACHI et al., 1979). LEVERING et al. (1988) postulierten, basierend auf Enzymmessungen, dass die Glucose-Oxidation hauptsächlich durch die membrangebundene Glucose-Dehydrogenase erfolgt, wodurch erst gar kein NADPH entstehen würde und damit auch die Frage nach der Regeneration von NADP weniger signifikant wäre. Ein Vorschlag, der auch in die Richtung geht, eine übermäßige Produktion von NADPH zu vermeiden, ist eine duale NAD/NADP Coenzym-Spezifität der Glucose-6-Phosphat-Dehydrogenase und der 6-Phosphogluconat-Dehydrogenase aus dem Pentosephosphat-Weg (TONOUCHI et al., 2003). Auch wenn dieses Modell eher darauf abzielt, einen möglichen Mechanismus zur Generation von NADH zu postulieren, so würde auch eine Anhäufung von NADPH verhindert werden. Zusammenfassend kann man sagen, dass bezüglich der Regeneration von Cofaktoren in *Gluconobacter* bislang nur wenig bekannt ist. Das Vorliegen der genomischen Sequenz bietet jedoch eine gute Grundlage, derartige Mechanismen genauer zu untersuchen.

4.3 Chinoproteine

Die oxidative Fermentation ist im Bereich der Biotechnologie vor allem bei der Produktion von Essig und L-Sorbose, einem Vorläufer von Ascorbinsäure, bereits etabliert. Unter dem Begriff oxidative Fermentation versteht man die unvollständige Oxidation eines Substrats, die mit der Anhäufung des entsprechenden Produkts im Medium einhergeht. Für die industrielle Nutzung von Mikroorganismen ist vor allem die Herstellung von reinen Enantiomeren von Interesse, da die Synthese der Verbindungen im Zuge der chemischer Synthese eine aufwendige Schutzgruppenchemie erfordert.

In den Prozess der oxidativen Fermentation durch Essigsäurebakterien sind zwei verschiedene Typen membrangebundener Enzyme involviert: Flavoproteine und Chinoproteine. In beiden Gruppen existieren Proteine die identische Substrate umsetzen, wobei die Oxidationsprodukte jedoch unterschiedlich sind. So produzieren die Flavin-haltige Gluconat-Dehydrogenase und Sorbitol-Dehydrogenase aus *Gluconobacter* 2-Keto-Gluconat und D-Fructose, während die entsprechenden Chinoproteine 5-Keto-Gluconat und L-Sorbose freisetzen (ADACHI et al., 2003a).

In *G. oxydans* 621H konnte jedoch keine PQQ-abhängige Gluconat-Dehydrogenase identifiziert werden (diese Arbeit). Die Flavin-haltige Gluconat-Dehydrogenase wird von den ORFs 2631 (Cytochrom c-Untereinheit), 2632 (alpha-Untereinheit) und 2633 (beta-Untereinheit) kodiert. Im Fall der Flavin-haltigen Sorbitol-Dehydrogenase sind es die ORFs 0754 (Cytochrom c Untereinheit), 0755, 0756 (große Untereinheit) und 0757 (kleine Untereinheit), welche die genetische Information für dieses Protein enthalten. Es ist jedoch anzunehmen, dass das Enzym aufgrund einer natürlichen Mutation (frame shift) in der großen Untereinheit (ORF 0755, 0756) nicht funktionell ist. Die PQQ-enthaltende Sorbitol-Dehydrogenase wird in dem hier sequenzierten Stamm von den ORFs 2230 (*sldA*) und 2231 (*sldB*) kodiert.

Chinoproteine sind eine Klasse von Oxidoreduktasen, die einen Chinon-Cofaktor enthalten. Hierbei handelt es sich meistens um Pyrrolochinolin-Chinon (PQQ), aber gelegentlich sind auch andere Chinon-Cofaktoren wie Topachinon (TPQ), Tryptophan-Tryptophyl-Chinon (TTQ) oder Lysin-Tyrosyl-Chinon (LTQ) zu finden. Oxidoreduktasen, welche PQQ als Cofaktor binden, sind in der Regel in der Membran lokalisiert, während Enzyme mit TPQ, TTQ und LTQ als Cofaktor cytoplasmatische Proteine sind. Chino- und Chinohämoproteine, also Enzyme die zusätzlich ein Häm c enthalten, wurden neben Pyrimidin Nukleotid- und Flavin-abhängigen Dehydrogenasen als dritte Klasse der Redox-Enzyme eingestuft (MATSUSHITA et al., 2001).

4.3.1 Allgemeine Merkmale von Chinoproteinen

Bei den Chinoproteinen handelt es sich um primäre Dehydrogenasen, die ihre Reduktionsäquivalente direkt in die Atmungskette einspeisen. Derartige Enzyme sind charakteristisch für rudimentäre und wenig effiziente Atmungsketten (MATSUSHITA

et al., 1994). Aufgrund des Fehlens der Komplexe I und IV der Atmungskette (s. 3.6.5.1) liegt auch bei dem hier sequenzierten Stamm *G. oxydans* 621H eine solche rudimentäre Atmungskette vor. Demgegenüber besitzt der Stamm eine Vielzahl von membrangebundenen Dehydrogenasen, die Elektronen einspeisen. Durch die direkte Oxidation von Zuckern, Zucker-Alkoholen oder Alkoholen wird eine energieaufwendige Substrataufnahme und Ausscheidung der Produkte umgangen. Voraussetzung dafür ist, dass die Enzyme membrangebunden sind und die aktiven Zentren in den periplasmatischen Raum ragen, so dass ein Transport der Substrate nicht notwendig ist. Die schnelle und effiziente Oxidation von Substraten, auch von zellschädlichen Stoffen, belegt die biotechnologische Nutzbarkeit dieser Proteine. Die genannten Eigenschaften der PQQ-abhängigen Enzyme sind auch charakteristisch für das in dieser Arbeit untersuchte und sequenzierte Bakterium *G. oxydans* 621H. Einige der Chinoproteine aus *Gluconobacter*-Stämmen wurden bereits beschrieben und finden im Folgenden noch Erwähnung. Darüber hinaus zeigte die Analyse des Genoms aber auch das Vorhandensein von noch nicht beschriebenen Dehydrogenasen. Es liegt nahe, dass mit der Analyse dieser bislang unbekanntenen Dehydrogenasen neue biotechnologische Prozesse etabliert oder aber auch bestehende verbessert werden können.

Chinoproteine, welche PQQ als Cofaktor aufweisen, zeigen mit Ausnahme der hydrophoben, N-terminalen Region, untereinander hohe Homologien. Für die membrangebundene Glucose-Dehydrogenase aus *E. coli* wurde gezeigt, dass diese in der N-terminalen Region fünf transmembrane Helices aufweist (YAMADA et al., 1993). Ein ähnliches Bild lässt sich auch für das homologe Enzym aus *G. oxydans* postulieren (Abb. 43). Bei der Vorhersage der in Abb. 43 dargestellten transmembranen Helices handelt es sich jedoch um einen Computer-generierten Vorschlag, der experimentell erst noch nachgewiesen werden muss. Unterstützt wird diese Topologie aber durch die in der Literatur beschriebenen Befunde. Für die PQQ-abhängige, membrangebundene Glucose-Dehydrogenase aus *G. oxydans* wurde gezeigt, dass die Substitution einer Aminosäure die Substratspezifität verändert, bzw. erweitert. Durch die Mutation konnte das Enzym nicht nur die Oxidation von Glucose, sondern auch die von dem Disaccharid Maltose katalysieren (CLETON-JANSEN et al., 1991). Auch dieser Befund eröffnet ein neues Feld für den Einsatz in der Biotechnologie. Auf der Basis der genomischen Sequenz können gezielt Mutationen eingefügt werden, die

Gluconobacter-Stämme befähigen könnten, neben den bekannten Substraten, auch noch weitere ungewöhnliche Substrate umzusetzen, die von industriellen Interesse sind.

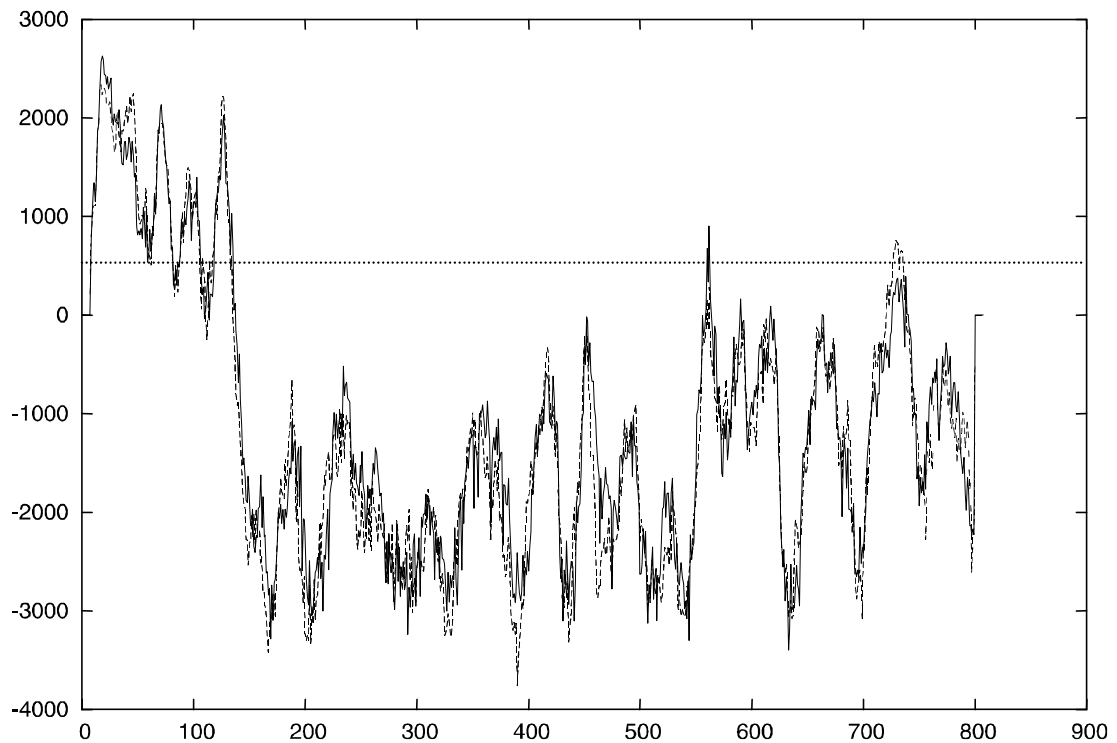


Abb. 43: Vorhersage der transmembranen Helices für die membrangebundene Glucose-Dehydrogenase aus *G. oxydans*

Die Ordinate gibt die Wahrscheinlichkeit für mögliche transmembrane Helices wieder. Scores die größer als 500 sind werden als signifikant bezeichnet. Die Abszisse zeigt die Position der Aminosäuren, beginnend vom N-Terminus. Die gepunktete horizontale Linie markiert den Cut-off (500) für die Wahrscheinlichkeit, dass es sich um eine transmembrane Helix handelt.

Die Eigenschaften von Chinoproteinen, zusammen mit dem einzigartigen Substratspektrum, legen biotechnologische Anwendungen solcher Enzyme nahe. Denkbar wäre z. B. der Einsatz als Biosensoren oder die biologische Umsetzung bestimmter Substrate zu nützlichen Verbindungen. Auch im Bereich der pharmazeutischen Industrie könnten die von *G. oxydans* katalysierten Prozesse Anwendung finden. Aus dem thermotoleranten Stamm *G. frateurii* wurde ein Enzym aufgereinigt, welches *meso*-Erythritol zu L-Erythrulose oxidiert. Auch bei diesem Enzym spekulierten die Autoren, dass es sich um eine PQQ-abhängige Dehydrogenase handelt (MOONMANGMEE et al., 2002). Die enantiomerenreine L-Erythrulose kann weiter zu dem Neurotransmitter γ -Aminobuttersäure oder einem Antiepileptikum

verarbeitet werden. Die molekulare Masse des aufgereinigten Enzyms beträgt etwa 80 kDa (MOONMANGMEE et al., 2002). Auffallend an diesem Enzym ist das breite Substratspektrum. Neben der Oxidation von *meso*-Erythritol zeigt diese Dehydrogenase auch signifikante Aktivitäten bei der Umsetzung von Glycerin, Ribitol, D-Arabitol, D-Sorbitol, D-Mannitol, D-Glucose und Acetaldehyd (MOONMANGMEE et al., 2002).

Ein breites Substratspektrum ist für PQQ-abhängige Dehydrogenasen zunächst nicht ungewöhnlich (ADACHI et al., 2003b), allerdings handelt es sich dabei um Substrate, welche bereits von bekannten und beschriebenen Dehydrogenasen aus *Gluconobacter* umgesetzt werden. Vor diesem Hintergrund drängt sich die Frage auf, ob es sich hierbei um ein neues, bislang unbekanntes Enzym handelt oder aber ob ein bekanntes Enzym neu beschrieben wurde. Denkbar wäre auch eine Mutation in dem entsprechenden Gen, so dass die Substratspezifität verändert ist, so wie es für die Glucose-Dehydrogenase gezeigt wurde (CLETON-JANSEN et al., 1991). Das grundsätzliche Problem bei den beschriebenen Untersuchungen zur Enzymcharakterisierung ist, dass oftmals keine Ansequenzierung der Proteinuntereinheiten erfolgt sind oder diese nicht veröffentlicht wurden. Als Folge werden dann schon bekannte Enzyme neu beschrieben und publiziert. Die Veröffentlichung der N-terminalen Sequenzen könnte dies verhindern und würde dazu beitragen, die Konfusion bezüglich der Nomenklatur der Dehydrogenasen aus *Gluconobacter*-Stämmen zu verringern. Diese Zusammenhänge werden im folgenden Kapitel noch einmal verdeutlicht.

4.3.2 PQQ-abhängige Proteine aus *G. oxydans*

In der Vergangenheit wurden vor allem aus *Gluconobacter*-Arten immer wieder neue Dehydrogenasen gereinigt, wobei die Enzyme zwar biochemisch charakterisiert wurden, die zugrunde liegende Aminosäuresequenz war aber nicht bekannt. So wurde aus *G. suboxydans* IFO 3257 eine membrangebundene Arabitol-Dehydrogenase aufgereinigt und es wurde gezeigt, dass es sich um ein Chinoprotein handelt. Das Enzym ist aus zwei unterschiedlichen Untereinheiten mit molekularen Massen von 82 und 14 kDa aufgebaut. Als besonders herausragend stellten die Autoren die breite Substratspezifität dar. Die Arabitol-Dehydrogenase setzt nicht nur Pentitole, sondern auch andere Polyhydroxy-Alkohole wie D-Sorbitol, D-Mannitol, Glycerin, *meso*-Erythritol und 2,3-Butandiol um (ADACHI et al., 2001).

Aus *G. suboxydans* IFO 3255 wurde eine PQQ-abhängige Sorbitol-Dehydrogenase isoliert, welche ein ähnlich breites Substratspektrum aufwies, wie die Arabitol-Dehydrogenase (SUGISAWA & HOSHINO, 2002). Zunächst wurde beschrieben, dass dieses neue Enzym aus einer Untereinheit mit einer molekularen Masse von etwa 80 kDa aufgebaut ist. Es zeigte sich jedoch, dass für eine aktive membrangebundene Sorbitol-Dehydrogenase aus *Gluconobacter* die Genprodukte von *sldA*, *sldB* und die Anwesenheit von PQQ notwendig sind (MIYAZAKI et al., 2002). Es ist davon auszugehen, dass bei der Aufreinigung der Sorbitol-Dehydrogenase SldB durch die Reinigungsschritte verloren ging. Das Genprodukt von *sldB* ist ein Polypeptid, das aus 126 hydrophoben Aminosäuren besteht. SldB besitzt offensichtlich eine Chaperon-ähnliche Funktion und hat Einfluss auf die korrekte Faltung von SldA um eine aktive Sorbitol-Dehydrogenase zu formen (SHINJOH et al., 2002a; HOSHINO et al., 2003). Weitere Überprüfungen der Ergebnisse bezüglich der Arabitol- und Sorbitol-Dehydrogenase, einschließlich einer N-terminalen Aminosäuresequenz, zeigten, dass es sich um ein und dasselbe Enzym handelt (MATSUSHITA et al., 2003). Diese Befunde verdeutlichen, wie wichtig sowohl die Untersuchung der zugrunde liegenden Sequenz entsprechender Enzyme, als auch die biochemische Charakterisierung ist. Auch im Genom von *G. oxydans* wird *sldB* gefunden. Es liegt zusammen mit *sldA* in einem Cluster vor (ORF 2230 und 2231). Das Genprodukt von *sldB* weist im Einklang mit der Literatur 126 Aminosäuren auf.

4.3.2.1 Die membrangebundene Alkohol-Dehydrogenase aus *G. oxydans*

Neben den erwähnten Zucker- und Polyol-Dehydrogenasen, ist *Gluconobacter* auch für membrangebundene, PQQ-abhängige Alkohol-Dehydrogenasen bekannt. Die Oxidation von Ethanol zu Essigsäure durch Essigsäurebakterien ist bei der Produktion von Essig ein charakteristischer Prozess (DE LEY et al., 1984). Das „prominenteste“ Chinoprotein aus *Gluconobacter* ist daher vermutlich die membrangebundene Alkohol-Dehydrogenase. Hierbei handelt es sich um ein Chino-Hämo-Protein, also ein Enzym, welches neben PQQ auch noch Cytochrom c enthält. Die Alkohol-Dehydrogenase der Essigsäurebakterien fungiert als primäre Dehydrogenase in der Ethanol-Oxidase-Atmungskette. Die Elektronen werden direkt auf Ubichinon übertragen (MATSUSHITA et al., 1994).

Die PQQ-abhängigen Alkohol-Dehydrogenasen (ADHs) werden in drei verschiedene Klassen (ADH I, ADH II, ADH III) eingeteilt, wobei ADHs der Klasse I und insbesondere der Klasse II eine außerordentlich breite Substratspezifität aufweisen. Dadurch katalysieren diese Enzyme auch die Oxidation von ungewöhnlichen Substraten (MATSUSHITA et al., 2002). Typische Substrate für ADHs sind neben Ethanol 1-Butanol, 1,2-Propandiol (TOYAMA et al., 1995), Polyethylen-Glycol (YASUDA et al., 1996), Tetrahydrofuryl-Alkohol (ZARNT et al., 2001) und Polyvinyl-Alkohol (SHIMAO et al., 1986). Das Genom von *G. oxydans* enthält zwei Gene, die für die beiden Untereinheiten der ADH aus diesem Organismus kodieren. Die große Untereinheit stellt die eigentliche Dehydrogenase (ORF 2455, 757 Aminosäuren) dar, während die kleinere Untereinheit ein c-Typ Cytochrom ist (478 Aminosäuren). Weiterhin wurde festgestellt, dass die Untereinheit I ein und die Untereinheit II drei Häm c-Moleküle enthalten. Zwei der Hämgruppen der Untereinheit II und die der Untereinheit I sind dabei in den intramolekularen Elektronentransport der Alkohol-Dehydrogenase auf Ubichinon involviert (MATSUSHITA et al. 1996). In vielen Essigsäurebakterien weist dieses Enzym noch eine weitere Untereinheit mit einer molekularen Masse von etwa 15 kDa auf, die Funktion dieser Untereinheit ist jedoch nicht geklärt (KONDO & HORINOUCI, 1997a). In *G. oxydans* 621H wird dieses Polypeptid von ORF 2124 kodiert.

Die Alkohol-Dehydrogenase aus *Acetobacter polyoxogenes* besteht z. B. aus zwei Untereinheiten (78 kDa und 44 kDa; TAYAMA et al., 1989), die aus *Acetobacter aceti* weist dagegen noch eine dritte Untereinheit von etwa 20 kDa (AdhS) auf (MATSUSHITA et al., 1992). Es wurde gezeigt, dass die Cytochrom c-Untereinheit essentiell für die Ethanol-Oxidase-Atmungskette ist und möglicherweise als Bestandteil einer alternativen, Cyanid-insensitiven Atmungskette von *G. suboxydans* dient (MATSUSHITA et al. 1991) Darüber hinaus wurde gezeigt, dass die durch die membrangebundene Alkohol-Dehydrogenase katalysierte Reaktion reversibel ist, so dass das Protein sowohl Ubichinon reduzieren als auch Ubichinol oxidieren kann (MATSUSHITA et al., 1999).

Neben der oben beschriebenen membrangebundene Alkohol-Dehydrogenase wurde aus *G. frateurii* eine membrangebundene cyclische Alkohol-Dehydrogenase aufgereinigt.

Diese katalysiert ausschließlich die Oxidation zyklischer Alkohole zu den korrespondierenden Ketonen, nicht aber die Rückreaktion (MOONMANGMEE et al., 2001). Das aufgereinigte Protein wies eine molekulare Masse von 83 kDa auf und konnte die Oxidation verschiedener cyclischer Alkohole katalysieren. Geringe Aktivitäten wurden bei der Oxidation aliphatischer sekundärer Alkohole verzeichnet. Im Gegensatz zu der oben beschriebenen Alkohol-Dehydrogenase der Ethanol-Atmungskette wurden jedoch keine primären Alkohole oxidiert. Signifikante Aktivitäten zeigte das Enzym außerdem bei der Umsetzung von 2,3-Butandiol, 2-Butanol, Acetaldehyd und *meso*-Erythritol. Anzumerken bleibt, dass keine Sequenzinformationen für dieses Protein verfügbar sind, so dass die Zuordnung zu einem Gen aus dem in dieser Arbeit sequenzierten Stamm 621H nicht möglich ist.

4.3.2.2 Chinat-Dehydrogenase

3-Dehydrochinat stellt ein Intermediat des Shikimat-Weges dar. Dieser Stoffwechselweg dient zur Synthese der aromatischen Aminosäuren sowie anderer aromatischer Verbindungen. Die Synthese beginnt mit einer Kondensation von Erythrose-4-phosphat und Phosphoenolpyruvat. Kürzlich wurde aus der Cytoplasmamembran von *Acinetobacter calcoaceticus* eine PQQ-abhängige Chinat-Dehydrogenase aufgereinigt (ADACHI et al., 2003c). Hohe Enzymaktivitäten waren auch in den Membranfraktionen von *G. melanogenus* IFO 3294, *G. oxydans* IFO 3292 und *G. oxydans* IFO 3244 zu messen (ADACHI et al., 2003c). Die im SDS-Gel abgeschätzte molekulare Masse des aufgereinigten Proteins betrug 88 kDa. Erste Analysen zur oxidativen Fermentation von Chinat zu 3-Dehydrochinat zeigten in der frühen Wachstumsphase verschiedener *Gluconobacter*-Stämme, wie *G. oxydans* IFO 3244, IFO 3292, *G. melanogenus* IFO 3294 und *G. liquefaciens* IFO 12388 eine Ausbeute von nahezu 100 %. Weiterhin stellte sich heraus, dass *Gluconobacter*-Stämme geeignete Biokatalysatoren für die Oxidation von Chinat sind (ADACHI et al., 2003d). Durch die cytoplasmatische Dehydrochinat-Dehydrogenase wurde das Produkt weiter zu 3-Dehydroshikimat oxidiert.

Es stellte sich die Frage ob ein derartiges Enzym möglicherweise auch in dem hier sequenzierten Stamm 621H vorkommt. Ein BLAST-Abgleich mit der NCBI (<http://www.ncbi.nlm.nih.gov/>) der nicht-charakterisierten PQQ-abhängigen

Dehydrogenasen aus dem Genom zeigte, dass der ORF 0494 42 % Ähnlichkeit zu einer putativen Chinat-Dehydrogenase aus *Pseudomonas putida* (NELSON et al., 2002) und 40 % Ähnlichkeit zu einer Chinat/Shikimat-Dehydrogenase aus *A. calcoaceticus* (ELSEMORE & ORNSTON, 1994) aufweist. Die aus der Aminosäuresequenz abgeleitete molekulare Masse des mutmaßlichen Proteins betrug 85 kDa. Allerdings wies dieser ORF auch 45 % Ähnlichkeit zu einer PQQ-abhängigen Glucose-Dehydrogenase aus *P. syringae* auf (BUELL et al., 2003). Wie an anderer Stelle bereits erwähnt wurde, sind PQQ-abhängige Enzyme untereinander sehr homolog. Von daher ist nur anhand des BLAST-Ergebnisses keine Aussage über die Funktion des ORFs 0494 möglich. Das Vorliegen der Sequenz bietet jedoch die Möglichkeit, diesen ORF gezielt zu klonieren und zu exprimieren, um genaue biochemische Untersuchungen durchzuführen.

In *A. calcoaceticus* wird die Chinat-Dehydrogenase von dem Gen *quiA* kodiert (ELSEMORE & ORNSTON, 1995) und befindet sich in einem Cluster *quiBCXA*. Die Gene *quiB* und *quiC* kodieren für eine Dehydrochinat-Dehydratase und eine Dehydroshikimat-Dehydratase. Die Funktion von *quiX* ist nicht geklärt; es wird jedoch vermutet, dass dieses Gen für ein Porin kodiert. In dem Genom von 621H findet man den ORF 0494, der Ähnlichkeiten zu QuiA aus *A. calcoaceticus* zeigt, isoliert im Genom vorliegen, umgeben von hypothetischen ORFs.

Über die Funktion des ORF 0494 kann somit nur gemutmaßt werden. Eindeutige Ergebnisse können erst durch biochemische Experimente erzielt werden. Interessant dabei ist nicht nur zu überprüfen, ob eine Aktivität für eine Chinat-Dehydrogenase vorliegt, sondern auch ob eventuell noch Aktivitäten für andere Substrate zu verzeichnen sind. Dieses kann aufgrund des breiten Substratspektrums der bislang charakterisierten PQQ-abhängigen Enzyme aus *Gluconobacter*-Stämmen als wahrscheinlich angesehen werden.

Die bereits untersuchten Chinoproteinen aus *Gluconobacter* unterstreichen das große oxidative Potential dieses Organismus. Die in der Regel breite Substratspezifität der Enzyme verdeutlicht aber auch, wie wichtig es ist, ein Protein nicht nur biochemisch zu untersuchen, sondern auch die zugrunde liegende Sequenz zu bestimmen. Die Analyse des Genoms von *G. oxydans* zeigte, dass dieser Organismus noch eine Vielzahl nicht

charakterisierter Dehydrogenasen aufweist. Da PQQ-abhängige Dehydrogenasen untereinander große Homologien aufweisen, konnten auch neue Chinoproteine identifiziert werden (s. 3.6.6). Diese können nun gezielt überproduziert und biochemisch untersucht werden, um das oxidative Potential dieses Organismus weiter zu erforschen.

4.3.3 Der Cofaktor PQQ

Der Redox-Faktor PQQ wurde erstmals 1979 in Bakterien identifiziert (SALISBURY et al., 1979), ist aber für Säugetiere und Menschen nicht weniger wichtig. Mäuse, die mit einer PQQ-Mangel-Diät ernährt wurden, zeigten Defizite beim Wachstum und waren nicht mehr in der Lage sich fortzupflanzen (KILLIGORE et al., 1989; STEINBERG et al., 1994). Dieser Cofaktor ist nicht nur für viele membrangebundene Dehydrogenierungsreaktionen in *Gluconobacter* notwendig, sondern auch Bestandteil verschiedener Nahrungsmittel wie Fleisch und Gemüse (FLÜCKINGER et al., 1988; KUMAZAWA et al., 1992). KASAHARA & KATO (2003) zeigten, dass eine PQQ-abhängige Dehydrogenase in Mäusen für den Abbau der Aminosäure Lysin notwendig ist. Die Analyse des menschlichen Genoms (PRAPHANPHOJ et al., 2001; KASAHARA & KATO, 2003) zeigte, dass ein hoch homologes Enzym auch in *Homo sapiens* vorkommt. Da PQQ offensichtlich als Redox-Cofaktor im Zuge des Lysin-Abbaus bei Säugern fungiert und nicht *de novo* synthetisiert wird, kann festgestellt werden, dass diese Verbindung in die Gruppe der B-Vitamine einzuordnen ist (KASAHARA & KATO 2003). Im Gegensatz dazu verursachte PQQ in Versuchen mit Zellkulturen den Zelltod durch oxidative Schädigung (NUKADA et al., 2003). Aus diesen Ergebnissen schlossen die Autoren, dass PQQ in Abhängigkeit von den Umweltbedingungen in verschiedenen biologischen Systemen als Antioxidans oder aber auch als Pro-Oxidans agieren kann. Diese Ergebnisse zeigen, dass für eine großtechnische Herstellung und vor allem für eine Anwendung von PQQ als Vitamin weitere Studien notwendig sind. Deutlich wird aber auch, dass ein positiver Effekt von PQQ als Nahrungsmittelsupplement nicht abzustreiten ist. Hier ergibt sich möglicherweise ein neues Anwendungsgebiet für den biotechnologischen Einsatz von *G. oxydans* zur Synthese des neuen Vitamins.

Um so wichtiger ist es, die Biosynthese von PQQ zu verstehen. Bislang ist bekannt, dass das aromatische, heterozyklische *o*-Chinon aus einem Peptid-Vorläufer (PpqA) gebildet wird, das ein konserviertes Glutamat und Tyrosin enthält. Alle Kohlenstoff- und Stickstoffatome der genannten Aminosäuren finden sich in dem Cofaktor wieder, da eine Kondensation von Glutamat mit Dopachinon erfolgt. Letzteres entsteht durch den Abbau des bereits erwähnten Tyrosin-Restes. Über Zwischenstufen oder Intermediate dieser Synthese ist jedoch bislang wenig bekannt. Allerdings ist es MAGNUSSEN et al. (2004) kürzlich gelungen, ein Intermediat der PQQ-Biosynthese aus einer Mutante von *Methylobacterium extorquens* zu isolieren und die Struktur dieses Intermediats aufzuklären. Der rekombinante Stamm wies einen Defekt in *ppqC* auf. Das isolierte Intermediat stellte somit das Substrat für PqqC dar. Es zeigte sich, dass es sich um ein vollständig reduziertes Derivat von PQQ handelte, dessen fünfgliedriger Ring nicht geschlossen war. Mit anderen Worten fehlte lediglich der Ringschluss für die vollständige Synthese des Cofaktors PQQ. Diese Ergebnisse zeigen, dass die Reaktion, die von PqqC katalysiert wird, den letzten Schritt bei der Synthese des Cofaktors darstellt.

Die Gene der PQQ-Synthese befinden sich in einem Cluster, welches fünf ORFs umfasst (*ppqA-E*). FELDER et al. (2000) ist es gelungen, dieses Cluster durch Transposonmutagenese in *G. oxydans* IFO 3293 zu mutagenisieren. Die Analyse ergab, dass offenbar nur die Gene *ppqB-E* in einem Operon transkribiert werden, während *ppqA* einen eigenen Promotor aufweist. Eine derartige Organisation zeigt auch das Cluster der PQQ-Biosynthesegene aus 621H (ORFs 2370-2373; diese Arbeit). Während bei den ORFs, welche für die Gene *ppqB-E* kodieren, Start- und Stopcodon jeweils überlappen, befindet sich *ppqA* 119 bp stromaufwärts von *ppqB*. HOUCK et al. (1991) zeigten, dass PqqA auf einem höheren Level exprimiert werden muss, da dieses Peptid ein Vorläufer von PQQ darstellt, wohingegen PqqB-E für die Biosynthese des Cofaktors notwendig sind.

4.4 *G. oxydans* im Dienst der Biotechnologie

Organismen der Gattung *Gluconobacter* wurden schon seit langem in der Biotechnologie eingesetzt. Hintergrund dafür sind die einzigartigen metabolischen Fähigkeiten, die für diesen Organismus charakteristisch sind (MACAULEY et al., 2001;

GUPTA et al., 2001). Ein Grund dafür sind die hohen Oxidationsraten, einhergehend mit einer geringer Biomasse-Produktion (DEPPENMEIER et al., 2002). Ein weiterer Grund sind aber auch die stereo- und regioselektiven Dehydrogenierungsreaktionen, welche von diesem Organismus katalysiert werden. Dadurch lässt sich eine aufwendige und kostenintensive Schutzgruppenchemie z.B. bei der Oxidation von Zuckerderivaten vermeiden. Im Folgenden werden Prozesse, in denen *G. oxydans* bereits im großtechnischen Maßstab eingesetzt wird, vorgestellt, sowie mögliche neue Einsatzgebiete diskutiert.

4.4.1 Essigsäureproduktion

Essig wurde bereits im Altertum zur Konservierung von Fleisch und Gemüse und sogar als Arznei verwendet. Die Dokumentation über die Herstellung von Essig aus Wein geht bis 4000 Jahre v. Chr. zurück. Die derzeitige weltweite Essigproduktion beträgt etwa 1×10^6 l/Jahr (LUSTA & RESHETILOV, 1998). Es existieren drei verschiedene Verfahren zur Herstellung von Essig (GREENSHIELDS, 1978). Beim sogenannten Oberflächenverfahren wird Wein in flachen Schalen gefüllt und Fruchtfliegen (*Drosophila*) sorgen für die Beimpfung des Materials. An der Oberfläche bildet sich eine Haut aus Zellen von *Acetobacter xylinum*. Der Nachteil dieses Verfahrens ist, dass die Säuregärung nur langsam vonstatten geht. Beim Fesselverfahren werden die Bakterien an ein Trägermaterial gebunden, die Belüftung der Behälter erfolgt von unten. Da die Bakterien immobilisiert sind, muss der fertige Essig kaum noch gefiltert werden. Inzwischen hat das Submersverfahren die beiden oben dargestellten Technologien weitgehend abgelöst. Hierbei wird kein Trägermaterial benötigt, da sich die Bakterien direkt in der alkoholischen Maische befinden und intensiv belüftet werden. Der Vorteil dieses Verfahrens ist vor allem der schnelle Herstellungsprozess. Für die Produktion von Essig werden jedoch eher Stämme der Gattungen *Gluconacetobacter* und *Acetobacter* und weniger *Gluconobacter*-Spezies (ASAI, 1968) eingesetzt.

Die Produktion von Acetat durch Essigsäurebakterien erfolgt durch die Oxidation von Ethanol in zwei aufeinander folgenden Schritten. Der erste Schritt wird durch eine membrangebundene Alkohol-Dehydrogenase katalysiert. In dem hier sequenzierten Stamm werden die Untereinheiten von den ORFs 2454 und 2455 kodiert. Es handelt sich bei dem Protein um ein Chinohämoprotein, also um eine PQQ-abhängige Alkohol-

Dehydrogenase, die zusätzlich eine Cytochrom c Untereinheit enthält. Der zweite Schritt wird durch eine ebenfalls membrangebundene Aldehyd-Dehydrogenase katalysiert (MATSUSHITA et al., 1994; ORF 1935, 1936 in *G. oxydans* 621H; diese Arbeit).

4.4.2 Produktion von Vitamin C

Die weltweite Produktion von Vitamin C wird auf etwa 100.000 Tonnen/Jahr geschätzt. Die Herstellung erfolgt durch das Reichstein Verfahren (REICHSTEIN & GRÜSSNER, 1934). Dieses beinhaltet eine Reihe chemischer Oxidationsschritte und einen Schritt, welcher durch Mikroorganismen katalysiert wird. Verwendung findet die L-Ascorbinsäure zum einen in der Lebensmittelindustrie, wo es als Antioxidans eingesetzt wird (HANCOCK & VIOLA, 2002), zum anderen ist es aber auch ein lebenswichtiges Vitamin. Menschen, Affen und Meerschweinchen haben im Lauf der Evolution die Fähigkeit dieses Vitamin selbst zu synthetisieren verloren. Ihnen fehlt das Enzym L-Gulono- γ -Lacton-Oxidase, welches den letzten Schritt der Vitamin C-Synthese, die Umsetzung von L-Gulono- γ -Lacton zu L-Ascorbinsäure, katalysiert.

Die Herstellung von Vitamin C schließt sowohl chemische als auch biotechnologische Schritte ein. Der biotechnologische Schritt besteht in der Umsetzung von D-Sorbitol zu L-Sorbose. Ascorbinsäure weist mehrere optische Isomere auf, jedoch ist nur das L-Isomer biologisch aktiv. Daher ist die stereo- und regioselektive Umsetzung von D-Sorbitol zur L-Sorbose durch *G. oxydans* von besonderer Wichtigkeit. Dadurch lässt sich eine aufwendige Schutzgruppenchemie vermeiden.

Die klassische Synthese von Vitamin C beginnt mit der chemischen Hydrogenierung von D-Glucose zu D-Sorbitol. Mit Hilfe von *G. oxydans* wird D-Sorbitol stereo- und regioselektiv, der Bertrand-Hudson-Regel folgend (KERSTERS et al., 1965), zu L-Sorbose oxidiert. Die Bertrand-Hudson-Regel besagt, dass Polyole in der D-Konfiguration, welche zwei sekundäre Hydroxylgruppen in cis-Stellung aufweisen, zu den korrespondierenden Ketosen oxidiert werden. L-Sorbose wird mit Aceton durch Kristallisation und Kondensation zu Sorbose-Diaceton umgesetzt. Die anschließende Oxidation mit einem Platin-Katalysator führt zur 2-Keto-L-Gulonsäure. Durch

Enolisierung und Lactonisierung entsteht L-Ascorbinsäure. Die Ausbeute beträgt dabei etwa 50 % (BOUDRANT, 1990).

Die Oxidation von D-Sorbitol wird in *Gluconobacter* durch die Sorbitol-Dehydrogenase katalysiert. Diese Klasse von Proteinen beinhaltet verschiedene Isoenzyme. Aus *G. suboxydans* wurde ein Flavoprotein isoliert, welches aus drei Untereinheiten besteht (63, 51, 17 kDa); eine davon beinhaltet ein Häm c (SHINAGAWA et al., 1982). CHOI et al. (1995) reinigten eine Sorbitol-Dehydrogenase aus *G. suboxydans* KCTC 2111 (ATCC 621), die ebenfalls aus drei Untereinheiten aufgebaut (75, 50, 14 kDa) ist. Dieses Enzym ist jedoch kein Flavoprotein sondern enthält als Cofaktor PQQ. Aus dem Stamm IFO 3255 wurde eine Sorbitol-Dehydrogenase gereinigt, die aus nur einer Untereinheit von 80 kDa bestehen soll (SUGISAWA & HOSHINO, 2002). Hier liegt jedoch die Vermutung nahe, dass es sich um die membrangebundene, PQQ-abhängige Sorbitol-Dehydrogenase handelt, welche aus zwei Untereinheiten aufgebaut ist (MATSUSHITA et al., 2003). Es ist anzunehmen, dass die kleine Untereinheit, kodiert durch *sldB*, im Zuge der Aufreinigung verlorengegangen ist. In dem hier sequenzierten Stamm 621H wird die PQQ-abhängige Sorbitol-Dehydrogenase, die aus zwei Untereinheiten besteht, von den ORFs 2230 und 2231 kodiert. Daneben finden sich noch eine cytoplasmatische (ORF 0010) und eine zweite membrangebundene Sorbitol-Dehydrogenase (Orf 0754-0757) in *G. oxydans* 621H (diese Arbeit). Diese Enzyme spielen aber offensichtlich keine Rolle bei der biotechnologischen Anwendung des Organismus zur Vitamin C-Synthese. Das erste Enzym liegt wie erwähnt im Cytoplasma vor und ist NADP-abhängig. Somit kommt diese Sorbitol-Dehydrogenase zur Produktion von L-Sorbose nicht in Betracht. Das membrangebundene FAD-abhängige Enzym ist offensichtlich nicht funktionstüchtig, da ein Frame-Shift im ORF 0755 vorliegt (s. 3.6.6).

Derzeit werden intensive Untersuchungen zur Verbesserung der Ausbeute des biologischen Schrittes durchgeführt. DE WULF et al. (2000) berichteten von einer optimierten Synthese von L-Sorbose durch die regiospezifische Dehydrogenierung von D-Sorbitol durch einen mutagenisierten Stamm von *Gluconobacter oxydans*. Der maximale Produktionsertrag an L-Sorbose erreichte 100% wobei 200 g/l Sorbitol eingesetzt wurden.

Seit einigen Jahren werden Bestrebungen unternommen, die Herstellung von Vitamin C von chemischen Verfahren hin zu biotechnologischen Umsetzungen zu verschieben, wobei der Schwerpunkt der Untersuchungen in der mikrobiellen Produktion von 2-Ketogulonsäure liegt. Dabei wird ein Prozess angestrebt, welcher vollständig durch membrangebundene Enzyme katalysiert wird. Hierbei würden die Substrate im Periplasma oxidiert, was sehr vorteilhaft für die Produktaufreinigung wäre. Darüber hinaus würden auch Nebenreaktionen durch den intrazellulären Stoffwechsel vermieden, da die Substrate gar nicht erst in die Zelle gelangen und somit nicht für andere Stoffwechselwege zur Verfügung stehen.

Wie oben beschrieben, stellt 2-Ketogulonsäure ein Schlüsselintermediat bei der Biotransformation dar, das chemisch einfach zur Ascorbinsäure umgesetzt werden kann. Um diese Verbindung zu produzieren wurden bislang zwei Wege beschritten, welche in Abb. 44 dargestellt sind. Ausgangssubstrat ist die D-Glucose, bzw. D-Sorbitol. Beim „Sorbitol-Weg“, wird das Polyol durch eine Sorbitol- und Sorbose-Dehydrogenase zu L-Sorboson umgesetzt. Durch die Reaktion der Sorboson-Dehydrogenase entsteht 2-Ketogulonsäure. Die Schwierigkeit bei diesem Verfahren liegt darin, die Produktion von 2-Ketogulonsäure zu optimieren. Es ist bislang auch kein natürlicher Bakterienstamm bekannt, der diesen Weg mit hohen Umsatzraten katalysiert. Wie bereits oben erwähnt erfolgt die mikrobielle Produktion von 2-Ketogulonsäure in zwei enzymatischen Schritten. In der ersten Reaktion wird L-Sorbose zu L-Sorboson oxidiert. Durch die Sorboson-Dehydrogenase erfolgt dann die Synthese von 2-Ketogulonsäure (SHINJOH et al., 1995). Die Aufgabe liegt nunmehr darin diese Enzyme in Reihe zu schalten. Einige *Gluconobacter*-Stämme sind bekannt, die als Ausgangspunkt für dieses Verfahren genutzt werden können. Leider gehört der in dieser Arbeit sequenzierte Organismus *G. oxydans* 621H nicht zu dieser Gruppe, da keine Gene zur Bildung der Sorbose- bzw. der Sorboson-Dehydrogenase identifiziert wurden. Allerdings gelingt diese Umsetzung in einigen anderen *Gluconobacter*-Stämmen sowohl durch cytoplasmatische als auch membrangebundene Enzyme, die L-Sorbose über Sorboson zu 2-Ketogulonsäure umsetzen. Die Ertragsraten sind allerdings derzeit noch zu gering. SHINJOH et al. (2002b) zeigten, dass die NADPH-abhängige L-Sorbose-Reduktase aus *G. oxydans* IFO 3291 zur Assimilierung von L-Sorbose beiträgt, nicht aber zur Oxidation von L-Sorbose. Im Cytoplasma dieses Organismus wird 2-Ketogulonsäure über Idonsäure zu Gluconat umgesetzt werden. Dieses kann dann z. B.

in den Pentosephosphat-Weg eingeschleust werden und steht damit für die Vitamin C-Produktion nicht mehr zur Verfügung (HANCOCK & VIOLA, 2002). Eine Mutante von *G. melanogenus* war in der Lage, 100 g D-Sorbitol zu etwa 60 g 2-Ketogulonsäure umzusetzen, jedoch wurden bei diesem Prozess auch Nebenprodukte gebildet (SUGISAWA et al., 1990). Die Abtrennung dieser Verunreinigungen würde bei einer industriellen Nutzung wieder aufwendige Reinigungsschritte erfordern, was somit keine Verbesserung des bestehenden Verfahrens darstellt. URBANCE et al., (2001) isolierten zwei Spezies der Gattung *Ketogulonigenium*, die L-Sorbose zu 2-Ketogulonsäure oxidieren können. Von diesen Stämmen fehlen jedoch bislang Ergebnisse bezüglich der Produktivität. SAITO et al. (1998) isolierten einen *Gluconobacter*-Stamm, der D-Sorbitol in 2-Ketogulonsäure umwandeln kann, die Ausbeute war für eine industrielle Nutzung jedoch zu gering. SHINJOH et al. (1995) entwickelten einen rekombinanten *Gluconobacter*-Stamm, der L-Sorbose zu 2-Ketogulonsäure oxidiert, die Ausbeute betrug dabei etwa 40 g/Liter. Unter fermentativen Bedingungen konnte die Ausbeute an 2-Ketogulonsäure jedoch nicht verbessert werden. SAITO et al. (1997) konstruierten einen rekombinanten *Gluconobacter* Stamm, der D-Sorbitol durch die Reaktion der Sorbitol-Dehydrogenase und der Sorboson-Dehydrogenase zu 2-Ketogulonsäure umsetzte. Dabei wurde die Sorboson-Dehydrogenase aus *G. oxydans* T-100 in den Stamm G624 kloniert. Durch chemische Mutation wurde der Idonat-Weg unterdrückt. Dieses Vorgehen sollte zum einen die Synthese von unerwünschten Nebenprodukten vermeiden und zum anderen die Umsetzung von 2-Ketogulonsäure zu L-Idonsäure mindern, um den Ertrag des gewünschten Produkts zu verbessern. Um die Ausbeute weiter zu erhöhen, wurde der originale Promotor durch einen von *E. coli* ersetzt. Die Produktion von 2-Ketogulonsäure aus D-Sorbitol betrug 130 g/l (SAITO et al., 1997). Trotz dieser enormen Ausbeute war die Synthese von 2-Ketogulonsäure durch einfache Fermentation für eine großtechnische Nutzung noch nicht zufriedenstellend.

Ein anderer Ansatz die chemischen Schritte der Vitamin C-Produktion durch mikrobielle Schritte zu ersetzen besteht darin, die Synthese von 2-Ketogulonsäure in nur einem einzigen Schritt zu bewerkstelligen. Erste Erfolge dahingehend erzielten ANDERSON et al. (1985). Die Arbeitsgruppe klonierten die 2,5-Diketo-Gulonsäure-Reduktase aus *Corynebacterium* sp. und produzierten dieses Protein in *Erwinia herbicola*. Dadurch konnte eine Umsetzung von D-Glucose in 2-Ketogulonsäure in einem einzigen Fermentationsschritt erreicht werden. Aber auch hier müssen noch

Optimierungen für eine Nutzung im industriellen Maßstab stattfinden. SUGISAWA et al. (1995) berichteten von einem *Gluconobacter*-Stamm (DSM 4025), welcher L-Gulono- γ -Lacton zu L-Ascorbinsäure umsetzt. Hier ist anzumerken, dass die erzielte Ausbeute von 15 mg Vitamin C/ml Kultur für einen großtechnischen Einsatz noch bei weitem zu gering ist, wenn man bedenkt, dass im Zuge der chemischen Vitamin C-Synthese im letzten Schritt nur geringe Verluste zu verzeichnen sind.

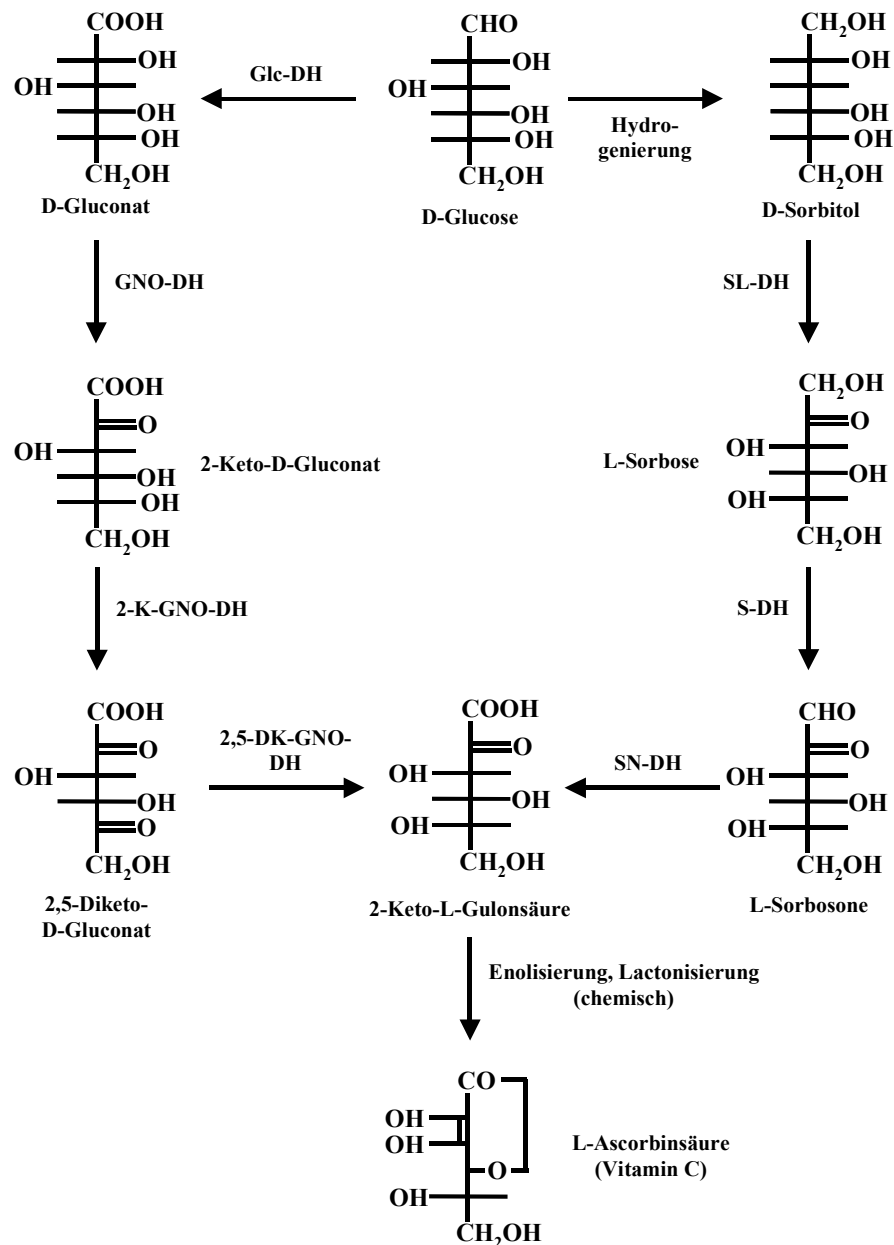


Abb. 44: Stoffwechselwege zur mikrobiellen Herstellung von 2-Ketogulonsäure
 Abkürzungen: Glc-DH: Glucose-Dehydrogenase; GNO-DH: Gluconat-Dehydrogenase; 2-K-GNO-DH: 2-Ketogluconat-Dehydrogenase; SL-DH: Sorbitol-Dehydrogenase; S-DH: Sorbose-Dehydrogenase; SN-DH: Sorboson-Dehydrogenase.

Der zweite in Abb. 44 dargestellte Weg, um zur 2-Ketoglulonsäure zu gelangen, beginnt ebenfalls mit Glucose, wobei das Substrat jedoch durch die Glucose-Dehydrogenase zu Gluconat umgesetzt wird. Mit Hilfe der Gluconat-Dehydrogenase erfolgt die Oxidation der Zuckersäure zu 2-Ketogluconat und abschließend die Umsetzung zu 2,5-Diketogluconat durch die katalytische Aktivität der 2-Keto-Gluconat-Dehydrogenase. Dieses Produkt kann durch die 2,5-Diketo-Gluconat-Reduktase zu 2-Ketogulonsäure umgewandelt werden.

JI & GAO (2001) beobachteten, dass *G. oxydans* ATCC 9937 bevorzugt Gluconat zu 2,5-Diketo-D-Gulonsäure umsetzt. Die NADP-abhängige Glucose-Dehydrogenase reduziert dieses Produkt weiter zu 2-Keto-L-Gulonsäure. Durch diesen zwei-Schritt Mechanismus ist eine Regeneration von NADP bzw. von NADPH nicht notwendig, bzw. wird durch die Glucose-Dehydrogenase bewerkstelligt. Pro Mol eingesetztem Substrat wurden 0.17 mol 2-Ketoglulonsäure erhalten. Durch einen Prozess, welcher nur membrangebundene Enzyme enthält, könnte die Ausbeute wahrscheinlich noch deutlich erhöht werden. Die Schwierigkeit besteht nach wie vor darin, sämtliche Reaktionsschritte von einem Organismus durchführen zu lassen. Dieses wäre nur durch die Herstellung eines rekombinanten Stammes möglich. Die Genomanalyse von *G. oxydans* 621H zeigte z. B. das Vorhandensein von Glucose- und Gluconat-Dehydrogenasen, jedoch konnte kein ORFs identifiziert werden, die für Proteine zur Produktion von Diketogluconat oder 2-Ketogulonsäure kodieren (diese Arbeit). In diesem Zusammenhang zeigten sich jedoch Erfolge beim gleichzeitigen Einsatz von *Gluconobacter oxydans* und *Corynebacterium* sp. Durch eine Co-Immobilisierung der beiden Organismen konnte eine Produktion von 2-Ketogulonsäure aus Gluconat erfolgen. Die Ausbeute betrug dabei 38 %. (AIGUO & PEIJI, 1998).

Zusammengefasst muss festgehalten werden, dass noch große Anstrengungen unternommen werden müssen, um ein Verfahren zur biologischen Vitamin C-Synthese zu entwickeln, das kommerziell konkurrenzfähig ist.

4.4.3 Miglitol und 1-Deoxynojirimycin

1-Deoxynojirimycin ist ein Zuckeranalogon, das die enzymatische Aktivität von Glycosidasen hemmen kann. Glycosidase-Inhibitoren spielen bei der Therapie diverser

Erkrankungen wie Diabetes, Influenza und HIV eine entscheidende Rolle. Die Struktur von 1-Deoxynojirimycin ist ähnlich der von D-Glucose, wobei drei der C-Atome dieselbe sterische Konfiguration zeigen. Innerhalb des Zucker-Rings ist jedoch ein Kohlenstoff-Atom gegen Stickstoff ausgetauscht. Miglitol ist ein Derivat von 1-Deoxynojirimycin (1-N-Hydroxyethyl-1-deoxynojirimycin; SCOTT & TATTERSALL, 1988), welches therapeutisch zur Behandlung von nicht-Insulin-abhängiger Diabetes mellitus eingesetzt wird (CAMPBELL et al., 2000). Andere Derivate von 1-Deoxynojirimycin zeigten einen hemmenden Effekt auf die HIV-1-Replikation (SHIMIZU et al., 1990).

Die Synthese von Miglitol stellt neben der Vitamin C-Produktion ein weiteres Beispiel dar, in dem durch eine Kombination von biotechnologischen und chemischen Schritten kommerzielle Produkte in einem großen Maßstab hergestellt werden. D-Glucose wird dabei erst chemisch durch reduktive Aminierung zu 1-Amino-Sorbitol umgesetzt (Abb. 45). Anschließend erfolgt die stereo- und regioselektive Oxidation von 1-Amino-D-Sorbitol zu 1-Amino-D-Sorbose durch *G. oxydans* (SCHEDEL, 2000), was eine Schlüsselreaktion bei der Synthese des Medikaments darstellt. Das Enzym, welches diese Reaktion katalysiert ist die schon beschriebene Sorbitol-Dehydrogenase. Durch einen chemischen, reduktiven Ringschluß wird das gewünschte Produkt Miglitol gebildet.

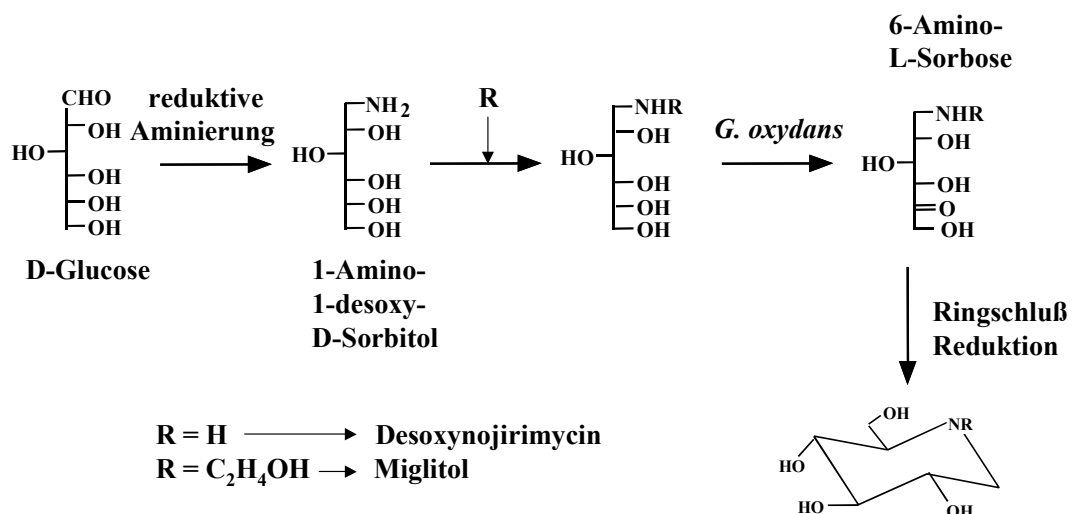


Abb. 45: Synthese von Miglitol (SCHEDEL, 2000)

4.4.4 Produktion von Dihydroxyaceton

Dihydroxyaceton ist eine Verbindung, die vor allem in der Kosmetikindustrie als Selbstbräunungsmittel eingesetzt wird. Darüber hinaus ist es ein Intermediat für viele weitere chemische Verbindungen (CLARET et al., 1994). Laut Angaben des Verbands der Vertriebsfirmen kosmetischer Erzeugnisse (Wiesbaden) liegt der Jahresumsatz von Selbstbräunungsprodukten derzeit bei rund 25 Millionen Euro. Diese Zahl verdeutlicht recht eindrucksvoll den industriellen Nutzen dieser Verbindung.

Dihydroxyaceton wird durch die Umsetzung von Glycerin unter Einsatz von *G. oxydans* hergestellt (RAINBOW & ROSE, 1963; PERLMAN, 1979). Verantwortlich für die Produktion von Dihydroxyaceton ist die membrangebundene, PQQ-abhängige Glycerin-Dehydrogenase aus *G. oxydans* (CLARET et al., 1994), welche von AMEYAMA et al. (1985) aufgereinigt und charakterisiert wurde. Kürzlich stellte sich heraus, dass es sich bei den Chinoproteinen der Glycerin-, Sorbitol- und Arabitol-Dehydrogenase um dieselben Enzyme handelt (MATSUSHITA et al., 2003). In dem Genom von 621H wird dieses Enzym von ORF 2230 und ORF 2231 (*sldA*, *sldB*) kodiert.

Die Fermentation erfolgt in diskontinuierlicher Kultur bei pH 6 und guter Belüftung, um eine ausreichende Versorgung mit Sauerstoff zu gewährleisten (SVITEL & STURDIK, 1994). Die anfängliche Konzentration von Glycerin liegt bei 200 g/l. Die Ausbeute an Dihydroxyaceton beträgt dabei etwa 90 % (CLARET et al., 1992). Ein großes Problem bei diesem fermentativen Prozess ist eine Feedback-Inhibierung: Aufgrund der ansteigenden Konzentration von Dihydroxyaceton wird die weitere Produktion dieser Verbindung gehemmt. Der Grund dafür ist, dass Dihydroxyaceton die Zellen irreversibel schädigt. Hohe Konzentrationen von Dihydroxyaceton inhibieren die Enzyme des Pentosephosphat-Wegs (OHREM & VOSS, 1995), die membrangebundene Glycerin-Dehydrogenase und ein Facilitator-Protein zur erleichterten Diffusion von Glycerin (CLARET et al. 1992). In dem Genom von 621H befinden sich die ORFs, welche für die Glycerin-Dehydrogenase (ORF 0748) und das Facilitator-Protein (ORF 0749) kodieren in dem Glycerin-Operon (s. 3.6.1.1).

Bestrebungen, die Produktion von Dihydroxyaceton zu verbessern und zu optimieren, gehen in die Richtung, immobilisierte Zellen von *G. oxydans* einzusetzen. (HOLST et al., 1985; TKAC et al., 2001). HEKMAT et al. (2003) ermittelten für eine weitere Optimierung des Prozesses die kritische Konzentration von Dihydroxyaceton, die bei 60 kg/m³ lag. Sie entwickelten einen neuen Reaktor, mit dem die Produktivität um 75 % auf 2,8 kg/(m³*h) Dihydroxyaceton erhöht werden konnte. Dieser Reaktor besteht aus einem Schütteltank, kombiniert mit einer durchlässigen Säule, in der sich die Zellen befanden. Durch Immobilisierung der Bakterien an ein hydrophobes Carrier-Material werden Zellen von *G. oxydans* an die Säule gebunden. Nach eigener Einschätzung der Autoren ist aber auch damit die maximale Produktivität noch nicht erreicht.

4.4.5 Produktion von Xylitol

Xylitol ist ein natürlich vorkommender Zucker-Alkohol. Kommerziell wird Xylitol durch die Hydrogenierung von D-Xylulose hergestellt. Da Xylitol einen ähnlich süßen Geschmack aufweist, wie z. B. Saccharose, wird es in der Lebensmittelindustrie oft als alternativer Süßstoff eingesetzt (PIERINI, 2001).

Durch die Nahrung aufgenommene Zucker bieten nach ihrer Hydrolyse durch Amylasen einen Angriffspunkt für die Bakterienpopulation der Mundhöhle. Die Folge davon ist ein Absinken des pH-Wertes. Dieser Vorgang wiederum führt zu einer Demineralisierung der Zähne. Studien haben gezeigt, dass es einen direkten Zusammenhang zwischen der Einnahme von Diät-Zuckern, bzw. Zucker-Ersatzstoffen, wie Xylitol, und der Ausbildung von Karies gibt (TOUGER-DECKER & van LOVEREN, 2003). Feldstudien belegten, dass beim Einsatz von Kaugummis, die Xylitol als Süßungsmittel enthielten, die Ausbildung von Karies präventiv verhindert werden konnte (KOVARI et al., 2003). Auch die Verwendung von Zahnpasta, welche Xylitol und Fluorid enthält, zeigte einen positiven Effekt im Bereich der Mundhygiene (IWATA et al., 2003).

SUZUKI et al. (2002) berichteten über die Produktion von Xylitol aus Arabitol durch *G. oxydans*. Dabei wird Arabitol zunächst durch die membrangebundene Arabitol/Sorbitol-Dehydrogenase zu Xylulose umgesetzt, anschließend erfolgt die Produktion von Xylitol durch die lösliche Xylitol-Dehydrogenase. SUGIYAMA et al. (2003a) reinigten aus der

cytoplasmatischen Fraktion des Stammes 621H die lösliche Xylitol-Dehydrogenase auf. Das Enzym reduzierte D-Xylulose zu Xylitol in Anwesenheit von NADH und bei einem optimalen pH-Wert von 5,0. Im Genom von *G. oxydans* 621H (diese Arbeit) wird das Enzym durch den ORF 2241 gebildet und weist im Einklang mit den Ergebnissen von SUGIYAMA et al (2003a) eine Zusammensetzung von 262 Aminosäuren auf. Mit Hilfe eines rekombinanten *Gluconobacter*-Stammes, der eine plasmidkodierte Kopie der Xylitol-Dehydrogenase trug, konnte die Produktion von Xylitol signifikant gesteigert werden. Während mit dem Wildtyp-Stamm aus 225 g/l Arabitol 27 g/l Xylitol gewonnen werden konnten, produzierte der rekombinante Stamm 57 g/l Xylitol. Ein Problem bei der Überexpression dieser cytoplasmatischen Dehydrogenase ist der Verbrauch des Cofaktors NADH. Nur wenn NADH in ausreichenden Mengen zur Verfügung steht, ist eine Umsetzung der Xylulose zu erwarten (SUGIYAMA et al., 2003a).

Der ORF 2241, der für diese Xylitol-Dehydrogenase in *G. oxydans* 621H kodiert (diese Arbeit), befindet sich innerhalb des Genoms nicht in einem Cluster, sondern liegt einzeln vor. Stromaufwärts des ORFs findet man Gene, die für die Synthese von S-Adenosyl-Methionin kodieren, stromabwärts sind ORFs, deren Genprodukte in die Biosynthese von Lysin involviert sind. Eine Organisation dieses ORFs in einem Cluster oder Operon hätte eventuell Hinweise auf mögliche Regulationsmechanismen bezüglich der Xylitol-Dehydrogenase geben können.

Untersuchungen von SUGIYAMA et al. (2003b) zeigten, dass die Xylitol-Produktion durch die Expression zweier zusätzlicher Proteine weiter gesteigert werden konnte. Eines dieser Enzyme ist ein bifunktionales Enzym, das sowohl eine Transaldolase- als auch eine Glucose-6-Phosphat-Isomerase-Aktivität aufweist. Kodiert wird das Gen im Stamm 621H von ORF 0315. Dieser ORF befindet sich innerhalb des Cluster, welches für die Enzyme des Pentosephosphat-Weges kodiert. Da *G. oxydans* eine Phosphofruktokinase als Schlüsselenzym der Glykolyse fehlt, erscheint es durchaus sinnvoll die Isomerisierung von Fructose-6-Phosphat zu Glucose-6-Phosphat und umgekehrt mit den Enzymen des Pentosephosphat-Weges zu verknüpfen.

Vor dem Hintergrund einer gesteigerten Aktivität der Xylitol-Dehydrogenase, die im Gegensatz zu den Enzymen des Pentosephosphat-Weges NADH-abhängig ist, war

dieser Befund zunächst äußerst erstaunlich. Diesbezüglich konnte jedoch von TONOUCHI et al. (2003) gezeigt werden, dass die Enzyme des Pentosephosphat-Weges in *G. oxydans* eine duale Coenzym-spezifität gegenüber NAD/NADH und NADP/NADPH aufweisen

Bei dem zweiten Enzym, das die Produktion von Xylitol durch *G. oxydans* steigert, handelt es sich um eine Ribulokinase, welche die Reaktion von D-Ribulose zu D-Ribulose-5-phosphat katalysiert (SUGIYAMA et al., 2003b). Ribulose-5-phosphat ist ebenfalls in den Pentosephosphat-Weg involviert. Im allgemeinen gilt der Pentosephosphat-Weg als Quelle für NADPH. Unter der Voraussetzung einer dualen Coenzym-Spezifität der Glucose-6-Phosphat-Dehydrogenase und der 6-Phosphogluconat-Dehydrogenase (TONOUCHI et al., 2003) kann der Pentosephosphat-Weg in *G. oxydans* zusätzlich als Quelle für NADH fungieren.

Es stellt sich die Frage, ob die Produktion von Xylitol aus Arabitol eine neue biotechnologische Anwendungsmöglichkeit für *Gluconobacter*-Stämme sein kann. Der oben beschriebene Einsatz von Xylitol als Zucker-Ersatzstoff deutet zunächst darauf hin (SUGIYAMA et al., 2003a,b). Ein Blick in einen Chemikalienkatalog muss allerdings zu einem anderen Urteil führen: Bei dem Substrat D-Arabitol handelt es sich um eine recht kostspielige Substanz, mit einem Preis von 4550 Euro pro Kilogramm. Der Preis von D-Xylulose und Xylitol beläuft sich auf 215 bzw. 165 Euro pro Kilogramm (SIGMA-ALDRICH Chemie GmbH, Taufkirchen, Chemikalienkatalog 2004-2005). Bei der chemischen Synthese von Xylitol wird dieses durch Hydrogenierung von D-Xylulose hergestellt. Zusammengefasst bedeutet dies, dass eine biotechnologische Produktion von Xylitol nur dann denkbar wäre, wenn Glucose als Ausgangssubstrat eingesetzt werden kann. Hier beläuft sich der Preis auf 39,80 €/kg (SIGMA-ALDRICH Chemie GmbH, Taufkirchen, Chemikalienkatalog 2004-2005). Denkbar wäre ein fermentativer Prozess, bei dem zunächst Glucose zu Arabitol umgesetzt wird, woraus dann mit Hilfe von *G. oxydans* Xylitol hergestellt werden könnte. Diesbezüglich wurde schon 1965 von der Arabitol-Produktion durch *Saccharomyces rouxii* berichtet (INGRAM & WOOD, 1965). Auch in neuerer Zeit gab es Berichte über die Produktion von D-Arabitol durch Mikroorganismen (LUCCA et al., 2002; NOZAKI et al., 2003). Allerdings würde auch ein derartiger zweistufiger Prozess weiterer Optimierung bedürfen. Da bei der rein chemischen Synthese von Xylitol keine

Schwierigkeiten wie z. B. eine aufwendige Schutzgruppenchemie auftauchen, ist es fraglich, ob ein derart etabliertes Verfahren durch eine biotechnologische Produktion ersetzt werden kann.

4.4.6 Produktion von Gluconat und Ketogluconaten durch *G. oxydans*

Die Produktion von Ketogluconaten in *G. oxydans* kann durch zwei voneinander topologisch getrennte Enzymsysteme erfolgen. Eines ist in der Membran lokalisiert, bei dem zweiten Enzymsystem handelt es sich um ein cytoplasmatisches System (AMEYAMA & ADACHI, 1982a, b; MATSUSHITA et al., 1994). Grundsätzlich erfolgt dabei zunächst die Oxidation der Glucose zu Gluconat. Diese wird anschließend zu den entsprechenden Ketogluconaten weiteroxidiert. Eine Übersicht über die Verwertung von Glucose durch *G. oxydans* ist in Abb. 47 wiedergegeben. Dargestellt ist sowohl die Produktion von Ketogluconaten als auch die Einschleusung der Glucose in den Intermediärstoffwechsel zum Aufbau von Zellbestandteilen. Dieses ist am Beispiel der Biosynthese der Aminosäuren wiedergegeben, die in Kapitel 3.6.3 vorgestellt wurde.

Die Produktion von 5-Ketogluconat ist vor allem für die chemische und die Lebensmittelindustrie von Interesse, da die Verbindung ein Vorläufer der L-Weinsäure ist (MATZERATH et al., 1995, KLASSEN et al., 1992). Das Produkt kommt z. B. in der Lebensmittelindustrie zum Einsatz, wo L-Weinsäure als Antioxidans eingesetzt wird. 5-Ketogluconat wird von Essigsäurebakterien beim Wachstum auf Glucose oder Gluconat produziert. Dabei bilden *Gluconobacter*-Stämme jedoch sowohl 5-Ketogluconat als auch 2-Ketogluconat, sowie 2,5-Diketogluconat.

4.4.6.1 Membrangebundene Produktion von Ketogluconaten

Die Oxidation von Gluconat durch die membrangebundenen Enzyme Gluconat-Dehydrogenase und 2-Ketogluconat-Dehydrogenase führt zur Bildung von 2-Ketogluconat und 2,5-Diketogluconat (SHINAGAWA et al., 1981, 1984). Ein Enzym für die Produktion von 2,5-Diketogluconat ist in *G. oxydans* 621H offensichtlich nicht vorhanden, da ein entsprechendes Gen nicht identifiziert werden konnte (diese Arbeit). Es bleibt zu prüfen, ob die heterologe Expression eines entsprechenden Gens in dem

Organismus möglich ist, um 2,5-Diketogluconat als Zwischenprodukt der Vitamin C-Synthese für industrielle Anwendungen zu bilden (siehe 4.4.2).

Im Genom des Stammes 621H wird die Gluconat-Dehydrogenase, die Gluconat zu 2-Ketogluconat umsetzt, von den ORFs 2631 (Cytochrom c-Untereinheit), 2632 (alpha-Untereinheit) und 2633 (gamma-Untereinheit) kodiert. Neben der Produktion von 2-Ketogluconat durch membrangebundene Enzyme wurde jedoch ebenso beobachtet, dass Gluconat zu 5-Ketogluconat umgesetzt werden kann (ADACHI et al., 2001). In der Vergangenheit wurden immer wieder Proteine aufgereinigt, biochemisch charakterisiert und nach der entsprechenden Substratspezifität benannt. Leider fehlte jedoch meistens eine Aminosäuresequenz. Dies führte dazu, dass ein und dasselbe Enzym mehrmals unter verschiedenen Namen beschrieben.

ADACHI et al. (2001) zeigten, dass die Arabitol/Sorbitol-Dehydrogenase eine Oxidation von Gluconat katalysiert, wobei als einziges Produkt 5-Ketogluconat entsteht. Dieses Protein war zuvor unter zwei verschiedenen Bezeichnungen publiziert worden (ADACHI et al., 2001; SUGISAWA & HOSHINO, 2002) Bestätigt wurde dieses Resultat von MATSUSHITA et al., 2003, die zeigten, dass Mutanten von *G. suboxydans* IFO 3257 mit einer Deletion im Sorbitol/Arabitol-Dehydrogenase-Gen beim Wachstum auf Gluconat kein 5-Ketogluconat mehr bildeten. 2-Ketogluconat wurde jedoch weiterhin produziert. Im Genom von 621H wird die beschriebene Dehydrogenase von *sldA* und *sldB* kodiert (ORF 2230, 2231).

Kürzlich wurde aus *G. suboxydans* IFO 12528 ein Chinoprotein isoliert, welches ebenfalls 5-Ketogluconat produziert. Bezeichnet wurde diese neue Dehydrogenase als Gluconat/Polyol-Dehydrogenase. Die Substratspezifität dieser Gluconat-5-Dehydrogenase war sehr breit und umfasste Polyole, Aminoderivate und sekundäre Alkohole. Die heterologe Expression dieser Dehydrogenase in *E. coli* resultierte bei der Umsetzung von Gluconat zu 5-Ketogluconat in einer 75 %igen Ausbeute (SALUSJARVI et al., 2004). Die Sequenz des Proteins ist unter der EBI (European Bioinformatics Institute) Accession-Nummer AJ577472 einzusehen. Ein Abgleich der dort publizierten Sequenz mit dem Genom des Stammes 621H zeigte, dass es sich bei diesem Protein um die bereits bekannte Sorbitol Dehydrogenase handelt, kodiert von *sldAB* (ORF 2230, 2231). Diese Ergebnisse verdeutlichen, wie wichtig es ist, die

entsprechenden Enzyme nicht nur biochemisch zu charakterisieren, sondern auch die Aminosäuresequenz zu bestimmen. Auf der anderen Seite zeigen die Ergebnisse aber auch eindrucksvoll, wie vielseitig die PQQ-abhängigen Enzyme aus *Gluconobacter*-Stämmen sind. Für die Untersuchung der bislang nicht-charakterisierten Dehydrogenasen stellt das Genom des in dieser Arbeit sequenzierten Organismus eine gute Grundlage dar. Die Daten ermöglichen eine Klonierung und Expression der Gene und eine Überproduktion der entsprechenden Proteine, um anschließend eine biochemische Charakterisierung vorzunehmen und das Substratspektrum zu ermitteln.

4.4.6.2 Cytoplasmatische Produktion von Ketogluconaten

Wie oben schon erwähnt weist *G. oxydans* neben dem membrangebundenen Enzymsystem auch ein cytoplasmatisches System für die Produktion von Ketogluconaten auf. Die cytoplasmatische Gluconat:NADP-5-Oxidoreduktase setzt Gluconat zu 5-Ketogluconat um (KLASEN et al., 1995). Im Genom von 621H umfasst die lösliche Gluconat-Dehydrogenase, kodiert von ORF 864, 256 Aminosäuren. Stromaufwärts von diesem Gen befindet sich eine Permease für die Aufnahme von Gluconat (ORF 865).

Intrazellulär wird Gluconat nicht nur zu Ketogluconat oxidiert, sondern auch in den intrazellulären Stoffwechsel eingebracht. (Abb. 47). Es wurde postuliert, dass die cytoplasmatische 5-Ketogluconat-Reduktase für die Regeneration von NADP zuständig ist (ADACHI et al., 1979b). Hierbei setzt das Enzym 5-Ketogluconat zu Gluconat um, wobei NADPH oxidiert wird. Diese Hypothese steht jedoch im Widerspruch zu der im Genom von *G. oxydans* 621H identifizierten Transhydrogenase (ORF 1600, 1601, 1602). Hinzu kommt, dass TONOUCHI et al. (2003) gezeigt haben, dass die Enzyme des Pentosephosphat-Weges sowohl eine Bindestelle für NAD als auch für NADP aufweisen. Dieses würde eine Anhäufung von NADPH zumindest minimieren. Da aber die Regulation der Gene der Transhydrogenase und der 5-Ketogluconat-Dehydrogenase nicht geklärt ist, kann diese Vermutung hier weder bestätigt noch widerlegt werden. WEENK et al. (1984) beobachteten, dass beim Wachstum von *Gluconobacter*-Spezies auf Glucose, das Substrat zunächst fast quantitativ zu Gluconat umgesetzt wird. Erst wenn die Glucose-Konzentration unter einen bestimmten Schwellenwert sinkt, tritt die Produktion von Ketogluconaten auf. Da

die entsprechenden Enzyme konstitutiv vorliegen (DE LEY & DOCHY, 1960; DE LEY & STOUTHAMER, 1959), wurden andere Mechanismen bezüglich der Regulation der Ketogluconat-Produktion vorgeschlagen (s. 3.6.4; Abb. 46). Das sogenannte Gluconat-Recycling-Modell besagt, dass die Produktion von Ketogluconaten durch die membrangebundene Gluconat-Dehydrogenase erfolgt. Die lösliche Ketogluconat-Reduktase setzt diese dann wieder zu Gluconat um (OLIJVE & KOK, 1979). Mit diesem Modell versuchte man die initiale Akkumulierung von Gluconat zu erklären. LEVERING et al. (1988) zeigten, dass die Aktivität der membrangebundenen Glucose-Dehydrogenase bei Wachstum auf Glucose in zellfreien Extrakten deutlich höher war, als die des entsprechenden löslichen Enzyms. Dieser Befund ist in sofern nicht weiter erstaunlich, da die Energiegewinnung von *Gluconobacter* ausschließlich durch membrangebundene Enzyme erfolgt. Die Elektronen der membrangebundenen Dehydrogenierungsreaktionen fließen direkt in die Atmungskette ein. Daraus folgt, dass die Aufnahme von Substraten ausschließlich zum Aufbau von Metaboliten für den intrazellulären Stoffwechsel erfolgt.

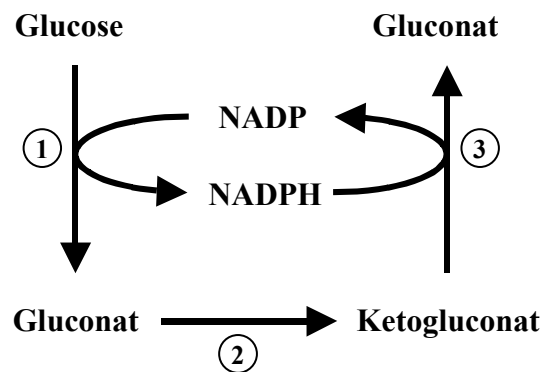


Abb. 46: Gluconat-Recycling Modell nach LEVERING et al., 1988 (mod.)

Die Zahlen beziehen sich auf die entsprechenden Enzyme: 1: lösliche Glucose-Dehydrogenase; 2: membrangebundene Gluconat-Dehydrogenase; 3: Ketogluconat-Reduktase.

Für die Produktion der L-Weinsäure ist vor allem die Synthese von 5-Ketogluconat interessant, da diese Verbindung chemisch einfach zu Tartrat umgesetzt werden kann. Im allgemeinen eignen sich hierzu membrangebundene Prozesse, bei denen das Substrat unvollständig zu dem gewünschten Produkt oxidiert wird. Zum einen stellt sich die

Aufreinigung des Produkts einfach da, zum anderen ist jedoch auch zu erwarten, dass die Ausbeute höher ist. Bei einer cytoplasmatischen Umsetzung der Substrate kommt es zur Einschleusung der Metabolite in den Intermediärstoffwechsel und zur Synthese von Nebenprodukten. Dies bedeutet, dass pro Mol eingesetzten Substrat weniger Produkt entsteht als bei membrangebundenen Prozessen. Da die Produktion von 5-Ketogluconat von *G. oxydans* durch membrangebundene Enzyme erst in neuerer Zeit bekannt wurde (ADACHI et al, 2001), gab und gibt es auch Versuche 5-Ketogluconat durch cytoplasmatische Prozesse vermehrt herzustellen. Diese Vorgehensweise ist nachfolgend beschrieben.

KLASEN et al. (1995) zeigten, dass die 5-Ketogluconat-Dehydrogenase aus *G. oxydans* spezifisch für Gluconat und 5-Ketogluconat ist. Es wurden weder Glucose, Sorbitol oder Mannitol oxidiert noch 2-Ketogluconat reduziert. Darüber hinaus ist dieses Enzym abhängig von NADP; mit NAD als Cofaktor war das Enzym nahezu inaktiv. Im Gegensatz zu den Chinoproteinen, welche die membrangebundene Produktion von Ketogluconaten katalysieren, weist das korrespondierende cytoplasmatische Enzym also nur ein enges, spezifisches Substratspektrum auf. Um eine Akkumulierung von 5-Ketogluconat zu erreichen, exprimierten HERMANN et al., (2003) die Gluconat:NADP-5-Oxidoreduktase über. Für die Überexpression wurde ein rekombinanter Stamm von 621H eingesetzt. Dieser trug auf einem Plasmid neben dem Gen für die Gluconat-Dehydrogenase auch das Gen *sthA* aus *E. coli*. Dieses Gen kodiert für eine cytoplasmatische Transhydrogenase und sollte gewährleisten, dass eine Oxidation von NADPH bzw. eine Regeneration von NADP erfolgt. Mit dem rekombinanten Stamm konnten 18,7 mM 5-Ketogluconat gebildet werden, während der Wildtyp nur 15,5 mM 5-Ketogluconat produzierte. Die Produktion von 2-Ketogluconat konnte jedoch nicht ausgeschaltet werden, allerdings blieb die Menge des gebildeten 2-Ketogluconats im Vergleich zum Wildtyp konstant.

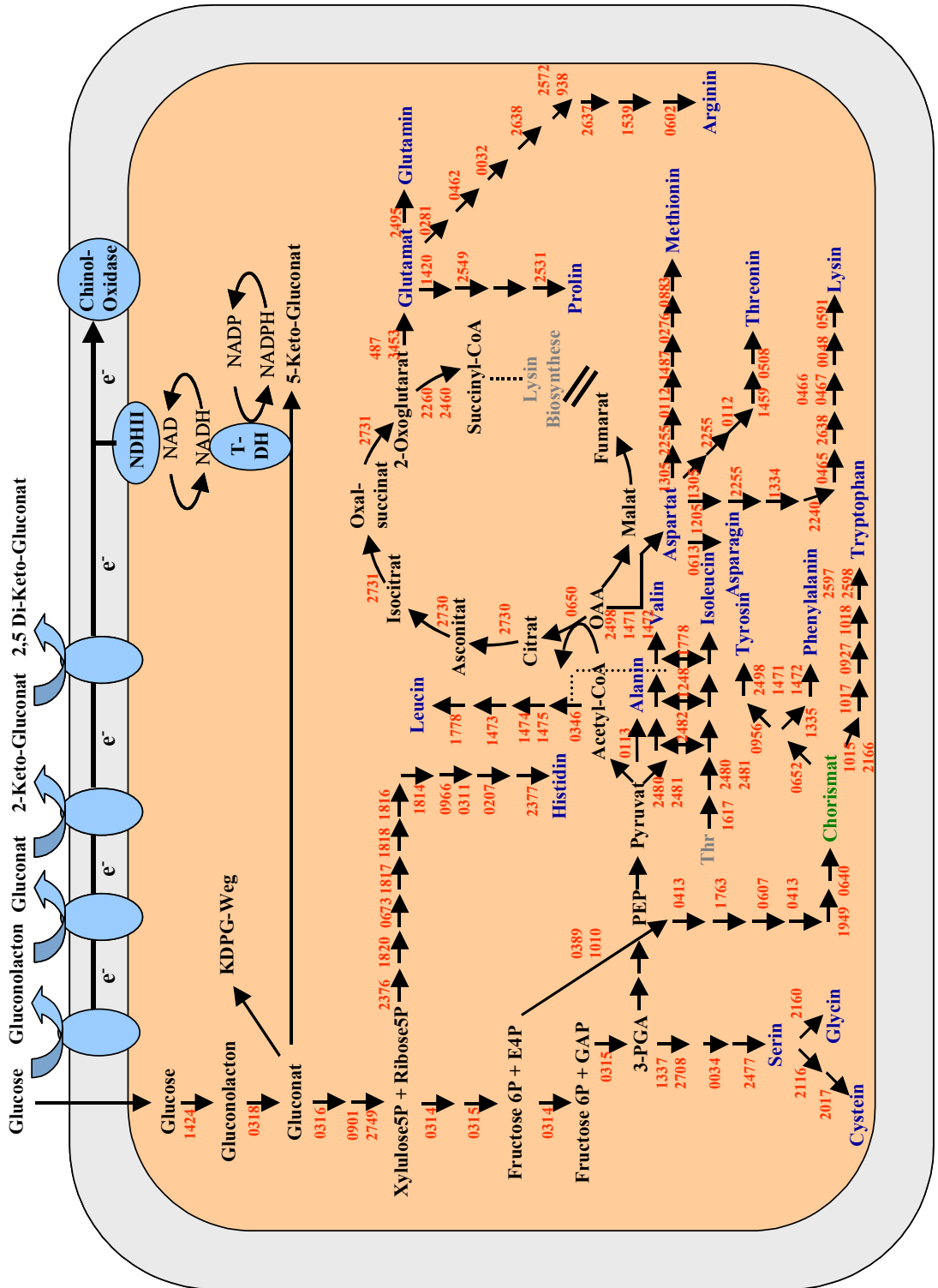


Abb. 47 (vorhergehende Seite): Übersicht über den Glucose Metabolismus in *G. oxydans*

Dargestellt ist die Bildung von Gluconat und Ketogluconaten durch *G. oxydans* sowie die Verstoffwechslung der Glucose für den intrazellulären Metabolismus. Dieser ist exemplarisch am Beispiel der Aminosäuresynthese dargestellt. Neben den gezeigten membrangebundenen Enzymen enthält das Genom eine Vielzahl weiterer membrangebundener Dehydrogenasen, die hier aus Gründen der Übersicht nicht gezeigt sind. Die roten Nummern an den Pfeilen geben jeweils die entsprechende ORF-Nummer an. Abk.: -P: phosphat; NDHII: NADH-Dehydrogenase von Typ II; T-DH: Transhydrogenase.

4.5 Ausblick

Die Entschlüsselung des Genoms von *G. oxydans* stellt ein Fundament für weitere Untersuchungen dar. Im Institut für Mikrobiologie und Genetik der Universität Göttingen ist die Herstellung eines DNA-Microarray geplant, der das gesamte Genom von *G. oxydans* umfasst. Dabei soll auf der Grundlage der in dieser Arbeit vorgestellten genomischen Sequenz jeder ORF mittels PCR amplifiziert werden, um anschließend eine genomweite Transkriptionsanalyse durchführen zu können. Die Anwendung der DNA-Chip Technologie stellt eine Möglichkeit dar, die Expressionslevel der einzelnen Dehydrogenasen unter verschiedenen Bedingungen zu untersuchen. Auch kann mit einem solchen Versuch gezeigt werden, ob und welche ORFs unter unterschiedlichen Bedingungen transkribiert werden. Darüber hinaus werden Hinweise auf intrazelluläre Regulationsmechanismen gewonnen.

Ebenfalls auf der Grundlage der hier vorgestellten genomischen Sequenz von *G. oxydans* können einzelne Dehydrogenasen gezielt kloniert, exprimiert und aufgereinigt werden. Dieses Vorgehen ermöglicht, bestimmte Enzyme zu charakterisieren und ihr Substratspektrum zu bestimmen. Dabei erhofft man sich, neue Einsatzmöglichkeiten für *G. oxydans* in der Biotechnologie zu finden.

5 ZUSAMMENFASSUNG

1. Im Rahmen dieser Arbeit wurde das Genom von *Gluconobacter oxydans* 621H vollständig sequenziert. Das Chromosom hat eine Größe von 2,7 Mb. Nach Abschluss der Sequenzierarbeiten betrug die Fehlerrate in der Konsensussequenz 1/289.582.
2. Neben dem Chromosom wurde in dem Genom von *G. oxydans* ein 163 kb großes zirkuläres Megaplasmid (pGOX1) gefunden. Darüber hinaus konnten vier weitere zirkuläre Plasmide, pGOX2 (26,6 kb), pGOX3 (14,5 kb), pGOX4 (13,2 kb) und pGOX5 (2,7 kb) identifiziert werden.
3. Auf dem Megaplasmid pGOX1 wurden 175 offene Leserahmen (ORFs) identifiziert. Davon konnten 55 ORFs (31 %) eine Funktion zugewiesen werden. In einem Gencluster wurden zahlreiche ORFs gefunden, die Ähnlichkeiten zu Genen aufwiesen, die für Proteine eines konjugativen DNA-Transfersystems kodieren. Dieses sind erste Hinweise darauf, dass *G. oxydans* 621H genetisches Material mittels Konjugation übertragen kann.
4. Auf dem 26,5 kb großen Plasmid pGOX2 wurden 18 ORFs identifiziert, von denen 13 eine (potentielle) Funktion zugeordnet werden konnte. Die Plasmide pGOX3 und pGOX4 enthielten Informationen für 18 ORFs bzw. 17 ORFs. Das Plasmid pGOX5 kodierte für fünf ORFs. Aufgrund der geringen Größe von nur 2,7 kb könnte dieses Plasmid die Grundlage zur Entwicklung eines neuen Klonierungsvektors für *G. oxydans* darstellen.
5. Für die Aufnahme von Substraten konnten in *G. oxydans* unterschiedliche Transportsysteme identifiziert werden. Darunter waren verschiedene ABC-Transporter, Permeasen, Symporter, Antiporter und ein Facilitator-Protein für die Aufnahme von Glycerin. In dem Genom wurde kein funktionstüchtiges Phosphotransferase-System (PTS-System) gefunden. Einzelne Komponenten des PTS-Systems lagen zwar vor, jedoch fehlten die Zucker-spezifischen Enzyme.
6. *G. oxydans* fehlt ein Gen, das für die 6-Phosphofruktokinase als Schlüsselenzym der Glykolyse, kodiert. Es wurde weiterhin festgestellt, dass der ORF 0315 für ein bifunktionales Enzym mit Glucose-6-Phosphat-Isomerase- und Transaldolase-

Aktivität kodiert. Die Analysen stehen in Übereinstimmung mit bereits publizierten biochemischen Daten, dass *G. oxydans* Glucose nicht über die Glykolyse sondern über den oxidativen Pentosephosphat-Weg verstoffwechselt.

7. In dem Genom von *G. oxydans* konnten keine Gene identifiziert werden, die für Enzyme zur Synthese von Phosphoenolpyruvat aus Pyruvat kodieren. Damit fehlt diesem Organismus nicht nur das Schlüsselenzym der Glykolyse sondern auch das der Gluconeogenese.
8. Demgegenüber besitzt der Organismus eine vollständige genetische Ausstattung für den oxidativen Pentosephosphat- und den 2-Keto-3-Desoxy-6-Phosphogluconat-Weg. Die intrazelluläre Verwertung von Substraten durch diese Stoffwechselwege dient nicht dem Energiegewinn, sondern lediglich dem Aufbau von Intermediaten zur Biosynthese von Zellmaterial.
9. *G. oxydans* ist nicht im Besitz eines vollständigen Zitratzyklus. Es konnten keine Gene identifiziert werden, die für die Succinat-Dehydrogenase kodieren. Ein unvollständiger Zitratzyklus impliziert, dass Substrate nicht vollständig oxidiert werden, da AcetylCoA nicht zu CO₂ aufoxidiert werden kann. Diese Einschränkung im Stoffwechsel ist sicherlich ein Hauptgrund für die unvollständigen Oxidationsreaktionen, die für *G. oxydans* charakteristisch sind.
10. Es konnte gezeigt werden, dass *G. oxydans* das genetische Potential zur Synthese sämtlicher Aminosäuren hat.
11. Die Rekonstruktion der Atmungskette von *G. oxydans* zeigte, dass der Organismus nur eine begrenzte Anzahl von Enzymen im Kernbereich des membrangebundenen Elektronentransportsystems aufweist. So fanden sich je eine Chinol-Oxidase vom *bo₃*- und *bd*-Typ sowie eine nicht-protonentranslozierende NADH-Dehydrogenase (NADH-DH Typ II). Es fehlten wichtige Schlüsselenzyme wie die Cytochrom c-Oxidase, der Komplex I (NADH-DH Typ I) und eine Succinat-Dehydrogenase. Die Möglichkeiten zur Protonentranslokation sind für diesen Organismus daher sehr begrenzt. Die Energiegewinnung erfolgt durch membrangebundene Dehydrogenierungsreaktionen von Alkoholen und Zuckern. Die Reduktionsäquivalente fließen dabei über den Ubichinon-Pool auf die entsprechenden Chinol-Oxidasen. Die Annotation zeigte weiterhin, dass Gene vorhanden sind, die für einen *bc₁*-Komplex kodieren. Die Funktion dieses Komplexes ist nicht geklärt, da *G. oxydans* keine Cytochrom c-Oxidase besitzt.

12. Der Organismus enthält eine Vielzahl von membrangebundenen Dehydrogenasen, die bereits aus verwandten Arten beschrieben wurden. Diese Enzyme sind verantwortlich für die unvollständige Oxidation von Zuckern und Alkoholen, für die *G. oxydans* charakteristisch ist. Darüber hinaus konnten im Genom mehr als 70 bislang nicht beschriebene Oxidoreduktasen identifiziert werden. Eine Analyse des Substratspektrums dieser Enzyme stellt ein großes Potential für neue biotechnologische Entwicklungen dar. Aufgrund der in dieser Arbeit erzielten Resultate können diese Enzyme nun gezielt produziert und charakterisiert werden.

6 LITERATURVERZEICHNIS

Adachi, O., Fujii, Y., Ghaly, M. F., Toyoma, H., Shinagawa, E., Matsushita, K., 2001. Membrane-bound quinoprotein D-arabitol dehydrogenase of *Gluconobacter suboxydans* IFO 3257: a versatile enzyme for the oxidative fermentation of various ketoses. *Biosci. Biotechnol. Biochem.* **65**: 2755-2762

Adachi, O., Matsushita, K., Shinagawa, E., Ameyama, M., 1979. Occurrence of old yellow enzyme in *Gluconobacter suboxydans*, and the cyclic regeneration of NADP. *J. Biochem.* **886**: 669-709

Adachi, O., Moonmangmee, D., Shinagawa, E., Toyoma, H., Yamada, M., Matsushita, K., 2003b. New quinoproteins in oxidative fermentation. *Biochim. Biophys. Acta* **164**: 10-17

Adachi, O., Moonmangmee, D., Toyoma, H., Yamada, M., Shinagawa, E., Matsushita, K., 2003a. New developments in oxidative fermentation. *Appl. Microbiol. Biotechnol.* **60**: 643-653

Adachi, O., Shinagawa, E., Matsushita, K., Ameyama, M., 1979b. Crystallization and properties of 5-keto-D-gluconate-reductase from *Gluconobacter suboxydans*. *Agric. Boil. Chem.* **43**: 75-83

Adachi, O., Tanasupawat, S., Yoshihara, N., Toyama, H., Matsushita, K., 2003d. 3-Dehydroquinone production by oxidative fermentation and further conversion to the intermediates in the shikimate pathway. *Biosci. Biotechnol. Biochem.* **67**: 2124-2131

Adachi, O., Yoshihara, N., Tanasupawat, S., Toyama, H., Matsushita, K., 2003c. Purification and characterization of membrane-bound quinoprotein quinone dehydrogenase. *Biosci. Biotechnol. Biochem.* **67**: 2115-2123

Aiguo, J. & Peiji, G., 1998. Synthesis of 2-keto-gulonic acid from gluconic acid by co-immobilized *Gluconobacter oxydans* and *Corynebacterium* sp. *Biotechnol. Lett.* **20**:

939-942

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410

Ameyama, M. & Adachi, O., 1982a. 5-Keto-D-gluconate reductase from *Gluconobacter suboxydans*. *Meth. Enzymol.* **89**: 198-202

Ameyama, M. & Adachi, O., 1982b. 2-Ketogluconate reductase from acetic acid bacteria. *Meth. Enzymol.* **89**: 203-210

Ameyama, M., Shinagawa, E., Matsushita, K. and Adachi, O., 1985. Solubilization, purification and properties of membrane-bound glycerol dehydrogenase from *Gluconobacter industrius*. *Agric. Biol. Chem.*: **49** 1001-1010

Ameyama, M., Shinagawa, E., Matsushita, K., Adachi, O., 1981. D-Glucose dehydrogenase of *Gluconobacter suboxydans*: solubilisation, purification and characterization. *Agric. Biol. Chem.* **45**: 851-861

Anderson, S., Berman, C. M., Lazarus, R., Miller, J., Stafford, K., Seymour, J., Light, D., Rastetter, W., Estell, D., 1985. Production of 2-keto-L-gulonate, an intermediate in L-ascorbate synthesis by a genetically modified *Erwinia herbicola*. *Science* **230**: 144-149

Anraku, Y. & Gennis, R. B., 1987. The aerobic respiratory chain of *Escherichia coli*. *Trends Biochem. Sci.* **12**: 262-266

Asai, T., 1968. Acetic acid bacteria: classification and biochemical activities. University of Tokyo Press, Tokyo

Bachmann, B. J., 1990. Linkage map of *Escherichia coli* K12. Edition 8. *Microbiol. Rev.* **54**: 130-197

Badger, J. H. & Olsen, G. J., 1999. CRITICA: coding region identification tool

invoking comparative analysis. *Mol. Biol. Evol.* **16**: 512-524

Bartosik, D., Szymanik, M., Wysocka, E., 2001. Identification of the partitioning site within the repABC-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. *J. Bacteriol.* **183**: 6234-6243

Bauchop, T. & Elsdén, S. R., 1960. The growth of microorganisms in relation to their energy supply. *J. Gen. Microbiol.* **23**: 457-469

Bertsova, Y., v. Bogachev, A., Skulachev, V., P., 2001. Noncoupled NADH:ubiquinone oxidoreductase of *Azotobacter vinelandii* is required for diazotrophic growth at high oxygen concentrations. *J. Bacteriol.* **183**: 6869-6874

Bizouarn, T., Althage, M., Pedersen, A., Tigerstrom, A., Karlsson, J., Johansson, C., Rydstrom, J., 2002. The organization of the membrane domain and its interaction with the NADP(H)-binding site in proton-translocating transhydrogenase from *E. coli*. *Biochim. Biophys. Acta* **1555**: 122-127

Bizouarn, T., Diggle, C., Quirk, P. G., Grimley, R. L., Cotton, N. P., J., Thomas, C. M., Jackson, J. B., 1996. Interaction of nucleotides with the NAD(H)-binding domain of the proton-translocating transhydrogenase of *Rhodospirillum rubrum*. *J. Biol. Chem.* **271**: 10103-10108

Bonfield, J. K. & Staden, R. 1996. Experiment files and their application during large-scale sequencing projects. *DNA Sequence* **6**: 109-117

Boonstra, B., French, C. E., Wainwright I., Bruce N. C., 1999. The *udhA* gene of *Escherichia coli* encodes a soluble pyridine nucleotide transhydrogenase. *J. Bacteriol.* **181**: 1030-1034

Boudrant, J., 1990. Microbiol processes for ascorbic acid biosynthesis: a review. *Enzyme Microb. Technol.* **12**: 322-329

Buell, C. R., Joardar, V., Lindeberg, M., Selengut, J., Paulsen, I. T., Gwinn, M. L.,

Dodson, R. J., Deboy, R. T., Durkin, A. S., Kolonay, J. F., Madupu, R., Daugherty, S., Brinkac, L., Beanan, M. J., Haft, D. H., Nelson, W. C., Davidsen, T., Zafar, N., Zhou, L., Liu, J., Yuan, Q., Khouri, H., Fedorova, N., Tran, B., Russell, D., Berry, K., Utterback, T., Van Aken, S. E., Feldblyum, T. V., D'Ascenzo, M., Deng, W. L., Ramos, A. R., Alfano, J. R., Cartinhour, S., Chatterjee, A. K., Delaney, T. P., Lazarowitz, S. G., Martin, G. B., Schneider, D. J., Tang, X., Bender, C. L., White, O., Fraser, C. M., Collmer, A., 2003. The complete genome sequence of the *Arabidopsis* and tomato pathogen *Pseudomonas syringae* pv. tomato DC3000. Proc. Natl. Acad. Sci. U.S.A. **100**: 10181-10186

Calhoun, M. W., Oden, K. L., Gennis, R. B., de Mattos, M., J., T., Neijssel, O. M., 1993. Energetic efficiency of *Escherichia coli*: effects of mutations in components of the aerobic respiratory chain. J. Bacteriol. **175**: 3020-3025

Campbell, L. K., Baker, D. E., Campbell, R. K., 2000. Miglitol: assessment of its role in the treatment of patients with diabetes mellitus. Ann. Pharmacother **34**: 1291-1301

Chester, N. & Marshak, D.R., 1993. Dimethyl sulfoxid-mediated primer T_m reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. Anal. Biochem. **209**: 284-290

Choi, E. S., Lee, E. H., Rhee, S. K., 1995. Purification of a membrane-bound sorbitol-dehydrogenase from *Gluconobacter suboxydans*. FEMS Microbiol. Lett. **125**: 45-50

Claret, C., Bories, A., Soucaille, P., 1992. Glycerol inhibition of growth and dihydroxyacetone production by *Gluconobacter oxydans*. Curr. Microbiol. **25**: 149-155

Claret, C., Salmon, J. M., Romieu, C., Bories, A., 1994. Physiology of *Gluconobacter oxydans* during dihydroxyacetone production from glycerol. Appl. Microbiol. Biotechnol. **41**: 359-365

Clarke, D. M., Loo, T. W., Gillam, S., Bragg, P.D., 1986. Nucleotide sequence of the pntA and pntB genes encoding the pyridine nucleotide transhydrogenase of *Escherichia coli*. Eur. J. Biochem. **158**: 647-653

Cleton-Jansen, A.-M., Dekker, S., van de Putte, P., Goosen, N., 1991. A single amino acid substitution changes the substrate specificity of quinoprotein glucose dehydrogenase in *Gluconobacter oxydans*. *Mol. Gen. Genet.* **229**: 206-212

De Ley, J & Stouthamer, A. H., 1959. The mechanism and localization of hexonate metabolism in *Acetobacter suboxydans* and *Acetobacter melanogenum*. *Biochim. Biophys. Acta* **34**: 171-183

De Ley, J., Dochy, R., 1960. On the localization of oxidase systems in *Acetobacter* cells. *Biochim. Biophys. Acta* **40**: 227-289

De Ley, J., Swings, J., Gossele, F., 1984. The genus *Gluconobacter*. In: Krieg, N. R. & Holt, J. G. (Ed.), *Bergey's manual of systematic bacteriology*, vol. 1. The Williams & Wilkins Co., Baltimore, S. 267-278

De Wulf, P., Soetaert, W., Vandamme, E. J., 2000. Optimized synthesis of L-sorbose by C(5)-dehydrogenation of D-sorbitol with *Gluconobacter oxydans*. *Biotechnol. Bioeng.* **69**: 339-343

Deppenmeier, U., Hoffmeister, M., Prust, C., 2002. Biochemistry and biotechnological applications of *Gluconobacter* strains. *Appl. Microbiol. Biotechnol.* **59**: 1513-1533

Dugaiczky, A., Boyer, H. W., Goodman, H. M., 1975. Ligation of *EcoRI* endonuclease-generated DNA fragments into linear and circular structures. *J. Mol. Biol.* **96**: 171-184

Eagon, R. G., 1963. Rate limiting effects of pyridine nucleotides on carbohydrate catabolic pathways of microorganisms. *Biochem. Biophys. Res. Comm.* **12**: 274-279

Elsemore, D. A. & Ornston, L. N., 1994. The *pca-pob* supraoperonic cluster of *Acinetobacter calcoaceticus* contains *quiA*, the structural gene for quinate-shikimate dehydrogenase. *J. Bacteriol.* **176**: 7659-7666

Elsemore, D. A. & Ornston, L. N., 1995. Unusual ancestry of dehydratases associated with quinate catabolism in *Acinetobacter calcoaceticus*. *J. Bacteriol.* **177**: 5971-5978

Ewing B., Hillier L., Wendl M.C., Green P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**: 175-85

Fani, R., Bazzicalupo, M., Damiani, G., Bianchi, A., Schipani, C., Sgaramella, V., Polsinelli, M., 1989. Cloning of histidine genes of *Azospirillum brasilense*: organization of the ABFH gene cluster and nucleotide sequence of the *hisB* gene. *Mol. Gen. Genet.* **216**: 224-229

Felder, M., Gupta, A., Verma, V., Kumar, A., Qazi, G. N., Cullum, J., 2000. The pyrroloquinoline quinone synthesis genes of *Gluconobacter oxydans*. *FEMS Microbiol. Lett.* **193**: 231-236

Fiedler, S. & Wirth, R., 1988: Transformation of bacteria with plasmid DNA by electroporation. *Anal. Biochem.* **170**: 38-44

Fluckiger, R., Woodtli, T., Gallop, P. M., 1988. The interaction of aminogroups with pyrroloquinoline quinone as detected by the reduction of nitroblue tetrazolium. *Biochem. Biophys. Res. Commun.* **153**: 353-358

Frank, A. C., Amiri, H., Andersson, S. G. E., 2002. Genome deterioration: loss of repeated sequences and accumulation of junk DNA. *Genetica* **115**: 1-12

Gibson, T., J., Rosenthal, A., Waterson, R., H., 1987. Lorist6, a cosmid vector with *Bam*HI, *Not*I, *Sca*I and *Hind*III cloning sites and altered neomycin phosphotransferase gene expression. *Gene* **53**: 283-286

Gillis, M. & De Ley, J., 1980. Intra-and intergenic similarities of the ribosomal ribonucleic acid cistrons of *Acetobacter* and *Gluconbacter*. *Int. J. Syst. Bact.* **30**: 7-27

Gordon, D., Abajian, C., Green, P., 1998: Consed: A graphical tool for sequence

finishing. *Genome Res.* **8**: 195-202

Greenfield, S. & Claus, G. W., 1972. Non-functional tricarboxic acid cycle and the mechanism of glutamate biosynthesis in *Acetobacter suboxydans*. *J. Bacteriol.* **112**: 1295-1301

Greenshields, R. N., 1978. Acetic acid: vinegar. In: Rose A. H. (Ed.) *Economic Microbiology*, vol. 2, Academic press, New york, pp 121-186

Grigoriev, A., 1998. Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* **26**: 2286-2290

Guo, F. B., Ou, H. Y., Zhang, C. T., 2003. ZCURVE: a new system for recognizing protein-coding genes in bacterial and archaeal genomes. *Nucleic Acids Res.* **31**: 1780-1789

Gupta, A., Singh, V. K., Qazi, G. N., Kumar. A., 2001. *Gluconobacter oxydans*: Its biotechnological applications. *J. Mol. Microbiol. Biotechnol.* **3**: 445-456.

Hancock, R. D. & Viola, R., 2002. Biotechnological approaches for L-ascorbic acid production. *Trends Biotechnol.* **20**: 299-305

Hawley, D. K. & Mc Clure, W. R., 1983. Compilation and analysis of *the Escherichia coli* promotor DNA sequences. *Nucl. Acids Res.* **11**: 2237-2255

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

Herrmann, U., Merfort, M., Jeude, M., Bringer-Meyer, S., Sahn, H., 2003. Biotransformation of glucose to 5-keto-D-gluconic acid by recombinatn *Gluconobacter oxydans* DSM 2343. *Appl. Microbiol. Biotechnol.* **64**: 86-90

Hoek, J. B. & Rydstrom, J., 1988. Physiological roles of nicotinamide nucleotide

transhydrogenase. *Biochem. J.* **254**: 1-10

Hofman, K. & Stoffel, W., 1993. A database of membrane spanning protein segments. *Biol. Chem. Hoppe-Seyler* **374**: 166

Holst, O., Lundbäck, H., Mattiasson, B., 1985. Hydrogen peroxide as an oxygen source for immobilized *Gluconobacter oxydans* converting glycerol to dihydroxyacetone. *Appl. Microbiol. Biotechnol.* **22**: 383-388

Hoshino, T., Sugisawa, T., Shinjoh, M., Tomiyama, N., Miyazaki, T., 2003. Membrane-bound D-sorbitol dehydrogenase of *Gluconobacter suboxydans* IFO 3255 – enzymatic and genetic characterization. *Biochim. Biophys. Acta.* **1647**: 278-288

Houck, D. R., Hanners, J. L., Unkefer, C. J., 1991. Biosynthesis of pyrroloquinoline quinone. 2. Biosynthetic assembly from glutamate and tyrosine. *J. Am. Chem. Soc.* **113**: 3162-3166

Hunte, C., Palsdottir, H., Trumpower, B. L., 2003. Protonmotive pathways and mechanisms in the cytochrome bc₁ complex. *FEBS Lett.* **545**: 39-46

Ingram, J. M. & Wood, W. A., 1965. Enzymatic basis for D-arabitol production by *Saccharomyces rouxii*. *J. Bacteriol.* **89**: 1186-1194

Inoue, H., Nojima, H., Okayama, H., 1990. High efficiency transformation of *E. coli* with plasmids. *Gene* **96**: 23-28

Iwata, C., Nakagaki, H., Morita, I., Sekiya, T., Goshima, M., Abe, T., Isogai, A., Hanaki, M., Kuwahara, M., Tatematsu, M., Robinson, C., 2003. Daily use of dentifrice with and without xylitol and fluoride: effect on glucose retention in humans *in vivo*. *Arch. Oral Biol.* **48**: 389-395

Jackson, B. J., 2003. Proton translocation by transhydrogenase. *FEBS Lett.* **545**: 18-24

Jaworowski, A., Campbell, H. D., Poulis, M. I., Young, I. G., 1981. Genetic

identification and purification of the respiratory NADH dehydrogenase of *Escherichia coli*. *Biochemistry* **20**: 2041-2047

Ji, A. & Gao, P., 2001. Substrate selectivity of *Gluconobacter oxydans* for production of 2,5-diketo-D-gulonic acid and synthesis of 2-keto-L-gulonic acid in a multienzyme system. *Appl. Biochem. Biotechnol.* **94**: 213-223

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., Watanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A., Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, M., Takeuchi, C., Yamada, M., Tabata, S., 2000. Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res.* **7**: 331-338

Kaneko, T., Nakamura, Y., Sato, S., Minamisawa, K., Uchiumi, T., Sasamoto, S., Watanabe, A., Idesawa, K., Iriguchi, M., Kawashima, K., Kohara, M., Matsumoto, M., Shimpo, S., Tsuruoka, H., Wada, T., Yamada, M., Tabata, S., 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. *DNA Res.* **9**: 189-197

Kerstens, K., Wood, W. A., De Ley, J., 1965. Polyol dehydrogenases of *Gluconobacter oxydans*. *J. Biol. Chem.* **240**: 965-974

Killgore, J., Smidt, C., Duich, L., Romero-Chapman, N., Tinker, D., Reiser, K., Melko, M., Hyde, D., Rucker, R. B., 1989. Nutritional importance of pyrroloquinoline quinone. *Science* **245**: 850-852

Klasen, R., Bringer-Meyer, S., Sahm, H., 1992. Incapability of *Gluconobacter oxydans* to produce tartaric acid. *Biotechnol. Bioeng.* **40**: 183-186

Klasen, R., Bringer-Meyer, S., Sahm, H., 1995. Biochemical characterization and sequence analysis of the gluconate:NADP 5-oxidoreductase gene from *Gluconobacter oxydans*. *J. Bacteriol.* **177**: 2637-2643

- Kondo, K. & Horinouchi, S.**, 1997a. Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from *Gluconobacter suboxydans* and their expression in *Acetobacter pasteurianus*. *Appl. Environ. Microbiol.* **63**: 1131-1138
- Kondo, K. & Horinouchi, S.**, 1997b. Characterization of an insertion sequence, *IS12528*, from *Gluconobacter suboxydans*. *Appl. Environ. Microbiol.* **63**: 1139-1142
- Kondo, K. & Horinouchi, S.**, 1997c. A new insertions sequence *IS1452* from *Acetobacter pasteurianus*. *Microbiology* **143**: 539-546
- Kovari, H., Pienihakkinen, K., Alanen, P.**, 2003. Use of xylitol chewing gum in daycare centers: a follow-up study in Savonlinna, Finland. *Acta Odontol Scand.* **61**: 367-370
- Kranz, R. G. & Gennis, R. B.**, 1983. Immunological characterization of the cytochrome *o* terminal oxidase from *E. coli*. *J. Biol. Chem.* **258**: 10614-10621
- Kulhanek, M.**, 1989. Microbial dehydrogenations of monosaccharides. *Adv. Appl. Microbiol.* **34**: 141-181
- Kumazawa, T., Seno, H., Urakami, T., Matsumoto, T., Suzuki, O.**, 1992. Trace levels of pyrroloquinoline quinone in human and rat samples detected by gas chromatography/mass spectrometry. *Biochim. Biophys. Acta* **1156**: 62-66
- Kurtz, S. & Schleiermacher, C.**, 1999. Fast Computation of Maximal Repeats in Complete Genomes. *Bioinformatics* **15**: 426-427
- Leisinger, T.**, 1965. Untersuchungen zur Systematik und Stoffwechsel der Essigsäurebakterien. *Zentralbl. Bakteriol. II Abt.* **119**: 239-376
- Levering, P. R., Weenk, G., Olijve, W., Dijkhuizen, L., Harder, W.**, 1988. Regulation of gluconate and ketogluconate production in *Gluconobacter oxydans* ATCC 621-H. *Arch. Microbiol.* **149**: 534-539

Levering, R. R., Weenk, G., Olijve, W., Dijkhuizen, L., Harder, W. 1988. Regulation of gluconate and ketogluconate production in *Gluconobacter oxydans* ATCC 621-H. Arch. Microbiol. **149**: 534-539

Lobry, J. R. & Sueoka, N., 2002. Asymmetric directional mutation pressures in bacteria. Genome Biology **3**: 0058.1–0058.14

Lobry, J. R., 1996. Asymmetric substitution patterns in the two DNA strands of Bacteria. Mol. Biol. Evol. **13** (5): 660-665

Lucca, M. E., Spencer, J. F., Figueroa, L. I., 2002. Glycerol and arabitol production by an intergeneric hybrid, PB2, obtained by protoplast fusion between *Saccharomyces cerevisiae* and *Torulaspora delbrueckii*. Appl. Microbiol. Biotechnol. **59**: 472-476

Lusta, K. A. & Reshetilov, A. N., 1998. Physiological and biochemical features of *Gluconobacter oxydans* and prospects of their use in biotechnology and biosensor systems. Appl. Biochem. Microbiol. **34**: 307-320

Macauley, S., McNeil, B., Harvey, L. M., 2001. The genus *Gluconobacter* and its applications in biotechnology. Crit. Rev. Biotechnol. **21**: 1-25

Magnussen, O. T., Toyama, H., Saeki, M., Schwarzenbacher, R., Klinman, J. P., 2004. The structure of a biosynthetic intermediate of pyrroloquinoline quinone (PQQ) and elucidation of the final step of PQQ biosynthesis. J. Am. Chem. Soc. **126**: 5342-5343

Matsushita, K., Fujii, Y., Ano, Y., Toyama, H., Shinjoh, M., Tomiyama, N., Miyazaki, T., Sugisawa, T., Hoshino, T., Adachi, O., 2003. 5-Keto-D-gluconate production is catalysed by a quinoprotein glycerol dehydrogenase, major polyol dehydrogenase, in *Gluconobacter* species. Appl. Environ. Microbiol. **69**: 1959-1966

Matsushita, K., Nagatani, Y., Shinagawa, E., Adachi, O., Ameyama, M., 1991. Reconstitution of the ethanol oxidase respiratory chain in membranes of quinoprotein

alcohol dehydrogenase-efficient *Gluconobacter suboxydans* subsp. α strains. J. Bacteriol. **173**: 3440-3445

Matsushita, K., Shinagawa, E., Adachi, O., Ameyama, M., 1987. Purification, characterization and reconstitution of cytochrome *o*-type oxidase from *Gluconobacter suboxydans*. Biochim. Biophys. Acta **894**: 304-312

Matsushita, K., Takaki, Y., Shinagawa, E., Ameyama, M., Adachi, O., 1992. Ethanol oxidase respiratory chain of acetic acid bacteria: reactivity with ubiquinone of pyrroloquinoline quinone-dependent alcohol dehydrogenases purified from *Acetobacter aceti* and *Gluconobacter suboxydans*. Biosci. Biotechnol. Biochem. **56**: 304-310

Matsushita, K., Toyama, H., Yamada, M., Adachi, O., 2002. Quinoproteins: structure, function, and biotechnological applications. Appl. Microbiol. Biotechnol. **58**: 13-22

Matsushita, K., Toyoma, H., Adachi, O., 1994. Respiratory chains and bioenergetics of acetic acid bacteria. Adv. Microb. Physiol. **36**: 247-301

Matsushita, K., Yakushi, T., Toyama, H., Shinagawa, E., Adachi, O., 1996. Function of multiple heme *c* moieties in intramolecular electron transport and ubiquinone reduction in the quinohemoprotein alcohol dehydrogenase-cytochrom *c* complex of *Gluconobacter suboxydans*. J. Biol. Chem. **271**: 4850-4857

Matsushita, K., Yakushi, T., Toyoma, H., Adachi, O., Myoshi, H., Tagami, E., Sakamoto, K., 1999. The quinohemoprotein alcohol dehydrogenase of *Gluconobacter suboxydans* has ubiquinol oxidation activity at a site different from the ubiquinone reduction site. Biochim. Biophys. Acta **1409**: 154-164

Matzerath, I., Kläui, W., Klasen, R., Sahm, H., 1995. Vanadate catalysed oxidation of 5-keto-D-gulonic acid to tartaric acid: the unexpected effect of phosphate and carbonate on rate selectivity. Inorg. Chim. Acta **237**: 203-205

Merkel, R., 2003. A survey of codon and amino acid frequency bias in microbial

genomes focusing on translational efficiency. *J Mol Evol.* **57** :453-66

Merkl, R., 2004. SIGI: Score-based identification of genomic islands. *BMC Bioinformatics* **5**: 22

Miyazaki, T., Tomiyama, N., Shinjoh, M., Hoshino, T., 2002. Molecular cloning and functional expression of D-sorbitol dehydrogenase from *Gluconobacter suboxydans* IFO3255, which requires pyrroloquinoline quinone and hydrophobic protein SldB for activity development in *E. coli*. *Biosci. Biotechnol. Biochem.* **66**: 262-270

Moonmangmee, D., Adachi, O., Shinagawa, E., Toyama, H., Theeragool, G., Lotong, N., Matsushita, K., 2002. L-erythrulose production by oxidative fermentation is catalysed by PQQ-containing membrane-bound dehydrogenase. *Biosci. Biotechnol. Biochem.* **66**: 307-318

Moonmangmee, D., Fujii, Y., Toyoma, H., Theeragool, G., Lotong, N., Matsushita, K., Adachi, O., 2001. Purification and characterization of membrane-bound quinoprotein cyclic alcohol dehydrogenase from *Gluconobacter frateurii* CHM 9. *Biosci. Biotechnol. Biochem.* **65**: 2763-2772

Morozova, I., Qu, X., Shi, S., Asamani, G., Greenberg, J. E., Shuman, H. A., Russo J. J., 2004. Comparative sequence analysis of the *icm/dot* genes in *Legionella*. *Plasmid* **51**: 127-147

Moshiri, F., Smith, E. G., Tarmino, J. P., Maier, R. J., 1991. Transcriptional regulation of cytochrome *d* in nitrogen-fixing *Azotobacter vinelandii*. Evidence that up-regulation during n_2 fixation is independent of *nifA* but dependent on *ntrA*. *J. Biol. Chem.* **266**: 23169-23174

Mrazek, J., Karlin, S., 1999. Detecting alien genes in bacterial genomes. *Ann. N Y Acad. Sci.* **870**: 314-29

Naya, H., Romero, H., Zavala, A., Alvarez, B., Musto, H., 2002. Aerobiosis increases the genomic guanine plus cytosine content (GC%) in prokaryotes. *J. Mol. Evol.* **55**:

260-264

Nelson, K., Paulsen, I., Weinel, C., Dodson, R., Hilbert, H., Fouts, D., Gill, S., Pop, M., Martins Dos Santos, V., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J., Timmis, K., Duesterhoft, A., Tummler, B., Fraser, C., 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ. Microbiol.* **4**: 799-808

Nozaki, H., Suzuki, S., Tsuyoshi, N., Yokozeki, K., 2003. Production of D-arabitol by *Metschnikowia reukaufii* AJ14787. *Biosci. Biotechnol. Biochem.* **67**: 1923-1929

Nukada, H. K., Urakami, T., Murphy, M. P., 2003. Antioxidant and pro-oxidant properties of pyrroloquinoline quinone (PQQ): implications for its function in biological systems. *Biochem. Pharmacol.* **65**: 67-74

Ochman, H. & Santos, S. R., 2003. Eyeing bacterial genomes. *Curr. Opin. Microbiol.* **6**: 109-113

Ohrem, H. L. & Voß, H., 1995. Inhibitory effects of dihydroxyacetone on *Gluconobacter* cultures. *Biotechnol. Lett.* **17**: 981-984

Olijve, W. & Kok, J. J., 1979. An analysis of the growth of *Gluconobacter oxydans* in chemostat cultures. *Arch. Microbiol.* **121**: 291-297

Overbeek, R., Larsen, N., Walunas, T., D'Souza, M., Pusch, G., Selkov, Jr., Liolios, K., Joukov, V., Kaznadzey, D., Anderson, I., Bhattacharyya, A., Burd, H., Gardner, W., Hanke, P., Kapatral, V., Mikhailova, N., Vasieva, O., Osterman, A., Vonstein, V., Fonstein, M., Ivanova, N., Kyrpides, N., 2003. The ERGOTM genome analysis and discovery system. *Nucleic Acids Res.* **31** (1): 164-71

Paulsen, I. T., Seshadri, R., Nelson, K. E., Eisen, J. A., Heidelberg, J. F., Read, T. D., Dodson, R. J., Umayam, L., Brinkac, L. M., Beanan, M. J., Daugherty, S. C., Deboy, R. T., Durkin, A. S., Kolonay, J. F., Madupu, R., Nelson, W. C., Ayodeji, B., Kraul, M., Shetty, J., Malek, J., Van Aken, S. E., Riedmuller, S., Tettelin, H., Gill, S. R., White, O., Salzberg, S. L., Hoover, D. L., Lindler, L. E., Halling, S. M., Boyle, S. M., Fraser, C. M., 2002. The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. Proc. Natl. Acad. Sci. U.S.A. **99**: 13148-13153

Perlman, D., 1979. Microbial Technology. Pepllar, H. J. and Perlman, D., (Eds) 2nd Ed., Vol. II, Academic Press ISBN: 0-12-551502-2

Pierini, C., 2001 Xylitol: A sweet alternative. Unique sweetener reduces tooth decay and infections. Vitamin Research News **15**: 1-11

Praphanphoj, V., Sacksteder, K. A., Gould, S. J., Thomas, G. H., Geraghty, M. T., 2001. Identification of the α -amino adipic semialdehyde dehydrogenase-phosphopantetheinyl transferase gene, the human ortholog of the yeast LYS5 gene. Mol. Genet. Metab. **72**:336-342

Rainbow, C. & Rose, A. H. (Eds.), 1963. Biochemistry of industrial microorganisms: Miscellaneous oxidative transformations (A. N. HALL). Academic Press, London, **17**: 607-612

Ramirez-Romero, M. A., Bustos, P., Girard, L., Rodriguez, O., Cevallos, M. A., Davila, G., 1997. Sequence, localization and characteristics of the replicator region of the symbiotic plasmid of *Rhizobium etli*. J. Microbiol. **143**: 2825-2831

Reichstein, T. & Grüssner, A., 1934. Eine ergiebige Synthese der 1-Ascorbinsäure (C-Vitamin). Helv. Chim. Acta **17**: 160-164

Rice, C. W. & Hempfling, W. P., 1978. Oxygen-limited continuous culture and respiratory energy conservation in *Escherichia coli*. J. Bacteriol. **134**: 115-124

Romano, A. H., Eberhard, S. J., Dingle, S. L., McDowell, T. D., 1970. Distribution of the phosphoenolpyruvate: glucose phosphotransferase system in bacteria. *J. Bacteriol.* **104**: 808-813

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A., Barrell, B., 2000. Artemis: Sequence visualization and annotation. *Bioinformatics* **16** (10): 944-945

Saito, Y., Ishii, Y., Hayashi, H., Imao, Y., Akashi, T., Yoshikawa, K., Noguchi, Y., Soeda, S., Yoshida, M., Niwa, M., Hosoda, J., Shimomura, K., 1997. Cloning genes coding for L-sorbose and L-sorbosone dehydrogenase from *Gluconobacter oxydans* and microbial production of 2-keto-L-gulonate, a precursor of L-ascorbic acid, in a recombinant *G. oxydans* strain. *Appl. Environ. Microbiol.* **63**: 454-460

Saito, Y., Ishii, Y., Hayashi, H., Yoshikawa, K., Noguchi, Y., Yoshida, S., Soeda, S., Yoshida, M., 1998. Direct fermentation of 2-keto-L-gulonic acid in recombinant *Gluconobacter oxydans*. *Biotechnol. Bioeng.* **58**: 309-315

Salisbury, S. A., Forrest, H. S., Cruse, W. B., Kennard, O., 1979. A novel coenzyme from bacterial primary alcohol dehydrogenases. *Nature* **280**: 843-844

Salusjarvi, T., Povelainen, M., Hvorslev, N., Eneyskaya, E. V., Kulminskaya, A. A., Shabalin, K. A., Neustroev, K. N., Kalkkinen, N., Miasnikov, A. N., 2004. Cloning of a gluconate/polyol dehydrogenase gene from *Gluconobacter suboxydans* IFO 12528, characterization of the enzyme and its use for the production of 5-ketogluconate in a recombinant *Escherichia coli* strain. *Appl. Microbiol. Biotechnol.* In press

Salzberg, S., Delcher, L., Kasif, S., White, O., 1998. Microbial gene identification using interpolated Markov models. *Nucleic Acids Res.* **26**: 544-548

Sambrook, J., Fritsch, E. F., Maniatis, T., 1989. *Molecular cloning: a laboratory manual* (2nd ed.). Cold Spring Harbour Laboratory Press, Cold Spring harbour, New York.

Sanger, F., Nicklein, S., Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Sci. USA* **74**: 5463-5467

Sauer, U., Canonaco, F., Heri, S., Perrenoud, A., Fischer, E., 2004. The soluble and membrane-bound transhydrogenases UdhA and PntAB have divergent functions in NADPH metabolism of *Escherichia coli*. *J. Biol. Chem.* **279**: 6613-6619

Schedel, M., 2000. Regioselective oxidation of aminosorbitol with *Gluconobacter oxydans*, a key reaction in the industrial synthesis of 1-deoxynojirimycin. In: Kelly, D. R. (Ed.) *Biotechnology, Vol. 8b. Biotransformations I*. Wiley-VCH, Weinheim, S. 296-308

Schneider, E. & Altendorf, K., 1985. All three subunits are required for the reconstruction of an active proton channel (F_0) of *Escherichia coli* ATP synthase (F_1F_0). *EMBO J.* **4**:515-518

Schultz, B. E. & Chan, S. I., 1998. Thermodynamics of electron transfer in *Escherichia coli* cytochrome *bo*₃. *Proc. Natl. Acad. Sci. USA* **95**: 11643-11648

Scott, A. R. & Tattersall, R. B., 1988. Alpha glucosidase inhibition in the treatment of non-insulin-dependent diabetes mellitus. *Diabet. Med.* **15**: 42-46

Sharrocks, A. D., Green, J., Guest, J. R., 1991. FNR activates and represses transcription in vitro. *Proc. R. Soc. Lond. B. Biol. Sci.* **245**: 219-226

Shibata, T., Ichikawa, C., Matsuura, M., Takata, Y., Noguchi, Y., Saito, Y., Yamashita, M., 2000. 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

Shimao, M., Ninomiya, K., Kuno, O., Kato, N., Sakazawa, C., 1986. Existence of a novel enzyme, pyrroloquinoline quinone-dependent polyvinyl alcohol dehydrogenase, in a bacterial symbiont, *Pseudomonas* sp. Strain VM15C. *Appl. Environ. Microbiol.* **51**: 268-275

Shimizu, H., Tsuchie, H., Yoshida, K., Morikawa, S., Tsuruoka, T., Yamamoto, H., Ushijima, H., Kitamura, T., 1990. Inhibitory effect of novel 1-deoxynojirimycin derivatives on HIV-1 replication. *AIDS* **10**:975-979

Shinagawa, E., Matsushita, K., Adachi, O., Ameyama, M., 1981. Purification and characterization of 2-keto-D-gluconate dehydrogenase from *Gluconobacter melanogenus*. *Agric. Biol. Chem.* **45**: 1079-1085

Shinagawa, E., Matsushita, K., Adachi, O., Ameyama, M., 1982. Purification and characterization of D-sorbitol dehydrogenase from membrane of *Gluconobacter suboxydans* var. *Agric. Biol. Chem.* **46**: 135-141

Shinagawa, E., Matsushita, K., Ameyama, M., Adachi, O., 1984. D-gluconate dehydrogenase, 2-keto-D-gluconate yielding from *Gluconobacter dioxyaceticus*: purification and characterization. *Agric. Biol. Chem.* **48**: 1517-1522

Shinjoh, M., Tazoe, M., Hoshino, T., 2002b. NADPH-dependent L-sorbose reductase is responsible for L-sorbose assimilation in *Gluconobacter suboxydans* IFO 3291. *J. Bacteriol.* **184**: 861-863

Shinjoh, M., Tomiyama, N., Asakura, A., Hoshino, T., 1995. Cloning and nucleotide sequencing of the membrane-bound L-sorbose dehydrogenase gene of *Acetobacter liquefaciens* IFO 12258 and its expression in *Gluconobacter oxydans*. *Appl. Environ. Microbiol.* **61**: 413-420

Shinjoh, M., Tomiyama, N., Miyazaki, T., Hoshino, T., 2002a. Main polyol dehydrogenase of *Gluconobacter suboxydans* IFO 3255, membrane bound D-sorbitol dehydrogenase, that needs product of upstream gene, *sldB*, for activity. *Biosci. Biotechnol. Biochem.* **66**: 2314-2322

Smith, L., 1954. Bacterial cytochromes; difference spectra. *Arch. Biochem. Biophys.* **20**: 299-314

Spiro, S., Roberts, R. E., Guest, J. R., 1989. FNR-dependent repression of the *ndh* gene of *Escherichia coli* and metal ion requirement for FNR-regulated gene expression. *Mol. Microbiol.* **3**: 601-608

Staden, R., 1980. A new computer method for the storage and manipulation of DNA gel reading data. *Nucleic Acids Res.* **8**: 3673-3694

Staden, R., Beal, K. F., Bonfield, J. K., 2000. The Staden package, 1998. *Methods Mol. Biol.* **132**: 115-130

Steinberg, F. M., Gershwin, M. E., Rucker, R. B., 1994. Dietary pyrroloquinoline quinone: growth and immune response in BALB/c mice. *J. Nutr.* **124**: 744-753

Stouthamer, A. H., 1959. *Antonie van Leeuwenhoek*, **25**: 241-264

Sugisawa, T. & Hoshino, T., 2002. Purification and properties of membrane-bound D-sorbitol dehydrogenase from *Gluconobacter suboxydans* IFO 3255. *Biosci. Biotechnol. Biochem.* **66**: 57-64

Sugisawa, T., Hoshino, T., Masuda, S., Nomura, S., Setoguchi, Y., Tatoe, M., Shinjoh, M., Someha, S., Fujiwara, A., 1990. Microbiol production of 2-keto-L-gulonic acid from L-sorbose and D-Sorbitol by *Gluconobacter melanogenus*. *Agric. Biol. Chem.* **54**: 1201-1209

Sugisawa, T., Ojima, S., Matzinger, P. K., Hoshino, T., 1995. Isolation and characterization of a new vitamin C producing enzyme (L-gulono- γ -lactone dehydrogenase) of bacterial origin. *Biosci. Biotech. Biochem.* **59**: 190-196

Sugiyama, M., Suzuki, S., Tonouchi, N., Yokozeki, K., 2003a. Cloning of the xylitol dehydrogenase from *Gluconobacter oxydans* and improved production of xylitol from arabitol. *Bioci. Biotechnol. Biochem.* **67**: 584-591

Sugiyama, M., Suzuki, S., Tonouci, N., Yokozeki, K., 2003b. Transaldolase/glucose-6-phosphate isomerase bifunctional enzyme and ribolokinase as factors to increase

xylitol production from D-arabitol in *Gluconobacter oxydans*. Biosci. Biotechnol. Biochem. **67**: 2524-2532

Suzuki, S., Sugiyama, M., Mihara, Y., Hashiguchi, K., 2002. Novel enzymatic method for the production of xylitol from D-arabitol by *Gluconobacter oxydans*. Biosci. Biotechnol. Biochem. **66**: 2614-2620

Svitel, J. & Sturdik, E., 1994. Product yield and by-product formation in glycerol conversion to dihydroxyacetone by *Gluconobacter oxydans*. J. Ferment. Technol. **78**: 351-355

Tayama, K., Fukaya, M., Okumura, H., Kawamura, Y., Beppu, T., 1989. Purification and characterization of membrane-bound alcohol dehydrogenase from *Acetobacter polyoxogenes* sp. nov. Appl. Microbiol. Biotechnol. **32**: 181-185

Tech, M. & Merkl, R., 2003. YACOP: Enhanced gene prediction obtained by a combination of existing methods. In silico Biology **3**: 0037

Tiller, E. R. & Collins, R. A., 2000. The contributions of replication orientation, gene direction, and signal sequences to base-composition asymmetries in bacterial genomes. J. Mol. Evol. **50**: 249-257

Tkac, J., Navratil, M., Sturdik, E., Gemeiner, P., 2001. Monitoring of dihydroxyacetone production during oxidation of glycerol by immobilized *Gluconobacter oxydans* cells with an enzyme biosensor. Enzyme Microb. Technol. **28**: 383-388

Tonouchi, N., Sugiyama, M., Yokozeki, K., 2003. Coenzyme specificity of enzymes in the oxidative pentose phosphate pathway of *Gluconobacter oxydans*. Biosci. Biotechnol. Biochem. **67**: 2648-2651

Touger-Decker, R., van Loveren, C., 2003. Sugars and dental caries. Am. J. Clin. Nutr. **78**: 881S-892S

Toyoma, H., Fujii, A., Matsushita, K., Shinagawa, E., Ameyama, M., Adachi, O., 1995. Three distinct quinoprotein alcohol dehydrogenases are expressed when *Pseudomonas putida* is grown on different alcohols. *J. Bacteriol.* **177**: 2422-2450

Trumpower, B. L., 1990. The protonmotive Q cycle. Energy transduction by coupling of proton translocation to electron transfer by the cytochrome bc₁ complex. *J. Biol. Chem.* **265**: 11409-11412

Urbance, J. W., Bratina, B. J., Stoddard, S. F., Schmidt, T. M., 2001. Taxonomic characterization of *Ketogulonigenium vulgare* gen. nov., sp. nov. and *Ketogulonigenium robustum* sp. nov., which oxidize L-sorbose to 2-keto-L-gulonic acid. *Int. J. Syst. Evol. Microbiol.* **51**: 1059-1070

Verma, V., Qazi, R., Cullum, J., Qazi, G. N., 1997. Genetic heterogeneity among keto-acid-producing strains of *Gluconobacter oxydans*. *World J. Micro. Biotech.* **13**: 289-294

Vosss, H., Wiemann, S., Wirkner, U., Schwager, C., Zimmermann, J., Erfle, H., Hewitt, N. A., Rupp, T., Ansorge, W., 1992. Automated DNA sequencing system revolving 1000 bases with Fluorescein-15-dATP as internal label. *Methods Mol. Cell. Biol.* **3**: 153-155

Walker, J. E., Saraste, M., Gay, N. J., 1982. The *unc* operon: nucleotide sequence, regulation and structure of ATP synthase. *Biochim. Biophys. Acta* **768**:164-200

Waterman, M., S., 1984. Efficient sequence alignment algorithms. *J. Theor. Biol.* **108**: 333-337

Weenk, G., Olive, W., Harder, W., 1984. Ketogluconate formation by *Gluconobacter* species. *Appl. Microbiol. Biotechnol.* **20**: 400-405

Williams, R., Cotton, N. P., Thomas, C. M., Jackson, J. B., 1994. 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

Wu, G., Hill, S., Kelly, M. J. S., Sawers, G., Poole, R. K., 1997. The *cydR* gene product, required for cytochrome *bd* expression in obligate aerobe *Acetobacter vinelandii*, is an FNR-like protein. *Microbiology* **143**: 2197-2207

Yamada, M., Sumi, K., Matsushita, K., Adachi, O., Yamada, Y., 1993. Topological analysis of quinoprotein glucose dehydrogenase of *Escherichia coli* and its ubiquinone-binding site. *J. Biol. Chem.* **268**: 12812-12817

Yasuda, M., Cherepanov, A., Duine, J. A., 1996. Polyethylene glycol dehydrogenase activity of *Rhodopseudomonas acidophila* derives from a type I quinohaemoprotein alcohol dehydrogenase. *FEMS Microbiol. Lett.* **138**: 23-28

Zarnt, G., Schrader, T., Andreesen, J. R., 2001. Catalytic and molecular properties of the quinohemoprotein tetrahydrofuryl alcohol dehydrogenase from *Ralstonia eutropha* strain Bo. *J. Bacteriol.* **183**: 1954-1960

Danksagung

Prof. Dr. Deppenmeier danke ich für die Überlassung des Themas, die Möglichkeit zum selbständigen wissenschaftlichen Arbeiten sowie die geleistete Unterstützung. Mein Dank gilt auch Herrn Prof. Dr. Gottschalk für die bereitwillige Übernahme des Referats, sowie sein stetiges Interesse am Fortgang meiner Arbeit. Herzlichen Dank auch, dass ich an vielen nationalen und internationalen Tagungen teilnehmen durfte.

Vielen Dank dem ehemaligen Labor 206: Marc Hoffmeister, Kai Hofmann, Frank Falinski, Hauke Schmidt, Marcel Grapp, Sonja Volland, Tomas Lehmann, Ray Hovey und Sabine Lentes. Es war ein einzigartiges Arbeitsklima, das ich gerne auch bis zum Ende meiner Promotion genossen hätte.

Allen ehemaligen und jetzigen Mitarbeitern des Göttinger Genomlabors möchte ich für die gute Zusammenarbeit danken. Insbesondere gilt dies für Rosa Martínez-Arias für die Einführung in Gap4, Heiko Liesegang für die bioinformatische Bewältigung der Daten, Florian Fricke für die Bewältigung der *Pile-ups* und für Axel Strittmatter, Arnim Wiezer, Silke Steckel, Silke Denker, Iwona Decker, Jarek Sobkowiak, Mechthild Bömecke, Rainer Merkl und Holger Brüggemann, die mir immer mit Rat und Tat zur Seite standen..

Für das Korrekturlesen danke ich Marc Hoffmeister und Claudia Wirth. Claudia danke ich auch für die vielen Gespräche, die weit über den Laboralltag hinausgingen.

Allen weiteren Mitgliedern des Instituts, die mich während meiner Doktorarbeit unterstützt haben, sei an dieser Stelle gedankt.

Meiner Stallbande, insbesondere Karen Herwig und Jennifer Lohs danke ich für die großartige Unterstützung und die Sorge um mein leibliches Wohl.

Meinem Mann Jörg sei für jegliche seelische und ideelle Unterstützung während meiner Promotion gedankt und herzlichen Dank auch meiner Familie, für die in vielerlei Hinsicht geleistete Unterstützung.

Lebenslauf

02.07.1973	geboren in Göttingen
1980 - 1984	Grundschule Herberhausen, Göttingen
1984 - 1986	Orientierungsstufe - Lutherschule in Göttingen
1986 - 1993	Hainberg-Gymnasium in Göttingen
Mai 1993	Abitur
August 1993 – Juni 1996	Ausbildung zur Biologielaborantin am Institut für Tierzucht und Haustiergenetik der Georg-August-Universität zu Göttingen
Oktober 1996	Immatrikulation an der Georg-August-Universität zu Göttingen für das Studienfach Biologie
Oktober 1998	Diplomvorprüfung in den Fächern Mikrobiologie, Zoologie, Chemie und physikalische Chemie
November 2000	Diplomprüfung in den Fächern Mikrobiologie, Genetik und Chemie
November 2000 – Oktober 2001	Anfertigung der experimentellen Diplomarbeit unter Anleitung von Priv. Doz. Dr. B. Averhoff am Institut für Mikrobiologie und Genetik der Georg-August-Universität zu Göttingen mit dem Titel: „Identifizierung und Charakterisierung neuer Kompetenzgene in <i>Thermus thermophilus</i> HB27“
Oktober 2001	Beginn der experimentellen Arbeit zur vorliegenden Dissertation