Metabolic Engineering of *C. glutamicum* for Amino Acid Production Improvement
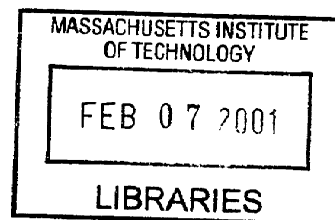
by

Mattheos A.G. Koffas

B.S. Chemical Engineering
National Technical University, Athens Greece, 1994

Submitted to the Department of Chemical Engineering in Partiall Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING

at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
October, 2000

Signature of the Author:—————————————
Department of Chemical Engineering
October 10, 2000

Certified by:—

Gregory Stephanopoulos
Professor of Chemical Engineering
Thesis Supervisor

Accepted by:———— \

Robert E. Cohen
St. Laurent Professor of Chemical Engineering
Chairman, Committee for Graduate Students

ABSTRACT

A central goal in metabolic engineering is the design of more productive biological systems by genetically modifying metabolic pathways. In this thesis we report such an optimization in the bacterial strain *Corynebacterium glutamicum* that is employed for the fermentative production of various amino acids such as lysine.

The main goal of the research presented here was the application of metabolic and genetic engineering tools in order to investigate the role of the pyruvate node in cellular physiology. This was achieved by integrating the tools of bioinformatics, recombinant DNA technology, enzymology and classical bioengineering in the context of control and genetically engineered strains of *C. glutamicum*.

First, the main anaplerotic pathway responsible for replenishing oxaloacetate, namely pyruvate carboxylase was targeted. After fruitless attempts to establish an *in vitro* enzymatic activity for this enzyme, our efforts were directed towards its gene identification. This was achieved by designing PCR primers corresponding to homologous regions among pyruvate carboxylases from other organisms. Utilizing these primers, a PCR fragment was isolated corresponding to part of the gene of the *C. glutamicum* pyruvate carboxylase. The sequence of the complete gene was finally obtained by screening a *C. glutamicum* cosmid library.

In order to investigate the physiological effect that this enzyme has on lysine production, recombinant strains and deletion mutants were generated. The presence of the gene of pyruvate carboxylase in a multicopy plasmid is not sufficient to yield a significant overexpresssion of this enzyme in *C. glutamicum*. Contrary to our expectations, overexpression of pyruvate carboxylase has a negative effect on lysine production but improves significantly the growth properties of *C. glutamicum*. A metabolic model was developed according to which pyruvate carboxylase overexpression increases the carbon flux that enters the TCA cycle, thus the higher growth. However due to the presence of a rate-limiting step in the lysine biosynthesis pathway this increased carbon flux does not translate into higher lysine production.

The role of aspartokinase, the first step in lysine biosynthesis, was explored as such a potential bottleneck. Its overexpression proves to increase the amount of lysine produced, however it leads to a lower growth and finally a lower productivity. Since pyruvate carboxylase and aspartokinase have opposite effects on cell physiology, the combination of the overexpression of these two enzymes was finally studied. By this simultaneous overexpression, we achieved to create a *C. glutamicum* recombinant strain with similar growth as that of the control but higher lysine production and productivity.

In the context of exploring the physiological role of pyruvate carboxylase, a biotinylated enzyme, two other enzyme that utilize biotin were also investigated namely acetyl-CoA-carboxylase and biotin ligase. The first enzyme was purified to completion and its N-terminal as well internal amino acid sequences were obtained. A cosmid from the *C. glutamicum* cosmid library was identified that most likely contains the gene of the latter enzyme.

In summary, in the present work we have achieved to prove unequivocally the presence of pyruvate carboxylase in *C. glutamicum*. We have also achieved to characterize the second biotinylated enzyme in this organism, namely acetyl-CoA-carboxylase. The physiological effect of both pyruvate carboxylase and aspartokinase was established and a metabolic model was developed based on these experimental results. This model finally led us to the construction of a new recombinant strain with improved lysine productivity. As such, this work stands as one of the few examples of a primary metabolite production improvement using metabolic engineering techniques.

Acknowledgments

Looking back into my years as graduate student, I would like to thank a number of people for their contribution in my work and my life:

First is the man that we went through this journey all together, from start to finish, my advisor Greg Stephanopoulos. He has been a teacher, a friend and the moving force behind all my work. During this time he always managed to bring the best out of me and most importantly, through his unbridled optimism, he gave me the motivation to keep on going. Thanks also go to another two people whose contribution has been invaluable: Anthony Sinskey who taught me everything I know about *Corynebacteria* and Martin Yarmush for all the input and support that he provided to me. I also appreciate the mentoring from the other members of my thesis committee, Daniel Wang, George Stephanopoulos and Arnold Demain.

My work would never have come into completion and life in the lab would not be as fun without the entire Stephanopoulos crew, past and present. From the old guard I would like to pay tribute to Aristos Aristidou for always being a good friend and a priceless source of wisdom, the lysine forerunners Sung Park and Cathryn Shaw-Reid, my pyruvate carboxylase colleague Roopa Ramamoorthi, Eduardo Agosin for his help and friendship and Roy Kamimura for our endless and relaxing philosophical discussions. It is hard to imagine how my thesis would have come to completion without the invaluable contribution of two later members of the group: Christopher Roberge and my lysine coworker and squash playmate Gyoo Yeol (Gary) Jung, for their insight, advice, help and patience with me. I will always treasure the memories of the time that I have spend with them (in and outside the lab) their uplifting words and encouragement and be grateful for simply making my life more beautiful. My deepest appreciation also goes to Saliya Silva for an enthusiasm that would always turn my scowl into a smile.

I have been blessed to have several gifted and dedicated undergraduates work by my side: Whitney Pine, Diana Donovan and Lee Lin. My admiration and respect for his brilliant work goes to Mike Jetten, a former member of the Sinskey group. Thanks also go

to my Latin instructor, Claudia von Canon, for letting me be part of her Latin Camera and for reinstating in me the love for the sublime world of the classics.

I would never have endured without the warmth and support of my three close friends Aphrodite, Penelope and Tina Giannakopoulos. I am proud to be their friend as they have been my thesis cheerleaders, my companions and in short a real family to me. I am also thankful to two other people whose passage through my Boston life brought me happiness and comfort, Konstantin Michailovits Linnik and Robert Jan Boer.

Finally, and for that reason most importantly, my deepest love, affection and gratitude goes to my parents, Georgios and Amalia and my brother Thomas. They inspired my faith in God, they raised my hope, and they have made me who I am. Their love and support has always been with me and has filled my heart with joy- even when there were miles separating us apart.

Trying to epitomize my thoughts as I put an end to my MIT years, I would like to quote one phrase that I first came across in my adolescence and has been with me ever since. It comes from Aristotle's Poetics and I dedicate it to all the people of the future who, at some point will find themselves in an overwhelming emotional crossroad, as I have been:

«... δι' ελέου καί φόβου περαίνουσα την των τοιούτων παθημάτων κάθαρσιν.»
"...(it is) through pity and fearful emotions that we finally reach our purgation".

Table of Contents

List of Figures

List of Tables

CHAPTER 1: BACKGROUND

1.A Metabolic Engineering

Altering metabolic pathways to improve cell properties and the chances of cell survival is as old as nature itself. The genomic and metabolic evolution of extremophiles is an example of such a natural adaptation process in bacteria (van der Oost, Ciaramella et al. 1998). The intentional manipulation of metabolic pathways by humans to improve the properties and productivity of microorganisms is similarly an established concept. Techniques such as genetic modifications via random mutagenesis have yielded, for example, improved strains of *Corynebacterium glutamicum* and its related species *Brevibacterium lactofermentum* and *Brevibacterium flavum* that excrete large amounts of amino acids into the fermentation medium (Kinoshita, Nakayama et al. 1958) (Shiio, Yoshino et al. 1990) (Yokota and Shiio 1988). Random mutagenesis relies heavily on chemical mutagens and creative selection techniques to identify superior strains for achieving a certain objective. Such traditional genetic approaches for strain improvement have been applied extensively in the past also in the areas of antibiotics, solvents and vitamin production among others.

Since the past decade, the development of recombinant DNA techniques has introduced a new dimension to pathway manipulation by offering, for the first time, the capability to construct specific metabolic configurations with novel, beneficial characteristics. Genetic engineering allows precise modification of specific enzymatic reactions in metabolic pathways, leading to the construction of well-defined genetic backgrounds. The redirection of cellular metabolism to create or enhance desirable attributes has been accomplished with a variety of novel techniques and applied towards an even greater variety of goals.

In this context, metabolic engineering has emerged as the technological and scientific discipline dealing with the introduction of specific modifications to metabolic pathways to improve cellular properties. Metabolic engineering involves manipulation of

enzymatic, transport, and regulatory functions of the cell by using recombinant DNA technology (Bouzier, Voisin et al. 1998) (Stephanopoulos and Vallino 1991). First, various analytical techniques are used to identify and subsequently determine fluxes through critical metabolic pathways in the cell or tissue of interest. This knowledge provides the rational basis for applying, in the second step, molecular biological techniques to enhance metabolic flux through a pathway of interest and minimize metabolic flow to undesired biosynthetically related products. Although a certain sense of direction is inherent in all strain improvement programs, the directionality of effort is a strong focal point of metabolic engineering, compared with random mutagenesis, because this directionality plays a dominant role in enzymatic target selection, experimental design and data analysis.

Although various terms have been coined over the past two decades to represent the increasing activity in pathway modification (pathway engineering, cellular engineering, in vitro evolution, etc) (MacQuitty 1988), (Nerem 1991), the term metabolic engineering has succeeded in capturing the ever-growing interest in this area. Furthermore, although initially embodied as a collection of examples from the chemical industry and biomedical research, metabolic engineering is quickly becoming a distinct scientific field. Its novel contribution lies in its emphasis on complete metabolic networks rather than individual reactions. To elaborate, as with all traditional fields of engineering, metabolic engineering too encompasses the two defining steps of analysis and synthesis. Because metabolic engineering emerged with DNA recombination as the enabling technology, its initial focus was on synthesis in the form of new pathway construction. As such, differentiation from genetic engineering was initially diffuse, and metabolic engineering could be considered as the technological manifestation of applied molecular biology. The real contribution of metabolic engineering emerged as soon as a need for a more rational approach to identifying promising targets of metabolic manipulation was articulated, replacing the previous, mostly ad hoc target selection process. In this sense the contribution of metabolic engineering emanates from pathway analysis, which yields an enhanced perspective on metabolism and cellular function, including consideration of reactions in their entirety rather than in isolation. Thus, metabolic engineering seeks to

analyze and then synthesize and design, using techniques and information developed from extensive reductionist research.

Metabolic engineering has found many applications, especially in microbial fermentation. It has been applied to increase the production of chemicals that are already produced by the host organism e.g. (Colon, Follettie et al. 1993) (Colon, Nguyen et al. 1995) (Follettie, Shin et al. 1988) (Ingram, Conway et al. 1987) (Tolan and Finn 1987), to produce desired chemical substrates from less expensive feed stocks, e.g. (Brabetz, Liebl et al. 1991), (Kumar, Ramakrishnan et al. 1992) (Zhang, Eddy et al. 1995), and to generate products that are new to the host organism e.g. (Ohta, Beall et al. 1991) (Ohta, Beall et al. 1991). Other challenges associated with metabolic engineering are the biosynthesis of secondary metabolites, the generation of organisms with desirable growth characteristics, and the manipulation of pathways for the production of chiral compounds as intermediates in the synthesis of pharmaceutical products (Jacobsen and Khosla 1998). Finally, although less widely appreciated, metabolic engineering techniques can also be applied for studying physiological systems and isolated whole organisms *in vivo* to elucidate the metabolic patterns that occur in different physiological states, such as "fed" or "fasted" as well as in disease (Yarmush and Berthiaume 1997).

In conclusion, a novel aspect of metabolic engineering is that it departs from the traditional reductionist paradigm of cellular metabolism, taking instead a holistic view. In this sense, metabolic engineering is well suited as a framework for the analysis of genome- wide differential gene expression data, in combination with data on protein content and *in vivo* metabolic fluxes. The insights of the integrated view of metabolism generated by metabolic engineering will have profound implications in biotechnological applications, as well as devising rational strategies for target selection for screening candidate drugs or designing gene therapies (Koffas, Roberge et al. 1999).

1.B An application of Metabolic Engineering: amino acid production

More than four decades ago, in 1957, a new bacterial strain was discovered, named *Corynebacterium glutamicum* (synonym *Micrococcus glutamicus*, Figure 1.1),

with a particular property of excreting L- glutamate. Since then, this strain has been the subject of extensive research and it has been shown to be an outstanding example of metabolic design, based on its flux properties. Today, *C. glutamicum* is mainly used for the biotechnological production of L-glutamate and L-lysine by fermentation. Figure 1.2 illustrates the metabolic pathway of L-lysine production.

L- glutamate and L-lysine are the two leading amino acids with the first being produced at approximately 10^6 tons per year and the second at approximately 500,000 tons per year (Eggeling and Sahm 1999). Indeed the increase in amino acid demand over the years is significant. Within ten years the total market has approximately doubled, with some amino acids displaying particularly large increase. For instance the world market for L-lysine has increased more than twenty fold in the past two decades. Estimates assume that the market is currently increasing by 10%-15% per year. This is a clear reflection of the improvement of living conditions around the world and in South America and China in particular. Part of the increased amino acid demand is due to increased pollution control since, with a balanced amino acid content in the feed, the manure contains less nitrogen.

The actual demand for the amino acids used as feed supplements can vary significantly. L-Lysine produced by fermentation competes with the natural L-lysine source soybean meal, where it is present in high concentrations. Since in 1997 the soybean meal market tightened globally, this led to an increased consumption of L-lysine made with *C. glutamicum* fermentation. This is also the case with other amino acids such as methionine, where chemically made methionine competes with fishmeal as the natural source of methionine. The worldwide decrease in sardine catches resulted in a decreased availability of this natural methionine source. Thus amino acid consumption is influenced by livestock and feed production throughout the world and the prices of natural sources. In addition to the specific amino acids required as feed supplements, amino acids are also used for a large variety of purposes, such as chemical building blocks, pharmaceutical products or food supplements.

The most common carbon sources for amino acid production by fermentation are starch hydrolysates, sucrose, high-test molasses (inverted can molasses), cane and sugar-

beet molasses. Most common nitrogen sources include ammonium sulfate and ammonia. During a continuous fermentation culture, L-lysine can accumulate to concentrations of up to 170 g/l (Eggeling and Sahm 1999). Usually lysine is present in the fermentation broth as a sulphate salt, due to the presence of ammonium sulphate in the culture. Appropriate restrained growth strategies can be used to maximize yield or productivity (Kiss and Stephanopoulos 1991), and yields of 54% L-lysine from glucose have been reported (Kawahara, Yoshihara et al. 1990).

For the downstream process of the L-lysine produced during the fermentation several different processes have been developed. One is the isolation of lysine by ion chromatography that can lead to the production of lysine at 98.5% purity. The second process involves biomass separation, evaporation and filtration to give a solution with lysine content of 50%. The third, very efficient procedure uses the direct spray drying of the entire fermentation broth, followed by granulation to yield a lysine sulfate preparation consisting of 47% L-lysine. This is the most efficient procedure since there are no losses in downstreaming and the waste volume is minimized.

1.C *Corynebacterium glutamicum*

Corynebacteria belong to the large group of Gram- positive bacteria with a high GC content, which constitutes the Actinomyces subdivision together with genera like *Streptomyces, Propionibacterium* or *Arthrobacter*. Chemotaxonomic studies comparing the cell wall composition (peptidoglycan structure, occurrence of mycolic acids) and lipid profiles suggest that the genera *Mycobacterium, Nocardia* and *Rhodococcus* are the closest relatives of *Corynebacterium* and the four genera are combined in the CMN group (Liebl 1992) (Barksdale 1970). The characteristic features of *Corynebacterium* are: nonsporing, nonmotile, not acid fast, straight or slightly curved rods, ovals or clubs (Figure 1.1); facultatively anaerobic to aerobic; predominant cell wall sugars are arabinose and galactose; mycolic acids present (Collins and Cummins 1986). The origin of the term "coryneform bacteria" is indeed a reference to their morphology and shape, as it is derived from the Greek word κορυνη meaning rod, club.

Nonmedical *Corynebacteria* can be isolated from a variety of different habitats (soil, water, plant material, animals) and normally grow well if cultivated aerobically at 30 °C in rich media. It is noteworthy that all strains of glutamic acid producing corynebacteria (*C. glutamicum* and similar strains) are dependent upon the presence of biotin in the growth medium (Abe, Takayama et al. 1967).

Certain saprophytic corynebacteria have a long tradition as industrial microorganisms in biotechnological production processes. The most prominent example is the production of amino acids by *C. glutamicum* and other similar bacteria, such as *Brevibacterium flavum, B. lactofermentum, B. roseum, B. divaricatum, C. lilium, C. herculis, Microbacterium ammoniaphilum* etc. The genome size of *C. glutamicum* is 3082 kb (Bathe, Kalinowski et al. 1996) and it is very likely that its entire genome sequence has already been completed (Eggeling and Sahm 1999).

The biotechnological importance of *C. glutamicum* for amino acid production has resulted in the development of cloning systems for this strain. Apart from a 29 kb *C. glutamicum* plasmid carrying streptomycin and spectinomycin- resistance determinants (Katsumata, Ozaki et al. 1984) all plasmids are cryptic. A multitude of cloning vectors for *C. glutamicum* and its related species have been constructed, most of which are shuttle vectors based on cryptic corynebacterial plasmids (Martin, Santamaria et al. 1987) (Archer, Follettie et al. 1989) (Hodgson, Krywult et al. 1990) (Messerotti, Radford et al. 1990) (Radford and Hodgson 1991) (Zhang, Praszkier et al. 1994).

Several methods for the introduction of recombinant plasmid molecules in *C. glutamicum* are available, such as protoplast transformation, conjugal transfer, polyethylene glycol (PEG)-induced protoplast fusion and phages (Katsumata, Ozaki et al. 1984) (Santamaria, Gil et al. 1984) (Yoshihama, Higashiro et al. 1985) (Karasawa, Tosaka et al. 1986) (Oki and Ogata 1968) (Patek, Ludvik et al. 1985) (Trautwetter, Blanco et al. 1987) (Trautwetter, Blanco et al. 1987) (Sonnen, Schneider et al. 1990) (Sonnen, Schneider et al. 1990) (Schafer, Kalinowski et al. 1990). However electroporation was shown to be a more convenient and much more efficient method for transformation of *C. glutamicum* (Liebl, Bayerl et al. 1989) (Haynes and Britz 1989) (Dunican and Shivnan 1989). An interesting feature of the molecular biology of *C.*

glutamicum is the broad expression of heterologous signals by this organism such as antibiotic resistance genes from *Escherichia coli* and various gram positive organisms such as *Streptomyces*. Furthermore a number of inducible *E. coli* promoters are functional in this organism, such as *lac*UV5, *tac* and *trp* (Morinaga, Tsuchiya et al. 1987).

C. glutamicum is one of the few saprophytic corynebacteria whose metabolism has been studied extensively. The uptake of glucose and fructose seems to be catalyzed by specific phosphoenolpyruvate:sugar phosphotransferase (PTS) systems. The main route for hexose breakdown, once inside the cell, is the Embden-Meyerhof pathway followed by the TCA cycle. Additionally, the key enzymes of the glyoxylate pathway, which are necessary for the assimilation of acetate have been demonstrated in these bacteria (Shiio, Otsuka et al. 1959) (Wendisch, de Graaf et al. 2000).

The question that naturally arises is: what peculiarities in the cell metabolism make *C. glutamicum* a superb amino acid producer? As in all bacteria lysine biosynthesis in *C. glutamicum* starts from aspartate, which is converted by aspartate kinase, aspartate semialdehyde dehydrogenase, dihydrodipicolinate synthase and dihydrodipicolinate reductase to L- piperideine-2, 6-dicarboxylate. At the aspartate semialdehyde branch point, carbon can enter into the metabolic pathways that produce threonine, methionine and isoleucine. At the L-piperideine-2, 6-dicarboxylate node is where one of the outstanding characteristics of the metabolism of *C. gltamicum* lies: there are two parallel routes from this point that can lead to the production of lysine, unlike any other bacterial species. Only recently enzymological and genetical studies have shown that these organisms can generate lysine either through the succinylase variant (four step pathway) or through the dehydrogenase variant (one step pathway) (Ishino, Yamaguchi et al. 1984) (Schrumpf, Schwarzer et al. 1991). Deletion mutants have shown that the single step pathway is not essential for the growth of *C. glutamicum* and that the lysine prototrophy still remains. However, in lysine hyperproducing strains, the final lysine yield is significantly lower when the dehydrogenase pathway is disrupted indicating that both pathways contribute to lysine formation in the parental strain (Schrumpf, Eggeling et al. 1992). A summary of the lysine biosynthesis pathway in *C. glutamicum* and *E. coli* is

given schematically in figures 1.2 and 1.3. The enzymes involved in the pathway are summarized in table 1.1.

The second important aspect of lysine formation from *C. glutamicum* is that there is no regulation of the synthesis of the enzymes of this pathway and that inhibition of enzyme activity was only found for aspartate kinase (inhibited by lysine plus threonine) (Nakayama, Tanaka et al. 1966) (Shiio and Miyajima 1969) (Tosaka, Takinami et al. 1978). This is a distinct *Corynebacterium* feature, in complete contrast with the lysine metabolism of other microbial organisms, such as *Escherichia coli* (Figure 1.3). The rather simple regulation strategies found in corynebacteria may explain the relative ease of obtaining high-level lysine producers. Single mutations introduced into the gene of aspartate kinase relieved this enzyme from feedback inhibition and led to the generation of lysine overproducer strains (Sano and Shiio 1970) (Kalinowski, Cremer et al. 1991) (Cremer, Eggeling et al. 1991). This is a further indication that a rather simple control of the lysine biosynthetic enzymes exists in *C. glutamicum*. Another enzyme whose overexpression also leads to significant lysine excretion is that of dihydrodipicolinate synthase (*dapA*), a surprising result considering the fact that this enzyme is not regulated by feedback mechanisms. When both aspartate kinase and dihydrodipicolinate synthase genes of *C. glutamicum* were overexpressed a further positive effect was observed on lysine accumulation (Eikmanns, Eggeling et al. 1993).

Overall a lysine overproducing *C. glutamicum* strain generally possesses the following characteristics: (i) feedback resistant aspartate kinase (ii) feedback resistant phosphoenolpyruvate carboxylase, one of the main anaplerotic enzymes in this organism (iii) much lower specific citrate synthase activity and (iv) much lower pyruvate kinase activity (Shiio, Yoshino et al. 1990). However there are examples in the literature that show that there might be several ways of obtaining high-producing strains and it might be that important features of these strains are still not known (Eikmanns, Eggeling et al. 1993). For example it has recently been shown that a highly active lysine exporter can significantly increase the rate of lysine production from *C. glutamicum* (Broer, Eggeling et al. 1993).

In conclusion, the detailed information on the biosynthetic pathways of *C. glutamicum* and their regulation and the availability of the genes involved permit a much better directed metabolic design, i.e. improvement of amino acid biosynthesis and excretion by manipulation of enzymatic and/or regulatory functions of *C. glutamicum* with the application of recombinant DNA technology. Even though such application has already found its way into the production of other amino acids such as threonine (Colon, Jetten et al. 1995), it has yet to be reported in a dramatic way in the case of lysine. Such an optimization using metabolic and genetic engineering tools has been the challenge of the work reported in this thesis.

In the present work, two particular *Corynebacterium* strains have been used extensively. The first is ATCC strain 21253 (*C. glutamicum*), a L-lysine producing strain that is a homoserine and leucine auxotroph due to the lack of homoserine dehydrogenase. The second is ATCC strain 21799 (*C. lactofermentum*). This is also a L-lysine producer that has acquired AEC resistance by random mutagenesis with the use of UV irradiation. It is also a leucine auxotroph.

1.D Motivation

A question that naturally arises in metabolic engineering is whether there is a critical branch point between catabolic and anabolic metabolism. This is highly unlikely due to the complexity of metabolism. However in a loose sense a node will be of kinetic importance at those points where carbon is partitioned between ATP generation on one hand and replenishment of TCA cycle intermediates for biosynthesis on the other. The quest for the identification of such nodes in *C. glutamicum* has led the focus of research in the past on the elucidation of the carbon flow in the central carbon metabolism in this strain. More specifically, research interest was more intense towards identifying the metabolic pathways responsible for the synthesis of oxaloacetate, which is the immediate precursor of amino acids of the aspartate family in these bacteria. In cells grown on carbohydrates, oxaloacetate can condense with acetyl-CoA to enter the tricarboxylic acid (TCA) cycle for energy generation, or it can be transaminated to aspartate, thus exiting

the TCA cycle for amino acid biosynthesis (Atkinson 1977) (Sanwal 1970). Oxaloacetate and other TCA-cycle intermediates used for biosynthesis are replenished via anaplerotic reactions.

Previous research (Vallino and Stephanopoulos 1993) has indicated that the yield and productivity of amino acids of the aspartate family would depend critically on the carbon flux through the anaplerotic pathway(s). Gubler et al (Gubler, Park et al. 1994) and Peters-Wendisch et al. (Peters-Wendisch, Eikmanns et al. 1993) obtained surprising results indicating that, in *C. glutamicum* one of the anaplerotic pathways, phosphoenolpyruvate carboxylase, is dispensable for growth and lysine production. In the above two studies, mutants deficient in phosphoenolpyruvate (PEP) carboxylase were constructed via a gene disruption method and were grown in glucose minimal medium. Both control and PEP-carboxylase-deficient mutant strains exhibited similar growth and lysine-production profiles, establishing the dispensability of PEP carboxylase (Figure 1.4). When pyruvate kinase, the enzyme that converts PEP into pyruvate was removed, lysine production dropped dramatically to about 50% of that of the control strain. This result indicates that pyruvate is a major source of oxaloacetate and hence lysine in *C. glutamicum*. The most common route for such an anaplerotic reaction between pyruvate and oxaloacetate is a pyruvate carboxylation. Finally, when both pyruvate kinase and PEP carboxylase were removed from *C. glutamicum* the amount of lysine produced dropped to negligible levels. A summary of these data is given in figure 1.4. Although both PEP carboxykinase (Jetten, Pitoc et al. 1994) and pyruvate carboxylase (Tosaka, Morioka et al. 1979) have been reported to be present in *B. lactofermentum* previous researchers have failed to provide unequivocal and reproducible evidence for the establishment of either of the two enzymes as the PEP-carboxylase-compensating enzyme.

In a separate study, it has been found that the apparent Km of pyruvate kinase (3.5 mM) for PEP is significantly lower than the associated Km for prosphoenolpyruvate carboxlase in *C. glutamicum* (10 mM). Under nominal conditions, phosphoenolpyruvate preferentially enters the pyruvate kinase branch. Any attempt to redirect the flow of carbon into the phosphoenolpyruvate carboxylase branch causes an increase in aspartate concentration due to the aspartate kinase kinetic bottleneck, and a drop in AcCoA

concentration due to the high activity of citrate synthase. The combination of high aspartate and low AcCoA concentrations results in a synergistic deactivation of phosphoenolpyruvate carboxylase activity and a precipitous drop in the phosphoenolpyruvate carboxylase flux. The net result of attempting to increase the prosphoenolpyruvate carboxylase split- ratio by blocking the pyruvate kinase branch typically results in the attenuation of both branches (Vallino 1991). Hence the phosphoenolpyruvate carboxylase node is strongly rigid. Coupled with the naturally higher activity of pyruvate kinase, there exists a much stronger driving force for PEP to enter the pyruvate kinase branch over the phosphoeolpyruvate carboxylase branch (Vallino 1991).

The above studies suggest that the PEP node may not play a significant anaplerotic role, thus making the pyruvate node a more probable one, assuming that a pyruvate carboxylating enzyme, which will serve the anaplerotic role, exists. Further more, NMR experimental work using ^{13}C labeled pyruvate further indicated that a pyruvate carboxylating pathway must be present that sustains a carbon flux nine times greater than that of the PEP carboxylating pathway (Park 1996). Most likely the unknown pyruvate carboxylating enzyme is pyruvate carboxylase. This is based on evidence that oxaloacetate formation by gluconeogenic enzymes such as oxaloacetate decarboxylase, malic enzyme and PEP carboxykinase is not sufficient to support growth on glucose. For example, a phosphoenolpyruvate carboxylase mutant of *Escherichia coli* was not able to grow on glucose even when transformed with a plasmid encoding the gene of phosphoenolpyruvate carboxykinase under the control of a *tac* promoter system (Chao and Liao 1993). Disruption of both genes of pyruvate carboxylase in *Saccharomyces cerevisiae* resulted in inability to grow on glucose despite the enhanced levels of phosphoenolpyruvate carboxykinase and malic enzyme (Brewster, Val et al. 1994). Similarly, a pyruvate carboxylase mutant of *Bacillus subtilis* was unable to grow on glucose even though high phosphoenolpyruvate carboxykinase activity was detected under the normal glycolytic conditions (Diesterhaft and Freese 1973; Brewster, Val et al. 1994). There is no *in vivo* evidence that oxaloacetate decarboxylase and malic enzyme can operate in the carboxylation direction. Furthermore, these carboxylation reactions are

very unlikely due to high free energy changes. All this evidence for the existence of pyruvate carboxylase makes the pyruvate node an obvious target and its control critical for the achievement of metabolite overproduction.

1.E Thesis Objectives

The main goal of the research presented here was the application of metabolic and genetic engineering tools in order to investigate the role of the pyruvate node in *C. glutamicum* physiology. This was achieved by integrating the tools of bioinformatics, recombinant DNA technology, enzymology and classical bioengineering in the context of wild type and genetically engineered strains of *C. glutamicum*. Once this physiological effect has been better understood, an optimal strategy was implemented for manipulating flux distribution to favor the overproduction of product.

In order to accomplish these objectives, the following specific aims were carried out

(A) Identify the enzymes that are responsible for depleting and producing pyruvate

(B) Determine the physiological effects that these enzymes have on cell growth and amino acid production.

(C) Develop a metabolic model for comprehending these physiological effects and identify combinations of other potential metabolic targets that can complement this model in order to achieve a production improvement.

1.F Method of Approach and Scope of Thesis

In order to meet the thesis objectives, a combination of experimental procedures belonging to the area of genetic engineering and bioinformatics, enzymology and classical bioengineering (as this is defined by bacterial fermentations) was applied.

23

In addressing the first objective, the first approach was by using *in vitro* enzymatic assays to measure the reaction rates of a possible pyruvate carboxylation reaction, the only step missing from the enzymatic scheme of the pyruvate node. Traditionally, biochemistry protocols rely upon enzymatic purification schemes to determine enzymatic reaction rates, but these rates can also be accurately probed by utilizing assays using cell extracts. For reasons that most likely have to do with the stability of the enzyme, *in vitro* pyruvate carboxylation was shown to be present, but contrary to our expectations this activity was low and independent of ATP. The enzymatic activity of pyruvate carboxylase was finally measured by the use of *C. glutamicum* permeabilized cells. This enzymatic verification was complemented by the simultaneous cloning of the gene of pyruvate carboxylase from this organism using homology search.

To complete the second objective, recombinant DNA technology was used in order to construct *C. glutamicum* and *C. lactofermentum* pyruvate carboxylase overexpressing strains. The physiological effect of these strains was then studied by the use of bacterial shake flask cultures and two major research questions were addressed: the ability of the recombinant strains to grow and their ability of to produce lysine. The productivity results were compared with control profiles and correlated with kinetic predictions. The technique of transconjugation was also used to construct mutant strains deficient in pyruvate carboxylase, by disrupting the gene of this enzyme within the *C. glutamicum* chromosome. The ability of the mutant strain to grow on various carbon sources was further assessed.

In order to meet the third objective, based on the data that the previous studies provided, a metabolic model was developed in order to identify potential targets for strain optimization. As such, aspartate kinase was chosen and the physiological effect that the overexpression of this gene has was also established. In a final step, the combined overexpression of pyruvate carboxylase and apsartokinase was performed, and similar physiological studies were conducted for the new recombinant strain.

On a separate level, we also tried to clarify further the role of biotin in amino acid production in general. To that extend, two more enzymes that utilize biotin were investigated, the first being acetyl-CoA-carboxylase and the second biotin ligase. The first

enzyme was purified to complete purity using standard procedures and establishing a chromatographic scheme. The N-terminal sequence as well as internal sequences of the purified protein was obtained. For the study of biotin ligase, a mutant *E. coli* strain with double phenotype was utilized in order to identify the gene of this enzyme from *C. glutamicum* as it has been described in the literature. Based on the screening of a cosmid library of this strain, potential candidates that contain this gene were identified and targeted for further investigation.

1.G Organization of Thesis Contents

The approach described above is further developed in each thesis chapter. Chapters 2,3,4 and 5 contain the experimental data that have been acquired during the completion of this thesis. Each one of these chapters is organized in a way that provides an introduction- summary of the accomplished work in the specific area and motivation for the next experiments- materials and methods, results and discussion sections.

Chapter 2 provides an overall introduction to the biotinylated enzymes in *C. glutamicum* and pyruvate carboxylation and gives the details of the discovery of the gene of pyruvate carboxylase in *C. glutamicum*. Chapter 3 discusses the experimental procedures and the results obtained that concern the physiological effects of pyruvate carboxylase on both cell growth and lysine production. The generation of plasmids that contain the gene of this enzyme as well as a deletion mutant are described, followed by batch culture data obtained by the use of the recombinant and mutant strains. A metabolic model is proposed at the end of the chapter that provides a theoretical explanation of the data presented. An experimental verification of this metabolic model is then described in detail in Chapter 4. Here the overexpression of aspartate kinase is initially described and the effect that this overexpression has on growth and production is presented based mostly on data available in the literature. In a final step, the simultaneous overexpression of both pyruvate carboxylase and aspartate kinase is described and the data that concern the physiology of the bicistronic recombinant strain are presented. Those data are based again on batch bacterial cultures and come as a confirmation to the proposed metabolic

model with the increase in lysine productivity obtained. Chapter 5 summarizes the attempts made in order to further understand the role of biotin in amino acid production in general. The study presented involves the purification of acetyl-CoA-carboxylase and the attempts to isolate the gene of biotin ligase from *C. glutamicum*.

Chapter 6 summarizes the contributions of this thesis and suggests future work to further explore other possible metabolic optimizations for product formation improvement. Finally, the literature is cited in alphabetical order in Chapter 7. All figures and tables are located at the end of the chapter in which they are referenced.
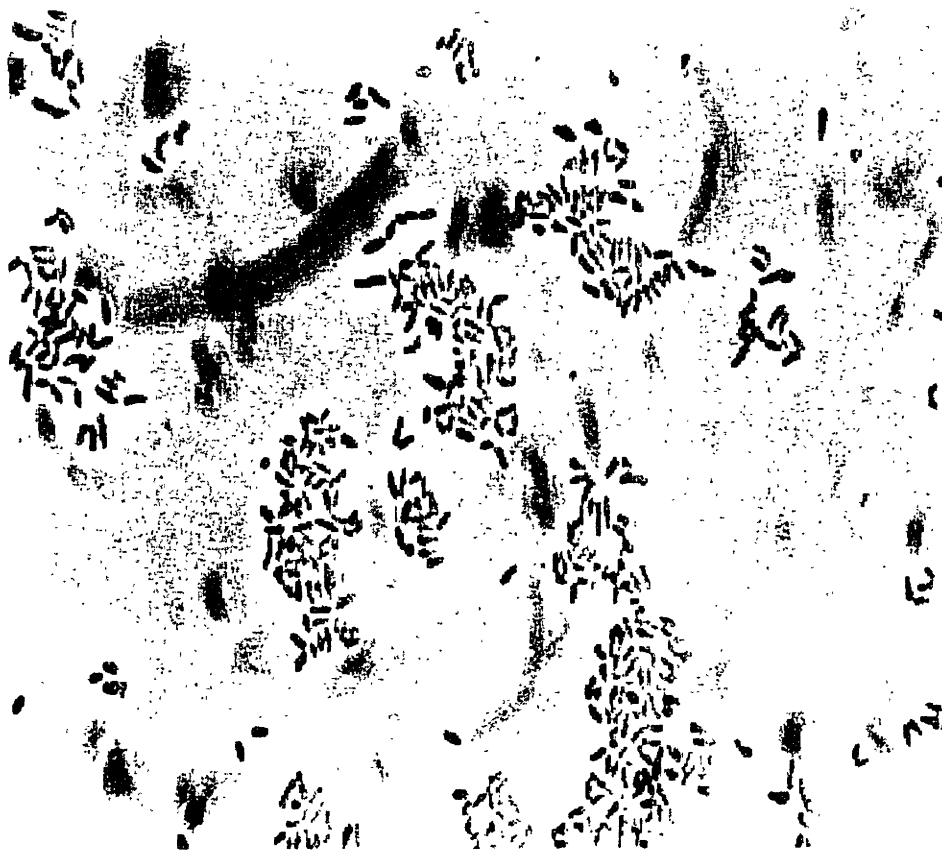
Figure 1.1 *Corynebacterium glutamicum*, an industrially important microbial strain
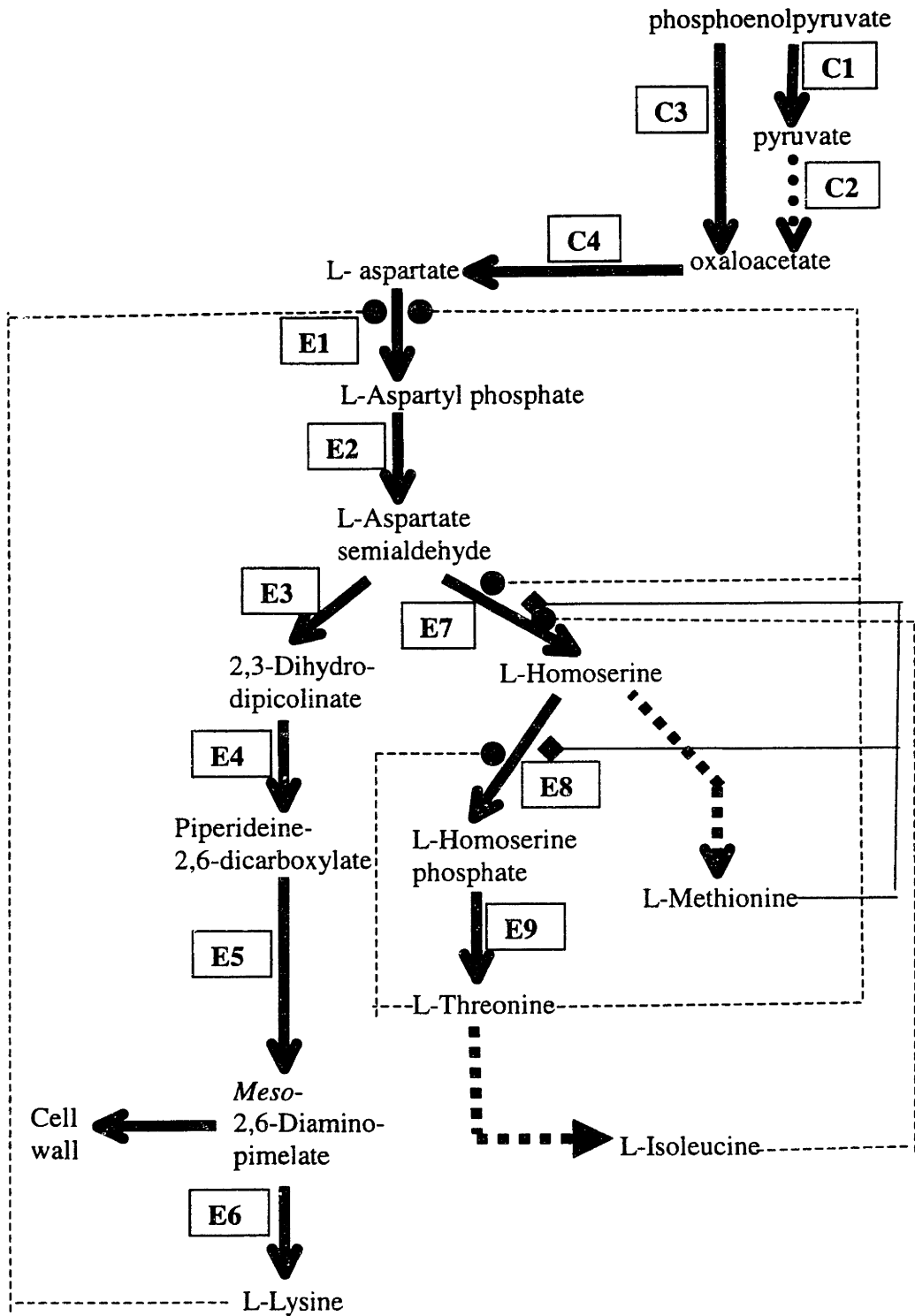
Figure 1.2 Regulation of the enzymes involved in the biosynthesis of L-lysine and L-threonine from aspartic acid in amino acid- producing corynebacteria. Thin lines: repression. Dashed lines: inhibition
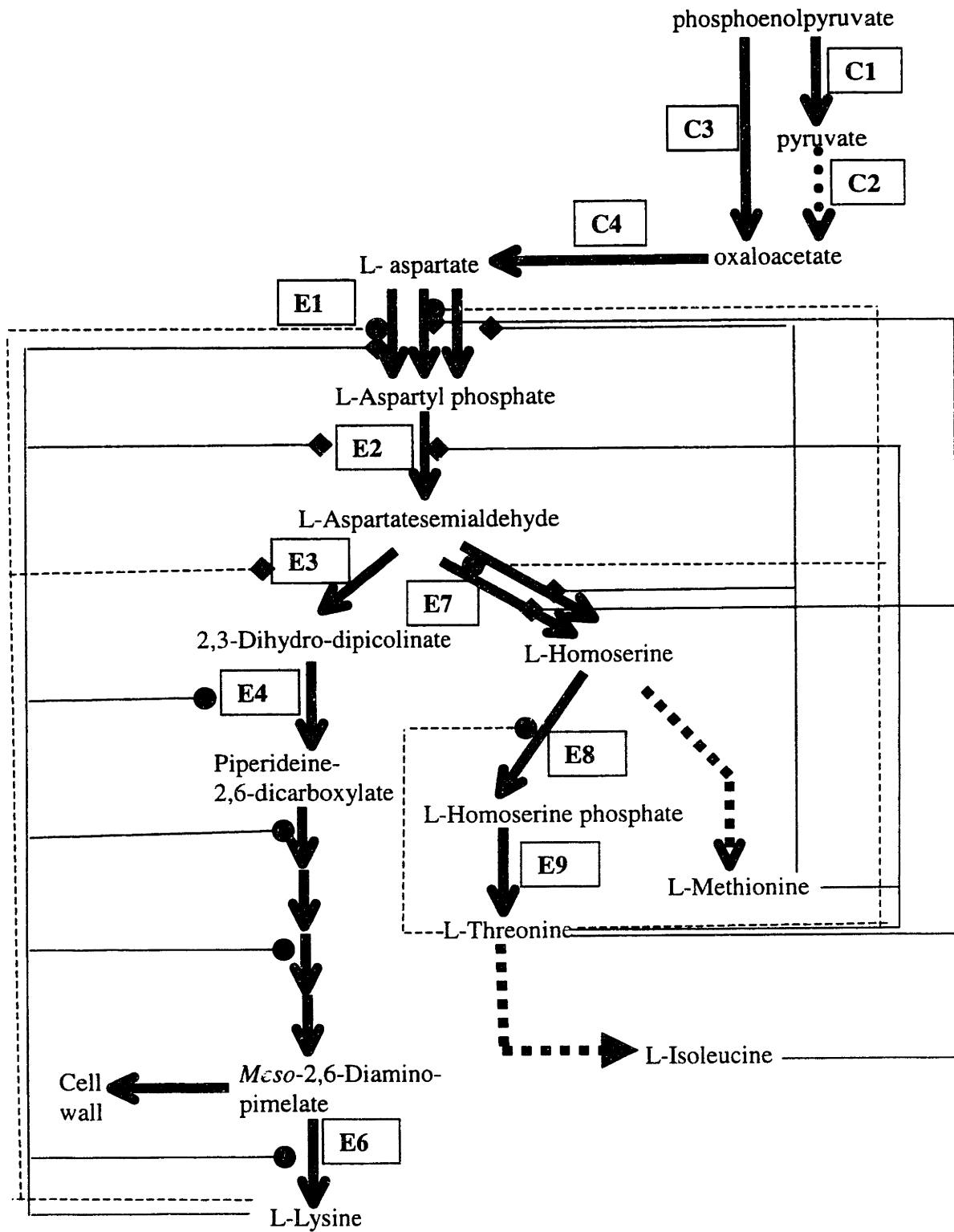
Figure 1.3 Regulation of the enzymes involved in the biosynthesis of L-lysine and L-threonine from aspartic acid in *E. coli*. Thin lines: repression. Dashed lines: inhibition.

Enzyme	Enzyme Name
C1	Pyruvate kinase
C2	Pyruvate carboxylase
C3	Phosphoenolpyruvate carboxylase
C4	Glutamate
E1	Aspartate kinase
E2	Aspartate semialdehyde dehydrogenase
E3	Dihydrodipicolinate synthase
E4	Dihydrodipicolinate reductase
E5	Diaminopimelate dehydrogenase
E6	Diaminopimelate decarboxylase
E7	Homoserine dehydrogenase
E8	Homoserine kinase
E9	Threonine synthase

Table 1.1 Enzyme abbreviations for the metabolic pathways shown in figures 1.2 and 1.3
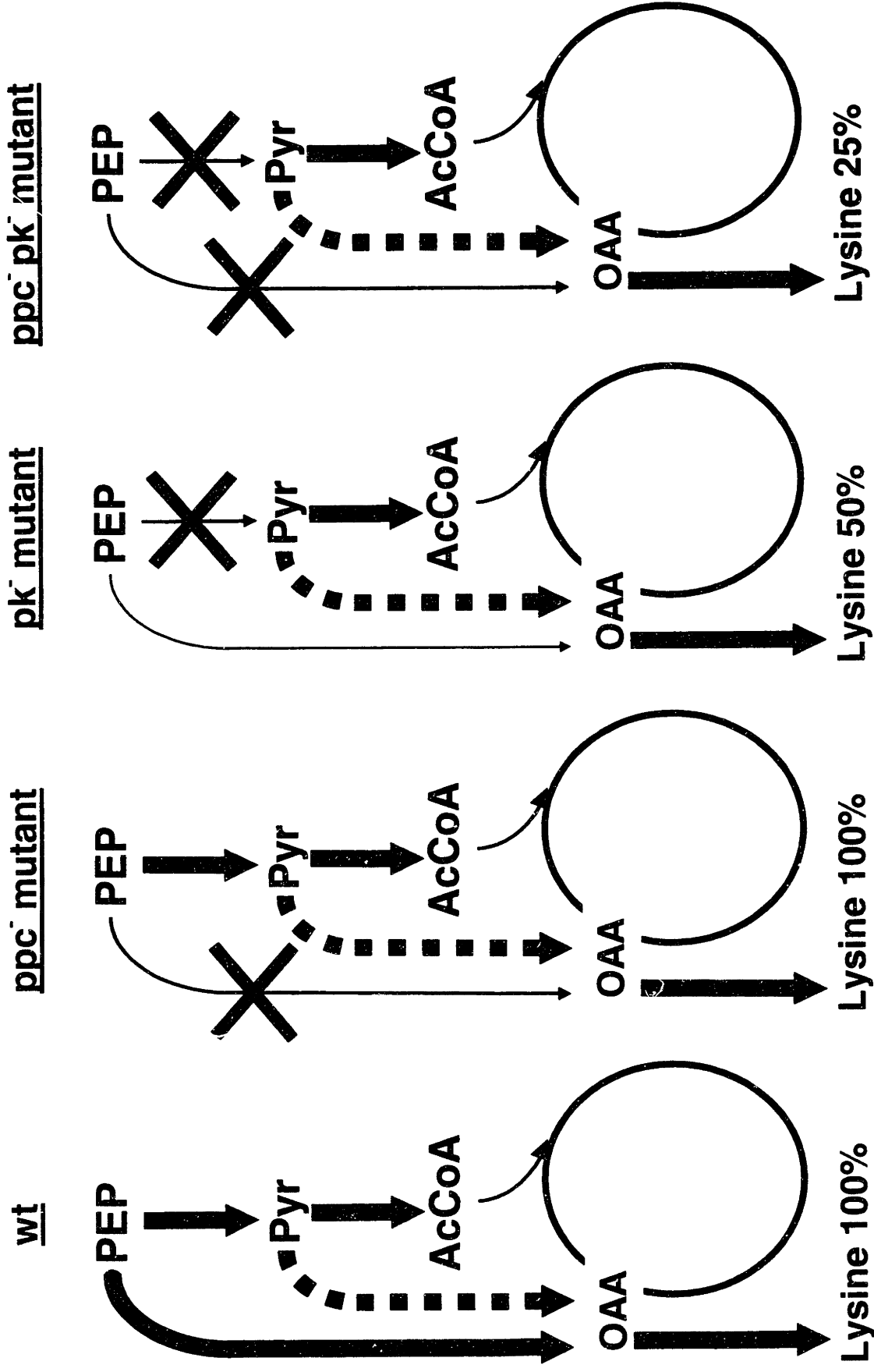
Figure 1.4 Phenotypic evidence for the existence of a pyruvate carboxylating enzyme, with the use of various *C. glutamicum* strains.

wt

PEP → Pyr → AcCoA → OAA

Lysine 100%

ppc⁻ mutant

PEP → Pyr → AcCoA → OAA

Lysine 100%

pk⁻ mutant

PEP → Pyr → AcCoA → OAA

Lysine 50%

ppc⁻ pk⁻ mutant

PEP → Pyr → AcCoA → OAA

Lysine 25%

CHAPTER 2: SEQUENCE OF THE *Corynebacterium glutamicum* PYRUVATE CARBOXYLASE GENE

2.A Introduction

2.A.1 Chapter overview

As indicated in the previous chapter, pyruvate carboxylase is an important anaplerotic enzyme replenishing oxaloacetate consumed for biosynthesis during growth, or lysine and glutamic acid production in industrial fermentations. Although accumulating experimental evidence pointed towards the existence of this enzyme in *C. glutamicum*, enzymatic efforts to detect its activity in protein crude extracts failed repeatedly.

This chapter starts with a review of similar work directed towards detecting *in vitro* pyruvate carboxylase activity in *C. glutamicum*. As this work was inconclusive, another approach was undertaken focusing on sequencing the gene of pyruvate carboxylase from this strain. For that purpose we used regions of homology from pyruvate carboxylase sequences of 12 different species (corresponding to catalytic and binding sites of the enzyme) to design polymerase chain reaction (PCR) primers for amplifying a fragment of the pyruvate carboxylase (*pc*) gene from *C. glutamicum* genomic DNA. An 850 bp fragment that was obtained with this approach was sequenced and found to have a significant homology to the pyruvate carboxylase gene of *Mycobacterium tuberculosis*. This 850 bp fragment was used to probe a *C. glutamicum* cosmid library and four candidate *pc* cosmids were identified. The complete sequence of the gene was obtained by directly sequencing one of the four cosmids by several rounds of primer synthesis and sequencing. The *C. glutamicum pc* sequence shows 64% homology with the *pc* gene of *M. tuberculosis* and 44% homology with the human *pc* gene. Regions of ATP, pyruvate and biotin binding have also been identified.

2.A.2 Genetic background on pyruvate carboxylase

Pyruvate carboxylase is one of the biotin-dependent enzymes that catalyze the carboxylation of pyruvate to form oxaloacetate (Attwood 1995) (Scrutton 1978). It has an anaplerotic function and helps replenish the oxaloacetate consumed for biosynthesis.

The two-step reaction mechanism catalyzed by pyruvate carboxylase is the following:

$$\text{Enzyme-biotin} + \text{ATP} + \text{HCO}_3^- \xrightarrow{\text{Mg}^{2+}} \text{Enzyme-biotin-CO}_2^- + \text{ADP} + \text{Pi} \quad (1)$$

$$\text{Enzyme-biotin-CO}_2^- + \text{pyruvate} \longrightarrow \text{Enzyme-biotin} + \text{oxaloacetate} \quad (2)$$
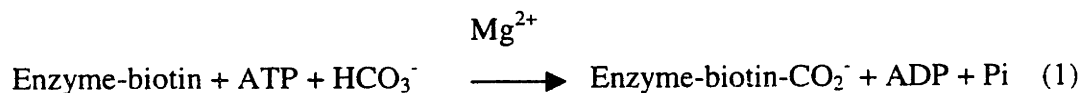
In reaction (1) the ATP-dependent biotin carboxylase domain carboxylates a biotin prosthetic group linked to a specific lysine residue in the biotin-carboxyl-carrier protein (BCCP) domain. Acetyl- coenzyme A activates reaction (1) by increasing the rate of bicarbonate dependent ATP cleavage. In reaction (2), the BCCP domain donates the CO_2 to pyruvate in a reaction catalyzed by the transcarboxylase domain (Attwood 1995). In the metabolism of carbohydrates, this two-step reaction is used for continuous replenishment of the tricarboxylic acid (TCA) cycle for anabolic purposes, such as amino acid synthesis.

To date, pyruvate carboxylase genes have been cloned and sequenced from four prokaryotes: *Rhizobium etli* (Dunn, Encarnacion et al. 1996), *Bacillus stearothermophilus* (Kondo, Kazuta et al. 1997), *Bacillus subtilis* (Genbank accession no. Z97025), and *Mycobacterium tuberculosis* (Genbank accession no. Z83018). Pyruvate carboxylase activity has been measured previously in *Brevibacterium lactofermentum* (Tosaka, Morioka et al. 1979) and *Corynebacterium glutamicum* (Peters-Wendisch, Wendisch et al. 1997). Whereas pyruvate carboxylase plays the major anaplerotic role in vertebrate tissues, especially in liver and kidney (Kornberg 1966) (Scrutton and Young 1972) (Jitrapakdee, Booker et al. 1996) and in yeast (Ruiz-Amil, De Torrontegui et al. 1965) (Stucka, Dequin et al. 1991), only a few prokaryotes, e.g. some *Bacillus* strains and

Rhodobacter capsulatus, use pyruvate carboxylase as the sole anaplerotic enzyme (Cazzulo, Sundaram et al. 1970) (Diesterhaft and Freese 1973) (Sundaram 1973) (Modak and Kelly 1995). In many bacteria e.g. enteric bacteria, cyanobacteria and streptomyces, another oxaloacetate forming enzyme, phosphoenolpyruvate carboxylase is the only enzyme used for replenishing the TCA cycle during growth on glucose (Chao and Liao 1993) (Owttrim and Coleman 1986) (Dekleva and Strohl 1988). In some other bacteria, such as *Pseudomonas citronellolis, P. fluorescens, Azotobacter vinelandii* and *Rhizobium etli*, both pyruvate carboxylase and phosphoenolpyruvate carboxylase have been detected (Higa, Milrad de Forchetti et al. 1976) (O'Brien, Chuang et al. 1977) (Liao and Atkinson 1971) (Scrutton and Taylor 1974) (Dunn, Encarnacion et al. 1996). However so far the physiological role of the one anaplerotic enzyme or the other in these bacteria has not been studied in detail.

Previous research has indicated that the yield and productivity of the aspartate family of amino acids depends critically on the carbon flux through anaplerotic pathways (Vallino and Stephanopoulos 2000). On the basis of metabolite balances, it can be shown that the rate of lysine production is less than or equal to the rate of oxaloacetate synthesis via the anaplerotic pathways. It has also been shown that lysine production is unaffected in a phosphoenolpyruvate carboxylase deficient mutant (Gubler, Park et al. 1994) (Peters-Wendisch, Eikmanns et al. 1993). However, lysine production is reduced in a pyruvate kinase deletion mutant, consistent with the viewpoint that pyruvate carboxylation is a significant anaplerotic activity leading to oxaloacetate production. In a phosphoenolpyruvate carboxylase-pyruvate kinase double mutant, lysine production is further decreased to 25% of the wild type levels. In this case, the only route for pyruvate formation is via the phosphoenolpyruvate: glucose transferase system, which produces one molecule of pyruvate for every molecule of glucose imported. In the double mutant there is less pyruvate available for the synthesis of oxaloacetate via pyruvate carboxylation and this limits the production of lysine (Park, Sinskey et al. 1997a). The presence of pyruvate carboxylation activity was also strongly suggested by ^{13}C labeling studies in *C. gutamicum*. In these experiments, carried out with *C. glutamicum* phosphoenolpyruvate carboxylase, pyruvate kinase single mutants and the

phosphoenolpyruvate carboxylase-pyruvate kinase double mutant, [1-^{13}C] pyruvate or [2-^{13}C] pyruvate was added to the fermentor and the fractional ^{13}C enrichment at individual carbon positions of lysine was measured. The pattern that was observed cannot be explained by sole operation of phosphoenolpyruvate carboxylation and supports the model that *C. glutamicum* has a pyruvate-carboxylating activity (Park, Shaw-Reid et al. 1997b). In order to understand fully the role of carbon flux through the anaplerotic pathway, it is necessary to clone and sequence the *C. glutamicum* pyruvate carboxylase gene as it has potential importance in increasing lysine and glutamic acid production.

2.B Materials and Methods

2.B.1 Bacterial strains, plasmids and media

C. glutamicum 21253 (*hom*⁻, lysine overproducer) was used for *in vitro* enzymatic assays and the preparation of chromosomal DNA. *Escherichia coli* DH5α (*hsdR*⁻, *recA*⁻) (Hanahan 1983) was used for transformations.

Plasmid pCR2.1 TOPO (Invitrogen) was used for cloning Polymerase Chain Reaction (PCR) products. The plasmid pRR850 was constructed from pCR2.1 TOPO and contained an 850 bp PCR fragment cloned in the pCR2.1 TOPO plasmid.

E. coli was grown in Luria-Bertani (LB) medium at 37 °C (Sambrook, Fritsch et al. 1989). *C. glutamicum* was grown in glucose minimal medium (Table 2.3) or LB medium at 30 °C. Where noted, ampicillin was used at the following concentrations: 100 μg/ml in plates and 50 μg/ml in liquid culture.

2.B.2 Protein manipulations

Preparation of Cell-Free Crude Extracts

For determination of *in vitro* enzymatic activities, cells grown in various media and phases were harvested by centrifugation at 4,000 g and 4°C for 15 min and the cell pellets were kept frozen, if necessary, at -20°C until use. The cell pellets were thawed on

ice and washed twice with cold buffer (100 mM Tris·HCl pH 7.5, 20 mM KCl, 5 mM MnSO$_4$, 10 mM MgCl$_2$). The cell pellets were resuspended in the same buffer to a final cell concentration between 20 to 50 g DCW/l. Cells were disrupted by homogenizing with glass beads (200-300 μm, Sigma Chemical Co.) in 5100 Bead Mill (SPEX Industries, Edison, NJ) for 5 min. The cell debris was removed by centrifugation at 47,000 g and 4°C for 1h. The supernatant was used for the enzyme measurements. This method produced typically 5 to 10 mg of protein per ml of supernatant. The protein concentrations were determined by the BCA method (Pierce Chemicals) using a bovine serum albumin as a standard, as instructed by the manufacturer.

Enzymatic assays

Enzymatic assays were performed in methylacrylate cuvettes (2.9 ml volume with 1 cm light path) in a Uvikon 810 spectrophotometer at room temperature. Most reactions were carried out in small volumes (1-2 ml) and activity was calculated typically in nmol/min/mg of protein using the following expression:

$$\text{Activity [nmol/min/mg of protein]}= \frac{1000V}{\varepsilon LPY}\frac{dA}{dt}$$

where V is the volume (in ml) of the assay mixture, L is the path length of cuvette (1 cm), P is the protein concentration of the cell-free extract (in mg/ml), Y is the volume of cell-free extract added (in ml), and dA/dt is the rate change of absorbance at the desired wavelength (in ΔOD/min). The extinction coefficients (ε) used were 6.22 mM^{-1}cm^{-1} for NADPH and NADH at 340 nm and 14.12 for thiophenolate at 412 nm.

Pyruvate Carboxylase enzymatic assay

Two different protocols were used to measure the activity of pyruvate carboxylase *in vitro*: one coupled with citrate synthase and one coupled with malate dehydrogenase.

The first protocol was based on the determination of absorbance increase at 412 nm due to the formation of thiophenolate using 5,5' dithio-bis (-2 nitorbenzoic acid)

(DNTB, Ellman's reagent) (Riddles, Blakeley et al. 1978) based on the following reactions:

$$\text{Pyruvate} + CO_2 \xrightarrow{\quad pc \quad} \text{Oxaloacetate} + \text{ATP}$$

$$\text{Oxaloacetate} + \text{Acetyl-CoA} \xrightarrow{\quad citrate-synthase \quad} \text{Citrate} + \text{HSCoA}$$

$$\text{HSCoA} + \text{DNTB} \xrightarrow{\quad\quad} \text{Thiophenolate}$$
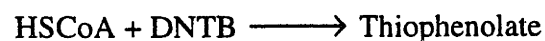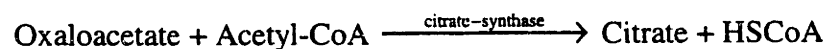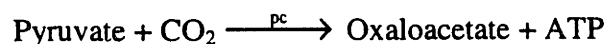
The reaction mixture contained in a volume of 1 ml: 100 mM Tris·HCl pH 7.5, 10 mM $MgCl_2$, 10 mM pyruvate, 50 mM $KHCO_3$, 0.2 mM acetyl-CoA, 10 untits of citrate synthase, 0.1 mg DNTB (from 10 mg/ml stock prepared in ethanol), and 250-500 µg protein crude extract. Blanks without pyruvate, bicarbonate, acetyl-CoA, ATP and crude extract were run.

The second protocol was based on the following reactions (Tosaka, Morioka et al. 1979):

$$\text{Pyruvate} + CO_2 \xrightarrow{\quad PC \quad} \text{Oxaloacetate} + \text{ATP}$$

$$\text{Oxaloacetate} + \text{NADH} \xrightarrow{\quad malate-dehydrogenase \quad} \text{Malate} + \text{NAD}^+$$
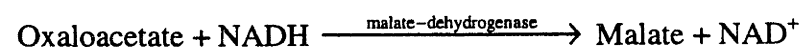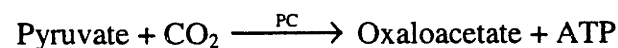
The activity was determined by the decrease in absorbance at 340 nm due to the oxidation of NADH in a reaction mixture containing 100 mM Tris·HCl pH 7.5, 10 mM $MgCl_2$, 10 mM pyruvate, 50 mM $KHCO_3$, 0.2 mM acetyl-CoA, 0.4 mM NADH, 10 units of malate dehydrogenase, and 250-500 µg of protein crude extract in a volume of 1 ml. Blanks without pyruvate, bicarbonate, NADH and protein crude extract were run.

The first protocol is more specific and sensitive than the first one since the background activity of NADH oxidation reactions does not interfere with the measurement.

37

SDS-Page electorphoresis

For separating proteins in sodium dodecyl sulfate (SDS) polyacrylamide gels, the protein crude extract is first diluted at a ratio 1:5 in sample buffer (1 ml 0.5 M Tris-HCl pH 6.8, 0.8 ml glycerol, 1.6 ml 10% SDS, 0.4 ml β-mercaptoethanol, 0.5% bromophenol blue, final volume 8 ml). The mixture is next boiled for 5 min.

Separately, the electrophoresis unit- a BioRad mini gel electrophoresis apparatus- is assembled as described by the manufacturer. SDS-Page gels used are purchased from BioRad (BioRad, Hercules CA) with the acrylamide concentration depending on the desired resolution. Both the upper and lower chambers of the apparatus are filled with running buffer (5× running buffer: 15 gr Tris-base, 72 gr glycine, 5 gr SDS in a final volume of 1 lt). After loading the protein sample on the well, the electrophoresis is run at approximately 200 V until the desired resolution is obtained. After the end of the electrophoresis, the gel is removed and processed further (staining or western blot transfer).

Western Blot

The protein transfer from the polyacrylamide gel to PVDF or nitrocellulose membranes is performed as described by Towbin et al. (Towbin, Staehelin et al. 1979). Initially the gel together with all other necessary material (sponges and filter paper) is equilibrated in transfer buffer (25 mM TrisBase, 192 mM glycine and 20% methanol). A sandwich is then prepared with the gel placed between the transfer membrane and filter paper. This sandwich is next placed between the sponges and after removing all air bubbles it is placed in the transfer apparatus. The apparatus is next filled with transfer/blotting buffer and the protein transfer is performed at 22 V overnight.

The membrane is next removed and incubated with 30 ml of blocking buffer (10% BSA or 2% casein) for 40 min. 20 µl of avidin-alkaline phosphatase conjugate is diluted in 30 ml of blocking buffer and the membrane is then placed in that solution for another 40 min. Next follow three washes of the membrane with 30 ml (each wash) of TTBS buffer (100 mM Tris-HCl, 150 mM NaCl, 0.1% Tween-20, pH 7.5), 10 min each. At the

final step, biotinylated proteins are detected using a cheminescence detection solution of 5-bromo-4-chloro-3-indoylphosphate-p-toluidine salt and nitroblue tetrazodium chloride purchased from Schleicher&Schuell. Color usually appears after a 10 min incubation of the membrane with 30 ml of this solution. To step the color development the membrane is washed with 50 ml water for 10 min.

2.B.3 DNA manipulations

C. glutamicum genomic DNA isolation

Cells grown overnight (500 ml culture on LB) are harvested by centrifugation at 8,000rpm for 10min in GSA rotor or equivalent (~10,000 g). Cells are then resuspended in 5 ml TE buffer (final volume ~10 ml) and transferred to 50 ml centrifuge tubes.

To this suspension, 20 mg lysozyme and 200 µl of mutanolysin solution (1 mg/ml) is added and the suspension is left to incubate for 1 hour at 37 °C while gently shaking. After the incubation, 1 ml of 0.5 M EDTA, 1 ml of 10% SDS and 1 ml of 5 M NaCl are added and the suspension is incubated on ice for 10 min.

2 mg of proteinase K (from a 20 mg/ml stock) is then added and a second incubation at 37° C follows for 60 min. 3.77g of sodium perchlorate is added to the cell slurry (final concentration of 1 M) and after gently mixing with the pipet the mixture is left to incubate for 30 min at room temperature.

A protein extraction follows using phenol: chloroform: isoamyl alcohol (25:24:1) and a centrifugation for 20 min at 5,000 rpm in GSA rotor or equivalent. A second extraction of the aqueous phase follows with 0.5 volumes of chloroform: isoamyl alcohol (24:1) and centrifugation for 20 min at 5,000 rpm (GSA rotor).

Two volumes of ice-cold ethanol are subsequently added to the aqueous phase and the DNA is spooled onto a glass rod and transferred to a new tube containing 5 ml of TE buffer. There the DNA is left to dissolve by rocking overnight at 4 °C. The RNA is removed by adding RNase to the DNA solution and incubating at 37 °C for 30 min. Rnase is further removed by a similar phenol: chloroform: isoamyl alcohol extraction as described previously, followed by a chloroform: isoamyl alcohol extraction.

To the aqueous phase isolated 1/10 volume of 3 M sodium acetate (pH 5.3) is added and two volumes of ice-cold ethanol. The DNA is spooled again on a glass rod and transferred to a 70% ethanol solution for 3 min. Finally the DNA is air dried for 5-10 min and is dissolved to a minimal volume of TE (0.5-1 ml).

This protocol is an adaptation from the original protocol published by Tomioka *et al.* (Tomioka, Shinozaki et al. 1981).

Dot Blot analysis

The dot blot analysis used is a modified version of the protocol published by Boehringer Manheim (Roche Biochemicals Inc.).

A DNA sample (probe) is prepared by standard Polymerase Chain Reaction (PCR) techniques and is resuspended (10 ng to 1 μg) in 16 μl TE buffer. It is subsequently denatured by boiling for 10 min and quickly chilled on ice for 5 min in order to prevent annealing. 4 μl of DIG-High Prime (label) is added, mix and centrifuge briefly. The labeling reaction takes place by overnight incubation at 37 °C. The reaction is stopped by adding 2 μl of 0.2 M EDTA (pH 8.0).

For the probe purification, an equal volume (22 μl) of pre-chilled (-20 °C) isopropanol is added to the previous mixture and incubated for one hour at –70 °C. The mixture is then spinned at 14,000 rpm for 15 min and the supernatant carefully removed. The pellet is then resuspended with 50 μl of cold 70% ethanol and again centrifuged at 14,000 rpm for 15 min. The DNA is dried under vacuum and resuspended in 50 μl of TE buffer. The labeled DNA probe can be stored safely at –20 °C.

For the cosmid preparation, an appropriate amount of DNA/ H_2O is transferred into eppendorff tube and heated at 95 °C for 5 min. The DNA is removed promptly and chilled on ice immediately. 5 μl of 20×SSC plus bromophenol blue as marker is then added (~ 10 μl dye/ ml solution).

The cosmid samples are then loaded onto a *Dot- Blotting manifold* loaded with a nylon membrane (the upper left corner of the membrane is labeled with a pencil for orientation recognition). Vacuum is applied for 2 min. The membrane is removed and denaturation follows.

The membrane is placed (DNA side up) to a wad of Whatman 3 MM paper steeped in denaturing solution (0.5 M NaOH/1.5 M NaCl). After a 5 min waiting period the membrane is transferred to a wad of Whatman 3 MM paper steeped in neutralization buffer (0.5 M Tris/ 1.5 M NaCl) for 1 min.

The membrane is subsequently rinsed briefly in 2× SSC buffer and dried over Whatman 3 MM paper. After wrapping it in Saran wrap the DNA is fixed to the membrane by exposure to UV light for 2 min each side.

A prehybridization step follows by placing the membrane in a plastic bag and adding 30 ml of hybridization buffer (5 × SSC, 0.1% N-lauroylsarcosine, 0.02% SDS, 1/10 volume 10× blocking reagent provided by the manufacturer). The sealed bag is then incubated at 65 °C for two hours. This step is necessary for blocking non-specific nucleic acid binding sites.

The DIG- labeled DNA probe is denatured as described previously and added to 30 ml prewarmed (65 °C) hybridization buffer. The mixture is next added to a bag containing the membrane with the immobilized cosmids and the hybridization takes place by overnight incubation at 65 °C.

The membrane is then washed twice with 100 ml of 2×SSC, 0.1% SDS, 15 min each wash, at room temperature. A second wash follows with 0.1 ×SSC, 0.1% SDS again at 65 °C under constant agitation. The SSC percentage can vary from 0.1× for highest stringency to 0.5 ×.

The final step involves the signal development and detection. Initially the membrane is washed with 100 ml washing buffer (0.5 M maleic acid, 0.75 M NaCl, 1.5% Tween 20, pH 7.5) and next incubated for 30 min in 100 ml blocking solution (provided by the manufacturer, Roche Biochemicals). 2 μl of anti-DIG- Alkaline Phosphatase conjugate is added to 30 ml of blocking buffer, generating the antibody solution, with which the membrane is then incubated for 30 min at room temperature. The membrane is next washed twice, 15 min each, with 100 ml washing buffer and equilibrated for 2-5 min in 20 ml alkaline phosphatase buffer (0.1 M Tris-HCl, 0.1 M NaCl, 50 mM $MgCl_2$, pH 9.5). About 0.5 ml (20 drops) of CSPD is then applied on the membrane and incubation follows for 10 min at 37 °C to enhance the luminescent reaction. Finally the signal is

detected by exposing the membrane to x- ray film (Kodak BIOMAX) at room temperature for more than two hours. The time of the exposure depends on the desired signal amount.

DNA sequencing

Automated DNA sequencing was performed by the MIT Biopolymers facility employing an ABI Prism 377 DNA sequencer.

Sequence analysis

The program DNA Strider Version 1.0 (Institut de Recherche Fondamentale, France) was used to invert, complement and translate the DNA sequence and find open reading frames in the sequence. The BLAST program (Altschul, Gish et al. 1990) from the National Center for Biotechnology Information (NCBI) was employed to compare protein and DNA sequences. Homology searches in proteins were done using the MACAW software (NCBI). PCR primers were designed with the aid of the Primer Premier software from Biosoft International. The compute pI/MW tool on the ExPasy molecular biology server (University of Geneva) was used to predict the molecular mass and isoelectric point of the deduced amino acid sequence.

Other DNA manipulations

Plasmid and cosmid DNA minipreparations were performed using Qiaprep spin columns following standard procedures described by the manifacturer (Qiagen). DNA was extracted from agarose gels with the Qiaex kit (Qiagen).

For large-scale high purity preparation of cosmid DNA for sequencing, the Promega Wizard kit was used following the procedures described by the manufacturer (Promega).

Standard heat shock techniques were used for *E. coli* transformation and agarose gel electrophoresis (Sambrook, Fritsch et al. 1989).

2.C Results

2.C.1 Identification of biotinylated proteins in *C. glutamicum*

Cell-free extracts of *C. glutamicum* wildtype grown in LB rich medium and glucose minimal medium were assayed for biotinylated proteins by SDS-PAGE and Western blot analysis. Two biotin containing proteins were detected (Figure 2.1). One band, located at approximately 80 kDa, has been identified as the biotin-carboxyl-carrier domain (BCCP) of acetyl-CoA-carboxylase (Jager, Peters-Wendisch et al. 1996). The second band, at 120 kDa, was expected to be the biotinylated pyruvate carboxylase subunit, as these subunits are in the range of 113-130 kDa (Attwood 1995).

2.C.2 *In vitro* enzymatic activities

Cell-free extracts of *C. gluamicum* grown in rich medium and in glucose minimal medium were tested for pyruvate carboxylase activity. Using the citrate-synthase-coupled assay a pyruvate carboxylating activity was observed of about 2 nmol/min/mg of protein. Even though this activity was pyruvate, bicarbonate and protein-dependent, the presence or absence of ATP seemed to have no effect on the enzymatic activity observed. The presence of ATP in the crude extract was verified using luminescence detection. Two methods for its removal were employed: the first involved the use of ATPase enzymes while the second the passage of the crude extract through a gel filtration column. In both cases there was no success in uncoupling the observed activity from a high ATP background. Various other conditions were also tested in order to establish their effect on the observed enzymatic activity (cold inactivation, aspartate inhibition etc.). When the malate-dehydrogenase-coupled assay was used no pyruvate carboxylating activity was detected.

It is possible that the weak enzymatic activity detected *in vitro* is due to the function of various gluconeogenic enzymes that exist in *C. glutamicum* such as oxaloacetate decarboxylase and malate dehydrogenase. None of these enzymes requires

the presence of ATP. The enzymatic measurement of pyruvate carboxylase activity was finally achieved by the use of *C. glutamicum* permeablized cells instead of protein crude extracts (Peters-Wendisch, Wendisch et al. 1997). At that point our efforts to prove the presence of pyruvate carboxylase in *C. glutamicum* were directed towards the discovery of the gene of this enzyme rather than establishing an *in vitro* enzymatic activity.

2.C.3 Cloning of a fragment of the gene of pyruvate carboxylase

The fact that pyruvate carboxylase belongs to the family of the biotin dependent enzymes facilitated the research effort to actually clone the gene of this enzyme. Highly conserved regions in previously cloned genes were used for that purpose. Almost all the pyruvate carboxylases that have been sequenced to-date contain a biotin carboxylase domain in their N- terminal region, a BCCP (biotin binding) domain in their C- terminal region and a transcarboxylase domain, with a binding site specific for pyruvate, in their central region (Dunn, Encarnacion et al. 1996).

In order to obtain a fragment of the gene of pyruvate carboxylase, a number of pyruvate carboxylase genes from various organisms were examined. A very strong similarity was established, on the amino acid level, among the motifs that correspond to an ATP binding site, the region around the pyruvate binding site and, of course, the sequence flanking the biotin attachment site. This conserved amino acid sequence extended to eight amino acids in the case of the ATP binding motif (YFIEXNXR), seven amino acids to a region adjacent to the pyruvate binding motif (ATFDVXX) and five amino acids to the biotin attachment region (AMKMX). Based on those similarities PCR primers were designed corresponding to those amino acid sequences as well as the region that corresponds to the pyruvate-binding site itself. Where the amino acids were different the primers were designed on the basis of *M. tuberculosis* because of its close relationship to *C. glutamicum* (both belong to the Actinomyces family, together with *Nocardia* and *Rhodococci*). The sequences of some of the primers used are given in Table 2.1.

Various combinantions of those oligonucleotides were next tried as PCR primers, using *C. glutamicum* genomic DNA as a template and *Taq* polymerase as the DNA

polymerase enzyme. The combination of the ATP corresponding primers with the pyruvate corresponding primers was expected to yield a DNA fragment of approximately 850 bp, based on the *M. tuberculosis* pyruvate carboxylase gene. In a similar way the combination of the pyruvate corresponding primers with the biotin corresponding primers was expected to yield a 1,550 bp DNA fragment (Figure 2.2).

A single 850 bp PCR band was produced when an oligonucleotide corresponding to the ATP binding site was used as the beginning primer and an oligonucleotide corresponding to a region close to the pyruvate binding site was used as the end primer. The beginning primer's sequence was: 5'- GTCTTCATCGAGATGAATCCGCG-3' while the sequence of the end primer was: 5'-CGCAGCGCCACATCGTAAGTCGC-3'. The PCR conditions were the following:

PCR Amplification of Pyruvate Carboxylase gene fragment

(A)   Mix the following reagents in a 0.2 ml PCR tube (in order):

Sterile distilled H_2O	65.0 μl
PCR Buffer	10.0 μl
dNTP mix (10 mM each)	2.0 μl
beginning primer (20 μM)	5.0 μl
end primer (20 μM)	5.0 μl
chromosomal DNA	10.0 μl (1.12 μg)
$MgCl_2$ (1 mM)	1.5 μl
Taq Polymerase	1.0 μl

(B)   Calculate Melting Temperature (Tm) for primers:

Tm= 67.5 + 34(decimal %GC content in primer)- (395/total# bases)

(C)   Carry out thermal cycling program:

Denaturing: 1minute at 94 °C

Repeat 35 times   1 minute at 94 °C

2 minutes at annealing temperature

3 minutes at 72 °C

Elongation: 10 minutes at 72 °C

(D)   Store amplified DNA at 4 °C.

The annealing temperature used for this amplification was 58 °C. This 850 bp band also appeared with other combinations of PCR primers derived from the ATP and pyruvate binding sites but not as a single band.

A single combination of a pyruvate-binding-site primer and a biotin-binding-site primer gave a band of molecular weight of approximately 1,550 bp. This fragment did not appear as a single DNA band but was amplified together with another three DNA bands of higher and lower molecular weights. The beginning primer for this PCR, (pyruvate binding site) was: 5'- GCGTCGCGGTGTAGCATTCAGCG-3' while the end primer, corresponding to the biotin binding site was: 5'-CGCTACTTCTACCTT-3'. The PCR conditions used for this specific PCR reaction were the same as the ones mentioned above with the exception of the annealing temperature that was lowered to 47 °C.

Both the 850 bp and the 1550 bp fragments were extracted from an agarose gel and cloned in the pCR2.1 TOPO vector (Invitrogen) according to the protocol provided by the manufacturer. The introduction of the PCR fragment into this vector takes advantage of the presence of nucleotide overhangs generated by *Taq* polymerase. Two plasmids were constructed with this method, pRR850 containing the 850 bp PCR fragment and pRR1550 containing the 1550 PCR fragment. The presence of the PCR fragments in these constructs was verified by restriction digests as well as PCR reactions using the same primers as above and the plasmids as templates.

2.C.4 Isolation of cosmids containing pyruvate carboxylase gene

Using primers M13 forward and M13 reverse that correspond to the region adjacent to the beginning and the end of the inserted PCR fragments the sequences of both the 850 bp as well as the 1550 bp PCR fragments were obtained. One of the open reading frames derived from the sequence of the 850 bp fragment gave a significant homology to the pyruvate carboxylase protein from *M. tuberculsosis* of about 60%. This is a clear indication that this 850 bp DNA fragment is part of the gene of pyruvate carboxylase from *C. glutamicum*, as it was expected based on the size of the fragment. On the other hand the sequence of the 1550 bp fragments gave no significant homology to any of the known pyruvate carboxylases that had been sequenced up to that point. Obviously this band was derived from a non-specific binding of the PCR primers to the *C. glutamicum* genomic DNA.

The 850 base pair fragment containing a portion of the *C. glutamicum* pyruvate carboxylase gene was used to probe a *C. glutamicum* genomic library. This cosmid library was prepared using standard procedures provided by Stratagene (Stratagene Inc., La Jolla CA) and employing superCos as the cloning cosmid. The cosmid library was a kind gift from Dr Phil Lessard, Department of Biology, MIT.

In the first round of screening, using the dot blot analysis method, 17 out of 291 cosmids appeared positive (Figure 2.3). This initial screening was done under non-stringent conditions such as low hybridization temperature (50 °C) and extended exposure time of the x-ray film (overnight). Because of the high number of positive cosmids a second round of screening was performed, even though the library construction method does not exclude the possibility of multiple cosmids containing the same piece of genomic DNA. The same probe was again used for the screening but this time the hybridization temperature was raised to 65 °C and the exposure of the x-ray film was lowered to half an hour. In order to verify the successful performance of the dot blot analysis a positive control was always present. This control consisted of an unlabeled 850 bp PCR fragment. The second round of screening yielded four cosmids with a positive signal (Figure 2.4). To confirm that these cosmids indeed contained the pyruvate

carboxylase gene, a PCR was performed using the four positive cosmids as templates and the same primers used to generate the 850 base pair probe from the *C. glutamicum* chromosomal DNA. The PCR conditions were identical to the ones described previously with the exception of the amount of DNA used during the reaction, which was reduced to about 500 ng. An 850 bp fragment was amplified from all four positive cosmids, which were designated as IIIF 10, IIE9, IIIG7 and IIIB7.

2.C.5 Sequencing strategy

The use of cosmids as templates for sequencing has been reported in the past for DNA fragments derived from Actinomyces species such as *Rhodococci* (Treadway, Yanagimachi et al. 1999). A similar approach was followed and in the case of the pyruvate carboxylase gene from *C. glutamicum.*

On the basis of the sequence of the 850 bp PCR fragment previously isolated, primers Begrev1 and Endfor1 were designed and used to sequence outwards from the beginning and the end of the 850 bp portion of the pyruvate carboxylase gene. The sequences of these primers corresponded to regions close to the 5' and 3' ends of the 850 bp DNA fragment respectively and were designed such as to avoid the creation of hairpin loops and secondary structures of the oligonucleotides. Cosmid IIIF10 was selected on a random basis and used as the sequencing template. The sequence was continued by designing new primers in a similar way as the one chosen for the previous two and "walking" across the gene as shown in Figure 2.5. The sequences of the primers used for performing the pyruvate carboxylase sequence are given in Table 2.2. Overall four sequencing cycles were performed, the first three of them both on downstream and upstream regions while the last one only on the downstream region. Some complications arose during the sequencing process, most likely due to the high GC content of the gene, a case very typical in Actinomyces genomes.

2.C.6 Sequence analysis

During the sequence process described above, we managed to sequence 3637 bp
of cosmid IIF10 and identified a 3420 bp open reading frame, which is predicted to
encode a protein of 1140 amino acids. The deduced protein is 63% homologous to the *M.*
tuberculosis pyruvate carboxylase and 44% homologous to the human pyruvate
carboxylase, and the newly discovered gene was named *pc* on the basis of this homology.
The deduced protein has a predicted isoelectric point of 5.4 and molecular mass of 123.6
kDa, which is similar to the subunit molecular mass of 120 kDa estimated by
SDS/polyacrylamide electrophoresis and western blotting. The predicted translational
start site, based on homology to the *M. tuberculosis* sequence is a GTG codon, as has
been observed in other bacterial sequences (Stryer 1988) (Keilhauer, Eggeling et al.
1993). The DNA sequence is now available in GenBank and has been assigned the
accession number AF038548.

The amino-terminal segment of the *C. glutamicum* pyruvate carboxylase contains
the hexapeptide GGGGRG which matches the GGGG(R/K)G sequence that is found in
all biotin-binding proteins and it is believed to be an ATP- binding site (Post, Post et al.
1990) (Fry, Kuby et al. 1986). A second region that is proposed to be involved in ATP
binding and is present in biotin- dependent carboxylases and carbamylphosphate
synthetase (Lim, Morris et al. 1988) is conserved in the *C. glutamicum* sequence. The
predicted *C. glutamicum* pyruvate carboxylase protein also contains a putative pyruvate-
binding motif, FLFEDPWDR, which is conserved in the transcarboxylase domains of
Mycobacterium, Rhizobium and human pyruvate carboxylase (Dunn, Encarnacion et al.
1996). Tryptophan fluorescence studies with transcarboxylase have shown that the Trp
residue present in this motif is involved in pyruvate binding (Kumer, Haase et al. 1988).
The carboxyl- terminal segment of the enzyme contains a putative biotin- binding site,
AMKM, which is identical to those found in other pyruvate carboxylases as well as the
biotin-carboxyl-carrier protein (BCCP) domains of other biotin- dependent enzymes.

2.D Discussion

Despite the importance of pyruvate carboxylase in the central metabolism of many organisms, only a few genetic studies on this enzyme have been performed. By the time that the sequencing of pyruvate carboxylase from *C. glutamicum* was accomplished, pyruvate carboxylase genes had been isolated, sequenced and characterized only from the eukaryotic organisms man, mouse, rat and *Saccharomyces cerevisiae,* and from the prokaryotes *R. etli,* and *Bacillus stearothermophilus* (Wexler, Du et al. 1994) (Zhang, Xia et al. 1993) (Jitrapakdee, Booker et al. 1996) (Lim, Morris et al. 1988) (Stucka, Dequin et al. 1991) (Dunn, Encarnacion et al. 1996) (Kondo, Kazuta et al. 1997). Additionally, the pyruvate carboxylase sequences of the mosquito *Aedes aegypti* and of *M. tuberculosis* are available in GenBank Data Library (Tu and Hagerdorn, 1994; Smith, 1994).

Pyruvate carboxylase belongs to the biotin containing enzymes, which include acyl-CoA carboxylases, oxaloacetate decarboxylases and transcarboxylases (Samols, Thornton et al. 1988) (Toh, Kondo et al. 1993). One of the most intriguing and interesting aspects in biotin-containing carboxylases is the similarity of their catalytic mechanism. This mechanism includes the carboxylation of a covalently attached biotin moiety and transcarboxylation from the carboxy-biotin to a specific acceptor molecule. This similarity in the past has led to the suggestion that organisms that contain multiple biotin-dependent enzymes might share common subunits (Alberts, Nervi et al. 1969) (Gerwin, Jacobsen et al. 1969). However at least in the cases of two organisms, this has not proved to be the case, *P. citronellolis* and *Bacillus* (Fall, Alberts et al. 1975) (Buckley, Libor et al. 1969).

Another more interesting idea by Lynen was based on the similarities of these enzymes in their reaction mechanisms and in their overall molecular properties. Lynen proposed that the biotin- dependent carboxylases might represent various stages in the evolution of an enzyme "family" from common ancestral genes (Lynen 1974) (Obermayer and Lynen 1976). Support for this attractive hypothesis is to be found in the high degree of sequence homology noted about the biotin attachment point in several

different enzymes from various sources. However this idea was disregarded for a long time based on the fact that very few sequences of those enzymes were available until recently and because there were no successful immunological data that could show that different biotin-containing enzymes share any common structural features. (Ballard, Hanson et al. 1970) (Sumper and Riepertinger 1972). However it is important to note that these studies were only concerned with these carboxylases in their native states. Since (1) changes in amino acid residues due to protein evolution are tolerated more readily at protein surfaces than in their interiors (Schulz and Schrimer 1979) and (2) antibodies react primarily if not exclusively with the amino acid side chains of surface residues (Schulz and Schrimer 1979), it would seem quite likely that failure to detect any immunological identity between the biotin carboxylases in their native states does not necessarily preclude the existence of common structural and functional features of these enzymes buries within the less rapidly evolving internal regions of these molecules.

The idea to look for homologies among various pyruvate carboxylases that had been sequenced at that point was based initially on the very well established high degree of sequence homology noted about the biotin attachment point (biocytin) in several different enzymes from various sources. If Lynen's evolution model was correct, we were expecting to find a very strong homology at the ATP binding site, which indeed was the case. There appears to be a consensus of the motif GGGG(R/K)G, not only among pyruvate carboxylases but many other enzymes that involve biotin not only as a cofactor but also as substrate (for example acetyl CoA carboxylase in the first case and biotin ligase in the second). The second ATP binding motif appears also to have a very strong consensus as mentioned earlier of YFIE (V/I) N as expected. This region was actually preferred for the design of PCR primers for the simple reason that it is closer to the second region that homology was sought, namely the pyruvate binding site.

Indeed the distance in all pyruvate carboxylases between the ATP binding motif and the biotin-binding motif (more than 2000 base pairs) made very difficult the exclusive use of these two binding sites for PCR purposes. Another binding motif was sought and for that purpose the pyruvate binding motif appeared to be a very attractive choice as it lies in between the previous two binding motifs in all pyruvate carboxylases

sequenced at that point. Again a very strong homology was identified in that region with a sequence FLXEDPW as well as in the region immediately upstream of that binding site with a sequence of ATFDV (A/S). In fact the latter region was better conserved than the pyruvate binding site itself, both on the protein as well as on the DNA level. This was the basis for selecting these regions as the intermediate regions between the ATP and the biotin binding sites. PCR primers were designed both as beginning primers (in combination with the biotin binding site primers) as well as end primers (in combination with the ATP binding site primers).

The first attempt to isolate a pyruvate carboxylase DNA fragment by PCR utilizing these primers, was by using a combination of pyruvate and biotin-binding-sites corresponding primers. One of these combinations succeeded to yield a DNA fragment with an approximate size of 1600 base pairs, similar to our expectations. This band was not unique, even when the annealing temperature was increased to 60 ° C. By subcloning this PCR fragment into pCR2.1 TOPO vector we were able to obtain its full DNA sequence utilizing primers M13 forward and reverse. Since none of the open reading frames gave any significant homology to other pyruvate carboxylase proteins, the combination of these sets of primers was disregarded.

The combination of ATP and pyruvate binding site primers was also able to yield PCR fragments of approximately 850 base pair, as expected based on the *M. tuberculosis* pyruvate carboxylase sequence. The primer corresponding to the pyruvate binding site gave a multitude of bands, even when the PCR conditions became more stringent by raising the annealing temperature and decreasing the amount of $MgCl_2$ present, while the primer corresponding to the adjacent region gave a single 850 base pair band even with less stringent conditions. It was exactly this single 850 bp DNA fragment that gave a strong homology to the *M. tuberculsosis* pyruvate carboxylase when cloned and sequenced in a similar way to the one mentioned above. The complete sequence of the gene of pyruvate carboxylase was finally obtained by directly sequencing one of the four positive cosmids detected to contain this gene. Phylogenetic analysis proved the close relationship between the *C. glutamicum* and the *M. tuberculosis* genes (Figures 2.6 and 2.7).

The sequencing of the gene of pyruvate carboxylase from *C. glutamicum* concludes extensive previous studies that had shown that phosphoenolpyruvate carboxylase is not the main anaplerotic enzyme for *C. glutamicum* since its absence does not affect lysine production (Gubler, Park et al. 1994) (Peters-Wendisch, Eikmanns et al. 1993). Moreover a number of studies have indicated the presence of a pyruvate-carboxylating enzyme, employing [13]C- labeling experiments and NMR and GC-MS analysis (Park, Shaw-Reid et al. 1997b) (Peters-Wendisch, Wendisch et al. 1996) or enzymatic assays with cell free extracts (Tosaka, Morioka et al. 1979) and permeable cells (Peters-Wendisch, Wendisch et al. 1997). However, due to a number of unsuccessful trials to reproduce some of the above-mentioned data, the finding of the gene of pyruvate carboxylase comes as the first direct and unequivocal evidence for the existence of this enzyme in *C. glutamicum*.

Apart from the AMKM sequence near the C-terminus of the *C. glutamicum* pyruvate carboxylase representing the biotin-binding site, the so called proline hinge (amino acid 1078 of the protein sequence) suggested to be necessary for the flexibility of the biotin carrier domain, and three glycine residues (positions 1090, 1096 and 1117) involved in the biotinylation of the enzymes can be found in the *C. glutamicum* pyruvate carboxylase (Leon Del Rio and Gravels 1994).

The independent study of this gene by another group showed that the transcript of pyruvate carboxylase has a size of about 3.5 kb, indicating that the gene is transcribed as a monocistronic message and is not organized as an operon (Peters-Wendisch, Kreutzer et al. 1998). A sequence similar to the −10 canonical sequence of *E. coli* that is common also in *C. glutamicum* promoters (with the exception of the *pheA* gene) (Follettie, Peoples et al. 1993) is also observed upstream of the pyruvate carboxylase gene (TACGAT) with four of six bases identical to the conserved sequence TA.aaT (Peters-Wendisch, Kreutzer et al. 1998). Sequence homology to the −35 motif of *E. coli* promoters is also observed (TTGATT) with three of six nucleotides identical to the consensus hexamer ttGcca, unlike other *C. glutamicum* promoters. The relatively low similarity to the consensus sequence might indicate that the pyruvate carboxylase promoter mediates a moderate expression of the pyruvate carboxylase gene and this corresponds to the relatively low

amount of the pyruvate carboxylase protein present in cell free extracts (Peters-Wendisch, Kreutzer et al. 1998). Finally a purine rich region (- is also present upstream of the –35 promoter motif, as has been previously document for *C. glutamicum* promoters (Follettie, Peoples et al. 1993).

There are precedents to the finding that *C. glutamicum* contains more than one enzyme to perform the anaplerotic function of regenerating oxaloacetate. *Pseudomonas fluorescens, Azotobacter vinelandii, P. citronellolis* and *Thiobacillus novellas* contain both phosphoenolpyruvate carboxylase and pyruvate carboxylase (O'Brien, Chuang et al. 1977) (Scrutton and Taylor 1974) (Milrad de Forchetti and Cazzulo 1976) (Charles and Willer 1984). *Zea mays* contains three isozymes of phosphoenolpyruvate carboxylase (Toh, Kawamura et al. 1994) and *Saccharomyces cerevisiae* contains two isozymes of pyruvate carboxylase (Brewster, Val et al. 1994), each differentially regulated.

With pyruvate, phosphoenolpyruvate and oxaloacetate occurring at one of the major crossroads of intermediary metabolism, it is not surprising that the enzymes catalyzing their interconversions have been subject of numerous investigations in a wide range of life forms (Keech and Wallace 1985).

Most cells have a combination of only three of these enzymes (namely pyruvate carboxylase, phosphoenolpyruvate carboxylase and pyruvate kinase) with which to negotiate this central triangle of metabolism. Some cells have four (e.g. *Salmonella typhimurium* and *E. coli*) have four-phosphoenolpyruvate carboxylase, phosphoenolpyruvate carboxytransphosphorylase, pyruvate kinase and phosphoenolpyruvate synthetase; *Azotobacter vinelandii* has pyruvate carboxylase, oxaloacetate decarboxylase, phosphoenolpyruvate carboxylase and pyruvate kinase; *P. citronellolis* has five- pyruvate carboxylase, oxaloacetate decarboxylase, phosphoenolpyruvate carboxylase, pyruvate kinase and phosphoenolpyruvate synthetase) (Theodore and Englesberg 1964) (Ashworth, Kornberg et al. 1965) (Scrutton and Taylor 1974) (O'Brien, Chuang et al. 1977). With the discovery of the existence of a pyruvate carboxylase gene in *C. glutamicum* the number of enzymes that participate in the phosphoenolpyruvate-pyruvate-oxaloacetate triangle in this organism rises to six (Figure 2.8). This presence of all six enzymes in one organism has not been reported previously

and leads to the formation of an interlocking system that would appear to require complex regulatory mechanisms to permit a proper flow of metabolites through the pathways and to prevent futile cycling. In fact, since five of these enzymes appear to be constitutive a very tight regulation through effectors would be expected to be important. Drawing a parallel with *P. citronellolis*, various metabolites would be expected to have reciprocal but reinforcing functions and possibly the expression of pyruvate carboxylase to be regulated by the carbon substrates of the growth medium. Such an expectation was strongly verified by the enzymatic assay results using permeabilized *C. glutamicum* cells that have been reported in literature (Peters-Wendisch, Wendisch et al. 1997).

Biochemical and genetic study of all six enzymes in coordination with other downstream activities may lead to the elucidation of the exact procedures necessary for maximizing the production of primary metabolites by this industrially important microorganism.
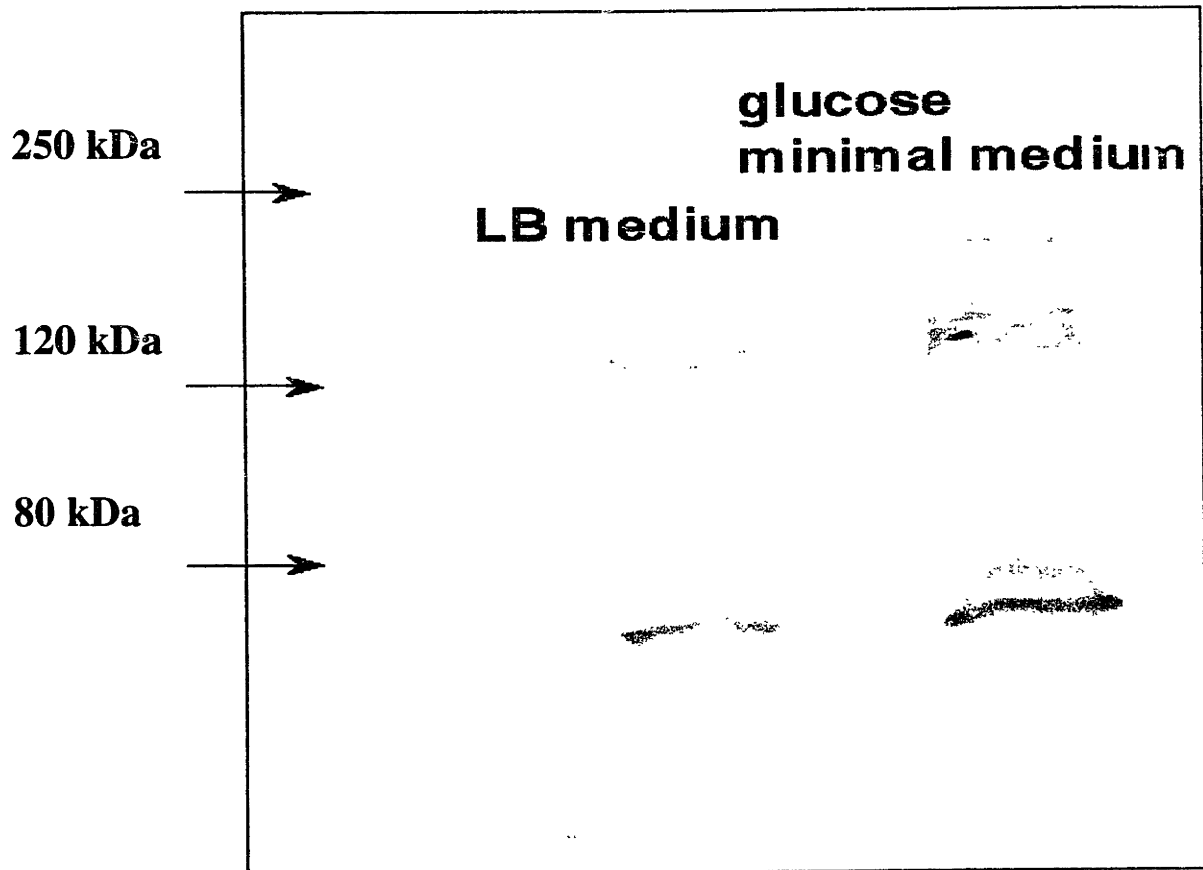
Figure 2.1 Western blot of SDS-PAGE of cell-free extracts (20 μg protein each) of *C. glutamicum* wildtype grown in LB rich medium (lane 1) and glucose minimal medium (lane 2). Protein standards are shown in lane 1.
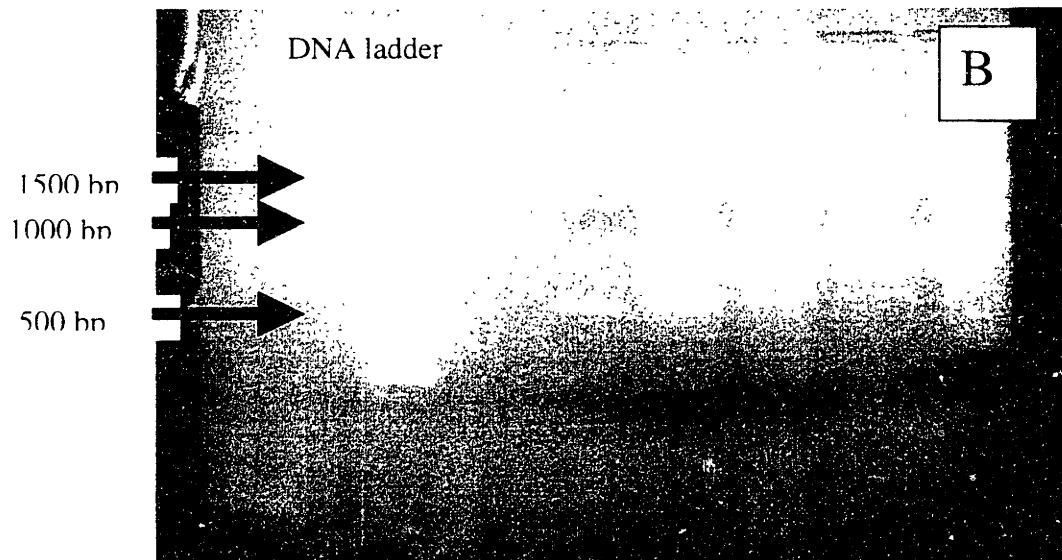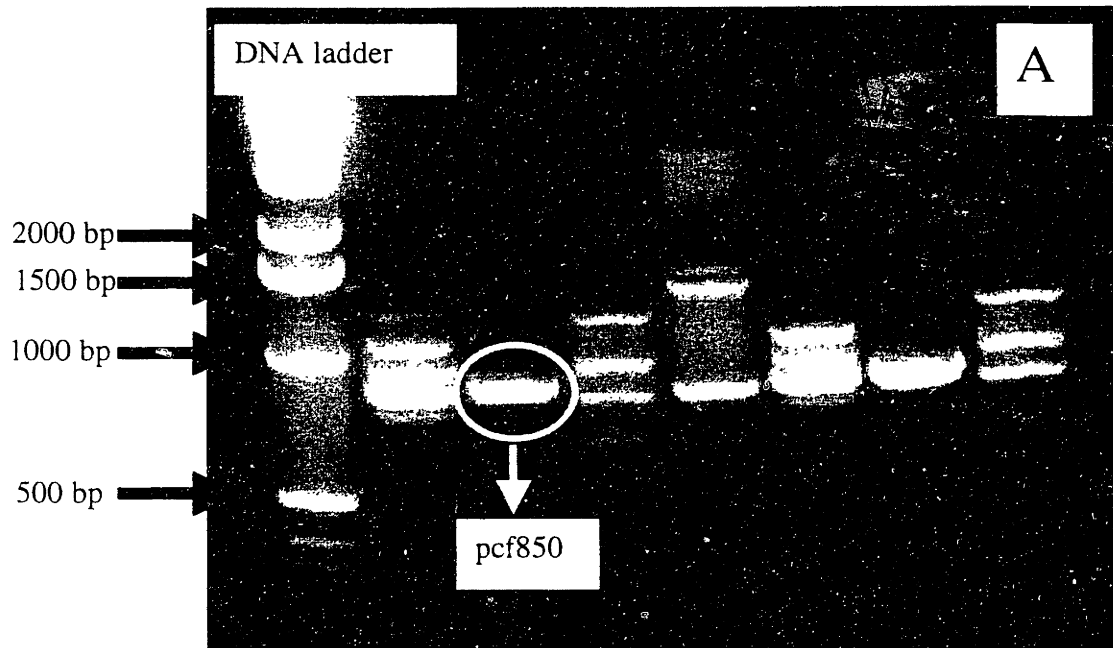
Figure 2.2 PCR results utilizing various combinations of primers with *C. glutamicum* genomic DNA as template. (A) PCR results using ATP-binding-site as beginning and pyruvate-binding-site as end primers. Band pcf850 was further used for cosmid library screening. (B) PCR results using pyruvate-binding-site as beginning and biotin-binding-site as end primers.
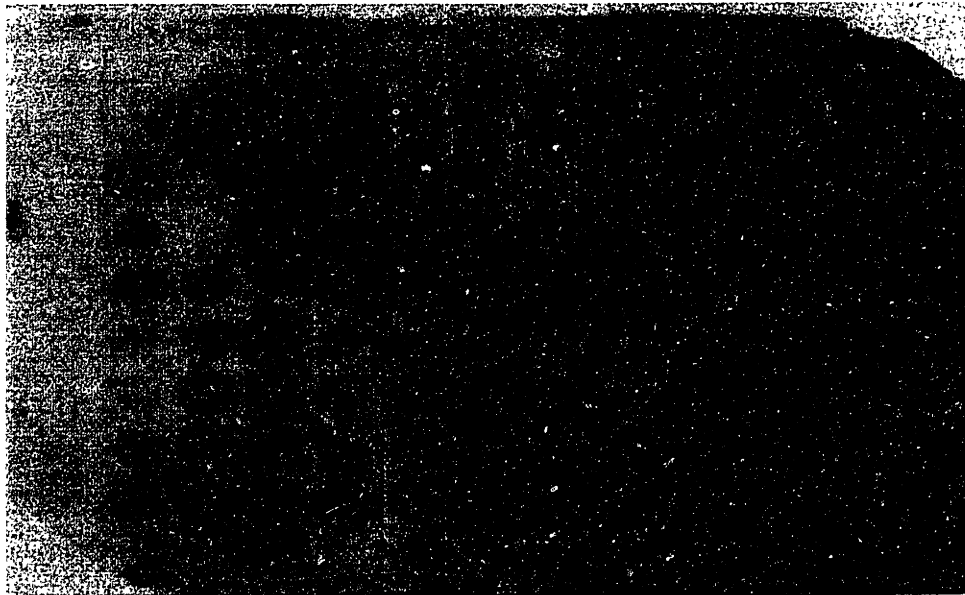
Figure 2.3 Initial screening of the cosmid library under less stringent conditions
(A) Cosmids IA1–IH11
(B) Cosmids IIA1–IIH11

Figure 2.4 Dot blot hybridization results for the 17 cosmids that appeared positive in the first screen. Cosmids IIE9, IIIF10, IIIG7 and IIIB7 were chosen for further analysis. The hybridization probe was included as a positive control.
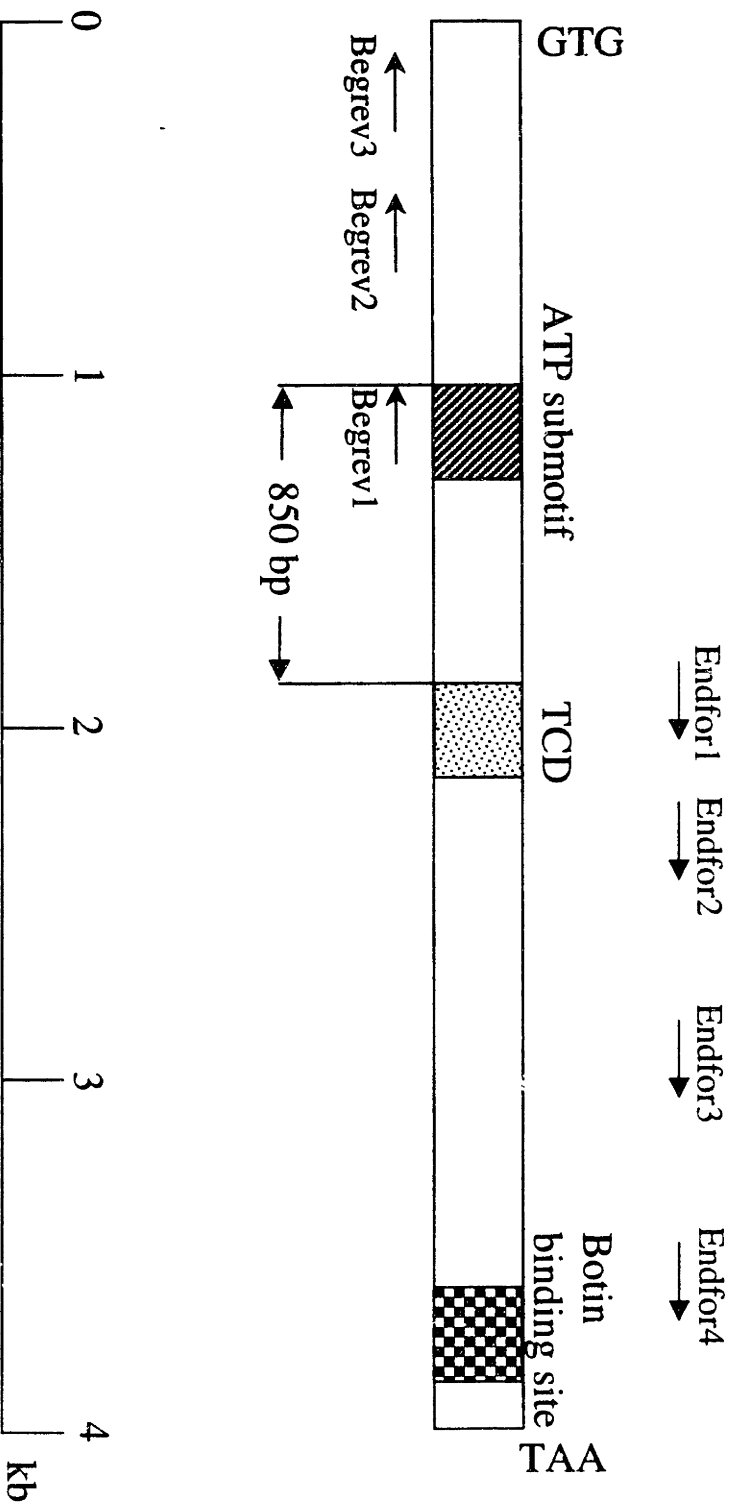
Figure 2.5 The strategy used for sequencing cosmid IIIF10. The 850 bp DNA fragment was first sequenced and primers Endfor1, Endfor2, Endfor3 and Endfor4 were used to obtain the sequence on one side of the 850 bp fragment and primers Begrev1, Begrev2 and Begrev3 were used to obtain the sequence on the other side of the 850 bp fragment.
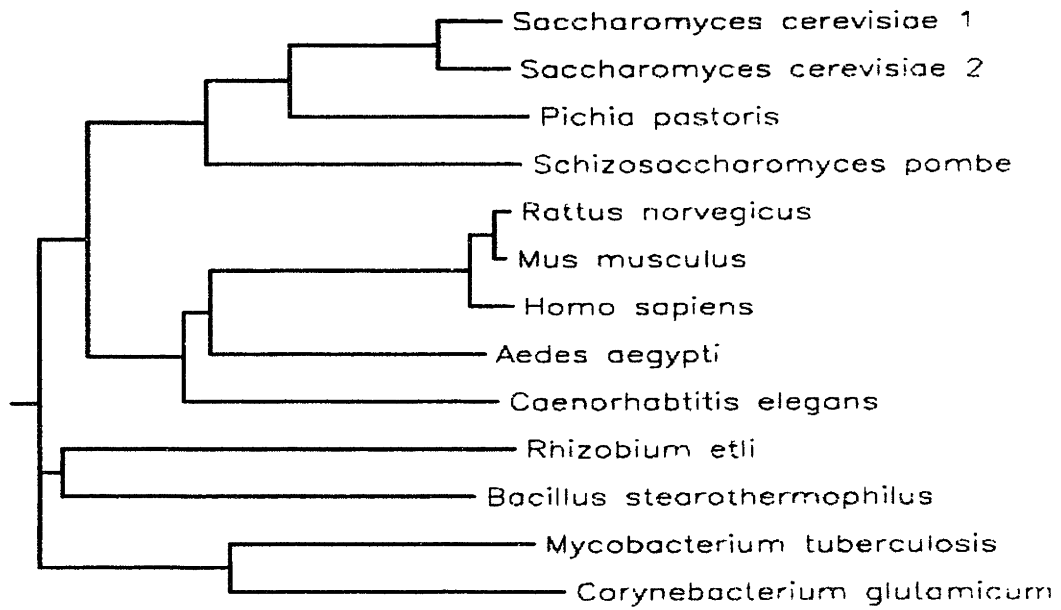
Figure 2.6 Rooted phylogenetic tree of various pyruvate carboxylases.
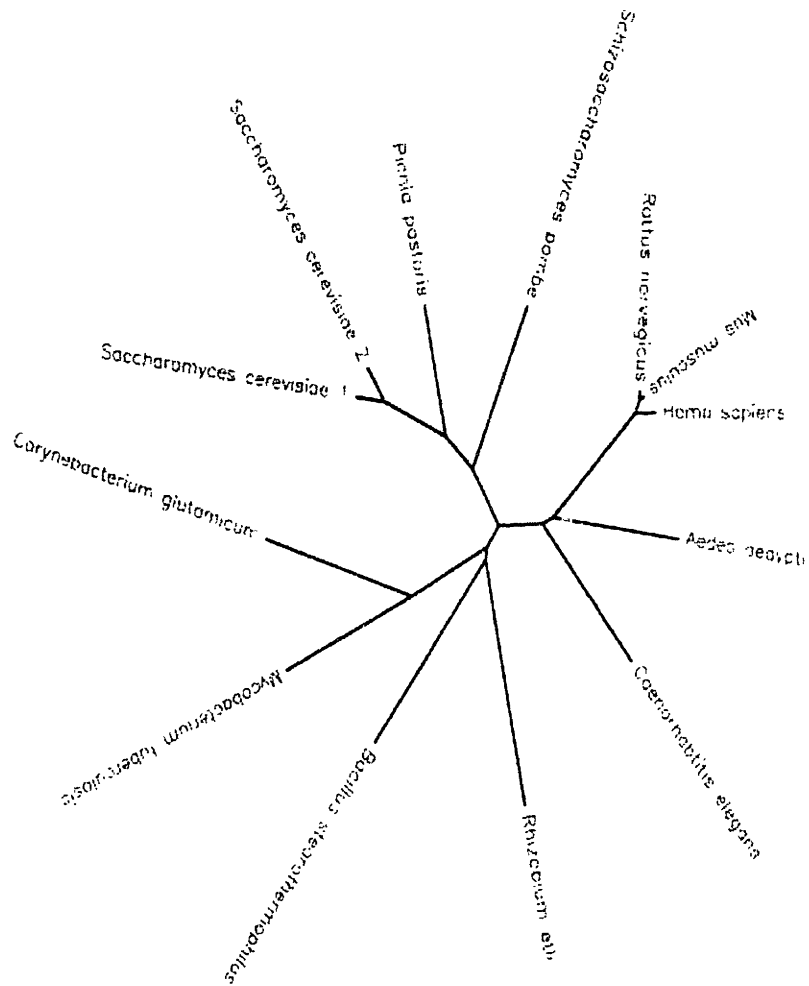
Figure 2.7 Unrooted phylogenetic tree of various pyruvate carboxylases
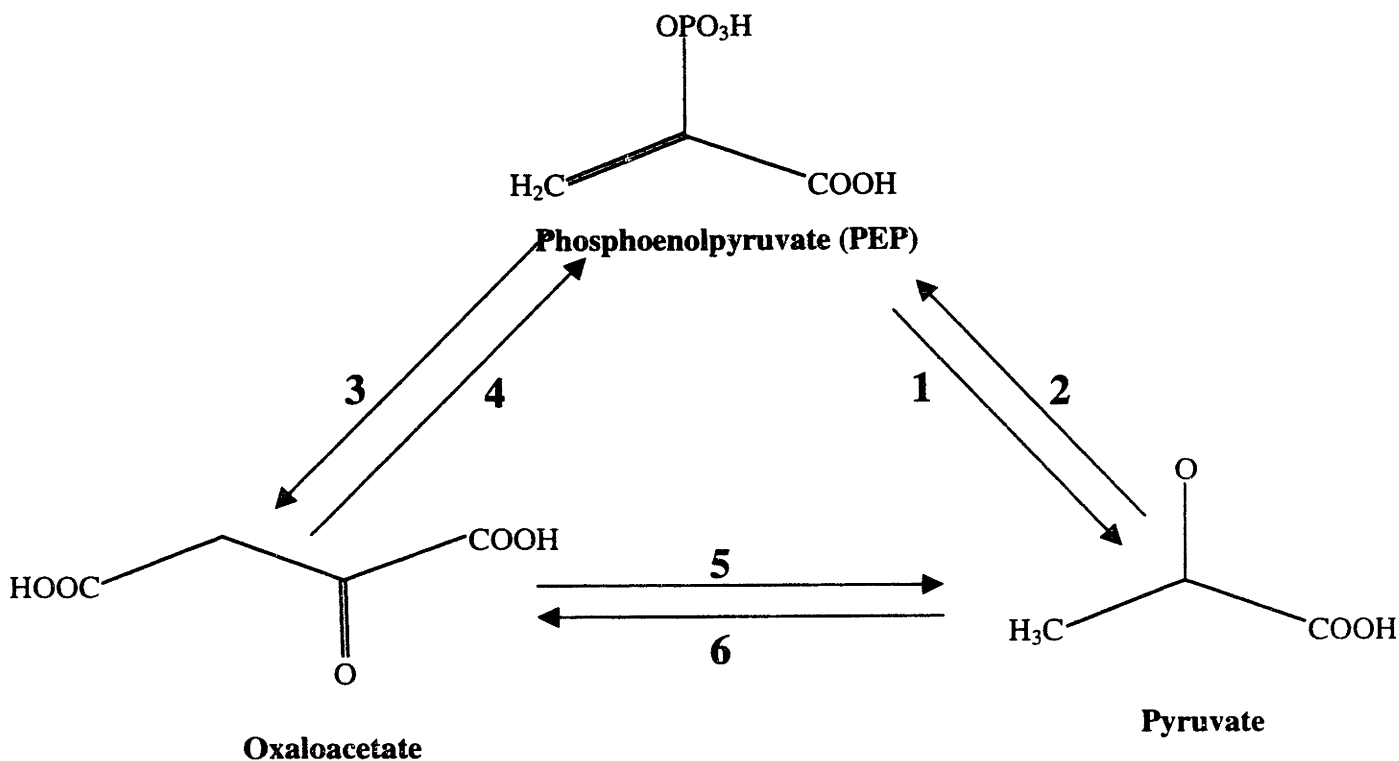
Figure 2.8 The metabolic pathways for the interconversion of oxaloacetate, phosphoenolpyruvate and pyruvate in *C. glutamicum*. Enzymes: *1* pyruvate kinase, *2* phosphoenolpyruvate carboxykinase, *3* phosphoenolpyruvate carboxylase *4* phosphoenolpyruvate synthetase *5* oxaloacetate decarboxylase *6* pyruvate carboxylase

Primer site	Primer sequence
ATP 1	GTC TTC ATC GAG ATG AAT CCG CG
ATP 2	CAG GTG GAG CAC ACG GTG AC
Pyruvate 1	CAT CGT AAG TCG CAC CGC CCC A
Pyruvate 2	CGC AGC GCC ACA TCG TAA GTC GC
Pyruvate (forward)	GCG TCG CGG TGT AGC ATT CAG CG
Biotin	CGC TAC TTC TAC CTT

Table 2.1 Primers used for isolating by PCR fragment of the gene of pyruvate carboxylase using C. glutamicum genomic DNA as template

Primer name	Primer sequence (5'-3')
Begrev1	TTC ACC AGG TCC ACC TCG
Begrev2	GAT GCT TCT GTT GCT AAT TTG C
Begrev3	GCG GTG GAA TGA TCC CCG A
Endfor1	CGT CGC AAA GCT GAC TCC
Endfor2	GGC CAT TAA GGA TAT GGC TG
Endfor3	ACC GCA CTG GGC CTT GCG
Endfor4	TCG CCG CTT CGG CAA CAC

Table 2.2 Primers used for sequencing the gene of pyruvate carboxylase

100× Mineral Salts

$MnSO_4$	200 mg/l
$Na_2B_4O_7 \cdot 4H_2O$	20 mg/l
$FeCl_3 \cdot 6H_2O$	200 mg/l
$ZnSO_4 \cdot 7H_2O$	50 mg/l
$CuCl_2 \cdot 2H_2O$	20 mg/l
$(NH_4)_6Mo_7O_{24} \cdot 4H_2O$	10 mg/l

(pH~2 to avoid precipitation)

Part A	Final conc.	Amt. For 100 ml 6× solution
D- Glucose	20 g/l	12 g
$Na_3Citrate \cdot 2H_2O$	1.1 g/l	0.66 g
NaCl	1 g/l	0.6 g
$MgSO_4 \cdot 7H_2O$	200 mg/l	120 mg
$Na_2EDTA \cdot 2H_2O$	75 mg/l	45 mg
$FeSO_4 \cdot 7H_2O$	25 mg/l	15 mg
$CaCl_2 \cdot 2H_2O$	50 mg/l	30 mg
100× Mineral Salts	10 ml/l	6 ml

Part B		
K_2HPO_4	8 g/l	4.8 g
KH_2PO_4	1 g/l	0.6 g
L-threonine	150 mg/l	0.9 g
L-methionine	80 mg/l	0.48 g
L-leucine	100 mg/l	0.6 g

Part C		
$(NH_4)_2SO_4$	5 g/l	3 g

Vitamins	Final conc.	Per 30 ml medium
D- Biotin (0.4 mg/ml)	1 mg/l	75 µl
Thiamine·HCl (0.5 mg/ml)	1 mg/ml	60 µl

Table 2.3 Glucose minimal medium for *C. glutamicum* growth

tggggcggggttagatcctggggggtttatttcattcactttggcttgaagtcgtgcaggtcaggggagtgttgcccgaaaacattgag

aggaaaacaaaaaccgatgtttgattgggggaatcggggggttacgatactaggacgcagtgactgctatcacccttggcggtctcttgt

pc ⟶
tgaaaggaataattactctaGTGTCGACTCACACATCTTCAACGCTTCCAGCATTCAAAAAGATCTTGGTAGCAAACCGCGGCGAAATC
         M S T H T S S T L P A F K K I L V A N R G E I
         1

GCGGTCCGTGCTTTCCGTGCAGCACTCGAAACCGGTGCAGCCACGGTAGCTATTTACCCCCGTGAAGATCGGGGATCATTCCACCGCTC
A V R A F R A A L E T G A A T V A I Y P R E D R G S F H R S
24

TTTTGCTTCTGAAGCTGTCCGCATTGGTACCGAAGGCTCACCAGTCAAGGCGTACCTGGACATCGATGAAATTATCGGTGCAGCTAAAA
 F A S E A V R I G T E G S P V K A Y L D I D E I I G A A K K
54

AAGTTAAAGCAGATGCCATTTACCCGGGATACGGCTTCCTGTCTGAAAATGCCCAGCTTGCCCGCGAGTGTGCGGAAAACGGCATTACT
 V K A D A I Y P G Y G F L S E N A Q L A R E C A E N G I T
84

TTTATTGGCCCAACCCCAGAGGTTCTTGATCTCACCGGTGATAAGTCTCGCGCGGTAACCGCCGCGAAGAAGGCTGGTCTGCCAGTTTT
 F I G P T P E V L D L T G D K S R A V T A A K K A G L P V L
114

GGCGGAATCCACCCCGAGCAAAAACATCGATGAGATCGTTAAAAGCGCTGAAGGCCAGACTTACCCCATCTTTGTGAAGGCAGTTGCCG
 A E S T P S K N I D E I V K S A E G Q T Y P I F V K A V A G
144

ATP binding submotiff
GTGGTGGCGGACGCGGTATGCGTTTTGTTGCTTCACCTGATGAGCTTCGCAAATTAGCAACAGAAGCATCTCGTGAAGCTGAAGCGGCT
[G G G R G M R] F V A S P D E L R K L A T E A S R E A E A A
174

TTCGGCGATGGCGCGGTATATGTCGAACGTGCTGTGATTAACCCTCAGCATATTGAAGTGCAGATCCTTGGCGATCACACTGGAGAAGT
F G D G A V Y V E R A V I N P Q H I E V Q I L G D H T G E V
204

TGTACACCTTTATGAACGTGACTGCTCACTGCAGCGTCGTCACCAAAAAGTTGTCGAAATTGCGCCAGCACAGCATTTGGATCCAGAAC
 V H L Y E R D C S L Q R R H Q K V V E I A P A Q H L D P E L
234

TGCGTGATCGCATTTGTGCGGATGCAGTAAAGTTCTGCCGCTCCATTGGTTACCAGGGCGCGGGAACCGTGGAATTCTTGGTCGATGAA
 R D R I C A D A V K F C R S I G Y Q G A G T V E F L V D E
264

ATP binding submotiff
AAGGGCAACCACGTCTTCATCGAAATGAACCCACGTATCCAGGTTGAGCACACCGTGACTGAAGAAGTCACCGAGGTGGACCTGGTGAA
K G N H V [F I E M N P R I] Q V E H T V T E E V T E V D L V K
294

GGCGCAGATGCGCTTGGCTGCTGGTGCAACCTTGAAGGAATTGGGTCTGACCCAAGATAAGATCAAGACCCACGGTGCAGCACTGCAGT
 A Q M R L A A G A T L K E L G L T Q D K I K T H G A A L Q C
324

GCCGCATCACCACGGAAGATCCAAACAACGGCTTCCGCCCAGATACCGGAACTATCACCGCGTACCGCTCACCAGGCGGAGCTGGCGTT
 R I T T E D P N N G F R P D T G T I T A Y R S P G G A G V
354

CGTCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATCACCGCACACTTTGACTCCATGCTGGTGAAAATGACCTGCCGTGGTTCCGACTT
R L D G A A Q L G G E I T A H F D S M L V K M T C R G S D F
384

TGAAACTGCTGTTGCTCGTGCACAGCGCGCGTTGGCTGAGTTCACCGTGTCTGGTGTTGCAACCAACATTGGTTTCTTGCGTGCGTTGC
 E T A V A R A Q R A L A E F T V S G V A T N I G F L R A L L
414

```
TGCGGGAAGAGGACTTCACTTCCAAGCGCATCGCCACCGGATTCATTGCCGATCACCCGCACCTCCTTCAGGCTCCACCTGCTGATGAT
  R    E    E    D    F    T    S    K    R    I    A    T    G    F    I    A    D    H    P    H    L    L    Q    A    P    P    A    D    D
444

GAGCAGGGACGCATCCTGGATTACTTGGCAGATGTCACCGTGAACAAGCCTCATGGTGTGCGTCCAAAGGATGTTGCAGCTCCTATCGA
  E    Q    G    R    I    L    D    Y    L    A    D    V    T    V    N    K    P    H    G    V    R    P    K    D    V    A    A    P    I    D
474

TAAGCTGCCTAACATCAAGGATCTGCCACTGCCACGCGGTTCCCGTGACCGCCTGAAGCAGCTTGGCCCAGCCGCGTTTGCTCGTGATC
 K    L    P    N    I    K    D    L    P    L    P    R    G    S    R    D    R    L    K    Q    L    G    P    A    A    F    A    R    D    L
504

TCCGTGAGCAGGACGCACTGGCAGTTACTGATACCACCTTCCGCGATGCACACCAGTCTTTGCTTGCGACCCGAGTCCGCTCATTCGCA
  R    E    Q    D    A    L    A    V    T    D    T    T    F    R    D    A    H    Q    S    L    L    A    T    R    V    R    S    F    A
534

CTGAAGCCTGCGGCAGAGGCCGTCGCAAAGCTGACTCCTGAGCTTTTGTCCGTGGAGGCCTGGGGCGGCGCGACCTACGATGTGGCGAT
 L    K    P    A    A    E    A    V    A    K    L    T    P    E    L    L    S    V    E    A    W    G    G    A    T    Y    D    V    A    M
564

               pyruvate binding motiff
GCCGTTTCCTCTTTGAGGATCCGTGGGACAGGCTCGACGAGCTGCGCGAGGCGATGCCGAATGTAAACATTCAGATGCTGCTTCGCGGCC
  R   |F    L    F    E    D    P    W    D    R    L|   D    E    L    R    E    A    M    P    N    V    N    I    Q    M    L    L    R    G    R
594

GCAACACCGTGGGATACACCCCGTACCCAGACTCCGTCTGCCGCGCGTTTGTTAAGGAAGCTGCCAGCTCCGGCGTGGACATCTTCCGC
  N    T    V    G    Y    T    P    Y    P    D    S    V    C    R    A    F    V    K    E    A    A    S    S    G    V    D    I    F    R
624

ATCTTCGACGCGCTTAACGACGTCTCCCAGATGCGTCCAGCAATCGACGCAGTCCTGGAGACCAACACCGCGGTAGCCGAGGTGGCTAT
 I    F    D    A    L    N    D    V    S    Q    M    R    P    A    I    D    A    V    L    E    T    N    T    A    V    A    E    V    A    M
654

GGCTTATTCTGGTGATCTCTCTGATCCAAATGAAAAGCTCTACACCCTGGATTACTACCTAAAGATGGCAGAGGAGATCGTCAAGTCTG
 A    Y    S    G    D    L    S    D    P    N    E    K    L    Y    T    L    D    Y    Y    L    K    M    A    E    E    I    V    K    S    G
684

GCGCTCACATCTTGGCCATTAAGGATATGGCTGGTCTGCTTCGCCCAGCTGCGGTAACCAAGCTGGTCACCGCACTGCGCCGTGAATTC
  A    H    I    L    A    I    K    D    M    A    G    L    L    R    P    A    A    V    T    K    L    V    T    A    L    R    R    E    F
714

GATCTGCCAGTGCACGTGCACACCCACGACACTGCGGGTGGCCAGCTGGCAACCTACTTTGCTGCAGCTCAAGCTGGTGCAGATGCTGT
 D    L    P    V    H    V    H    T    H    D    T    A    G    G    Q    L    A    T    Y    F    A    A    A    Q    A    G    A    D    A    V
744

TGACGGTGCTTCCGCACCACTGTCTGGCACCACCTCCCAGCCATCCCTGTCTGCCATTGTTGCTGCATTCGCGCACACCCGTCGCGATA
  D    G    A    S    A    P    L    S    G    T    T    S    Q    P    S    L    S    A    I    V    A    A    F    A    H    T    R    R    D    T
774

CCGGTTTGAGCCTCGAGGCTGTTTCTGACCTCGAGCCGTACTGGGAAGCAGTGCGCGGACTGTACCTGCCATTTGAGTCTGGAACCCCA
  G    L    S    L    E    A    V    S    D    L    E    P    Y    W    E    A    V    R    G    L    Y    L    P    F    E    S    G    T    P
804

GGCCCAACCGGTCGCGTCTACCGCCACGAAATCCCAGGCGGACAGTTGTCCAACCTGCGTGCACAGGCCACCGCACTGGGCCTTGCGGA
 G    P    T    G    R    V    Y    R    H    E    I    P    G    G    Q    L    S    N    L    R    A    Q    A    T    A    L    G    L    A    D
834

TCGTTTCGAACTCATCGAAGACAACTACGCAGCCGTTAATGAGATGCTGGGACGCCCAACCAAGGTCACCCCATCCTCCAAGGTTGTTG
  R    F    E    L    I    E    D    N    Y    A    A    V    N    E    M    L    G    R    P    T    K    V    T    P    S    S    K    V    V    G
864

GCGACCTCGCACTCCACCTCGTTGGTGCGGGTGTGGATCCAGCAGACTTTGCTGCCGATCCACAAAAGTACGACATCCCAGACTCTGTC
  D    L    A    L    H    L    V    G    A    G    V    D    P    A    D    F    A    A    D    P    Q    K    Y    D    I    P    D    S    V
894

ATCGCGTTCCTGCGCGGCGAGCTTGGTAACCCTCCAGGTGGCTGGCCAGAGCCACTGCGCACCCGCGCACTGGAAGGCCGCTCCGAAGG
 I    A    F    L    R    G    E    L    G    N    P    P    G    G    W    P    E    P    L    R    T    R    A    L    E    G    R    S    E    G
924
```

CAAGGCACCTCTGACGGAAGTTCCTGAGGAAGAGCAGGCGCACCTCGACGCTGATGATTCCAAGGAACGTCGCAATAGCCTCAACCGCC
K A P L T E V P E E E Q A H L D A D D S K E R R N S L N R L
954

TGCTGTTCCCGAAGCCAACCGAAGAGTTCCTCGAGCACCGTCGCCGCTTCGGCAACACCTCTGCGCTGGATGATCGTGAATTCTTCTAC
L F P K P T E E F L E H R R R F G N T S A L D D R E F F Y
984

GGCCTGGTCGAAGGCCGCGAGACTTTGATCCGCCTGCCAGATGTGCGCACCCCACTGCTTGTTCGCCTGGATGCGATCTCTGAGCCAGA
G L V E G R E T L I R L P D V R T P L L V R L D A I S E P D
1014

CGATAAGGGTATGCGCAATGTTGTGGCCAACGTCAACGGCCAGATCCGCCCAATGCGTGTGCGTGACCGCTCCGTTGAGTCTGTCACCG
D K G M R N V V A N V N G Q I R P M R V R D R S V E S V T A
1044

CAACCGCAGAAAAGGCAGATTCCTCCAACAAGGGCCATGTTGCTGCACCATTCGCTGGTGTTGTCACCGTGACTGTTGCTGAAGGTGAT
T A E K A D S S N K G H V A A P F A G V V T V T V A E G D
1074

biotin binding motif

GAGGTCAAGGCTGGAGATGCAGTCGCAATCATCGAGGCTATGAAGATGGAAGCAACAATCACTGCTTCTGTTGACGGCAAAATCGATCG
E V K A G D A V A I I [E A M K M E] A T I T A S V D G K I D R
1104

CGTTGTGGTTCCTGCTGCAACGAAGGTGGAAGGTGGCGACTTGATCGTCGTCGTTTCCTAAacctttctgtaaaaag
V V V P A A T K V E G G D L I V V V S *
1134

S. pombe S. cerevisiae 1 S. cerevisiae 2 R. etli P. pastoris M. musculus H. sapiens C. elegans B. stearothermophilus M. tuberculosis C. glutamicum — multiple sequence alignment

Block 1

	Sequence		
S. pombe	--MTSKYDALLHNQSTNTPFSKLQDRSSLLGEKFT--	VL	AI
S. cerevisiae 1	----------------MSQRKFAGLRDNFNLLGEKN---	IL	PI
S. cerevisiae 2	--------------MSSSKKLAGLRDNFSLLGEKN---	IL	PI
R. etli	MLKFQTVRGGLRVRRSSTAPVASPNVRRLEYKPIK	VM	AI
P. pastoris	----------MAEEDYLPVYQLRRDSSLLGTMN---	IL	PI
M. musculus	MLKFQTVRGGLRLLGVRRSSSAPVASPNVRRLEYKPIK	VM	AI
H. sapiens	MLKFRTVHGGLRLLGIRRTSTAPAASPNVRRLEYKPIK	VM	AI
C. elegans	-MRFSRIPP-IFANVVRQTHYRNYANGVIKPREFN---	VM	AI
B. stearothermophilus	----------------------MKTRRIR--	VL	AI
M. tuberculosis	--------------------------------MFS--	VL	AI
C. glutamicum	------------------------MSTHTSSTLPAFK--	IL	AV
Consensus		K VAHRGEI	

Block 2

	Sequence
S pombe	V TAH LSMHTV IYSYE RLSM RQK D SYPI KVGQYSPVG A
S. cerevisiae 1	I TAH LSMQTV IYSHE RLST KQK D AYVI EVGQYTPVG A
S cerevisiae 2	I SAH LSMRTI IYSHE RLSM RLK D AYVI EEGQYTPVG A
R. etli	V ACT LGIRTV VYSEQ TGQM RQK D AYLI RG--LAPVQ H
P pastoris	I TAH LSMNTV IYSHE RLSM RLK D AYVI ERGQYSPVQ A
M. musculus	V ACT LGIRTV VYSEQ TGQM RQK D AYLI RG--LAPVQ H
H sapiens	V ACT LGIRTV IYSEQ TGQM RQK D AYLI RG--LAPVQ H
C. elegans	V ALT LNKTSV IYAEQ KNSM RLK D AYLV KG--LPPVA T
B stearothermophilus	V ACT LGIRTV IYSKE VGSY RYK D AYLV EG--KKPIE D
M tuberculosis	A AAY LGVGTV VYPYE RNSQ RLK C SYQI DI--GHPVH S
C glutamicum	A AAL TGAATV IYPRE RGSF RSF S AVRI TE--GSPVK D
Consensus	R FR T A D R S G AYL

Block 3

	Sequence
S pombe	IDE VSI KRTGANLVH L NAEF RKVNEA MQFV SPEVIDS
S cerevisiae 1	IDE IJI QKHQVDFIH L NSEF DKVVKA ITWI PAEVIDS
S cerevisiae 2	MDE IEI KKHKVDFIH L NSEF DKVVKA ITWI PAEVIDS
R etli	IPD IKV KENGVDAVH L RADF QACQDA VRFI SPEVVRK
P pastoris	IDE IKI VKHNVNMIH C NSEF RKVEEN ILWV SDTVIDA
M musculus	IPD IKV KENGVDAVH L RADF QACQDA VRFI SPEVVRK
H sapiens	IPD IFV KENNVDAVH L RADF QACQDA VRFI SPEVVRK
C elegans	IDQ IET LKHNIDAIH L RSDF AACQNA IVFI SPDVMAR
B stearothermophilus	IEG IEI KAHDVDAIH F NIQF KRCREE IIFF NENHLDM
M tuberculosis	VDE VAT RRAGADAIY L NPDL AACAAA ISFV SAEVLEL
C glutamicum	IDE IGA KKVKADAIY L NAQL RECAEN ITFI TPEVLDL
Consensus	I A PGYGF SE A G GP

Block 4

	Sequence
S pombe	L D TK RAI IRCGVPVVPGTPG VEHYEEAEAFVKEYGL VII AM
S cerevisiae 1	V D VS RNL AKANVPTVPGTPG IETVEEALDFVNEYGY VII AF
S cerevisiae 2	V D VS RHL ARANVPTVPGTPG IETVQEALDFVNEYGY VII AF
R etli	M D VE RAI IAAGVPVVPGTNS INSLHEAHEFSNTYGF IIF AY
P pastoris	V D VS RNL YAANVPTVPGTPG IEDVAQATAFVEEYGY VII AF
M musculus	M D VE RAI IAAGVPVVPGTDS ISSLHEAHEFSNTFGF IIF AY
H sapiens	M D VE RAI IAAGVPVVPGTDA ITSLHEAHEFSNTYGF IIF AY
C elegans	M D VA RQA IEAGVQVVPGTPG ITTADEAVEFAKQYGF IIL AY
B stearothermophilus	F D VK RHA VNAGIPVIPGSDG VDGLECVVAFAEAHGY III AL
M tuberculosis	A N SR IAA REAGLPVLMSS-A SASVDELLSVAAGMPF LFV VA
C glutamicum	T D SR VTA KKAGLPVLAES-T SKNIDEIVKSAEGQTY IFV VA
Consensus	G K A A P KA G

Block 5

	Sequence
S pombe	MRV RSADTLKESFERARS LAS DGTVFI RFLDK K I
S cerevisiae 1	MRV REGDDVADAFQRATS RTA NGTCFV RFLDK K V
S cerevisiae 2	MRV REGDDVADAFQRATS RTA NGTCFV RFLDK K V
R etli	MRV HSYEELEENYTRAYP LAA NGALFV KFIEK R V
P pastoris	MRV REGDDIEDAFLRASS KTA NGTVFI RFLDK K V
M musculus	MRV HSYEELEENYTRAYS LAA NGALFV KFIEK R V
H sapiens	MRV HSYEELEENYTRAYS WPA NGALFV KFIEK R V
C elegans	IRR DKLEEVEEAFRRSYS QAA DGSLFV KFVER R V
B stearothermophilus	MRF RSKSEVKEAFERAKS KAA SDEVYV KLIEN K V
M tuberculosis	MRR GDIAALPEAIEAASR ESA DPTVYL QAVIN R V
C glutamicum	MRF ASPDELRKLATEASR EAA DGAVYV RAVIN R Q
Consensus	GGGRG V EX FG T F HIR Q

Block 6

	Sequence
S pombe	LMA KAGNVI H V VEIAPAKD DPKI QALYDDAIKI
S cerevisiae 1	LLA NHGNVV F V VEVAPAKT PREV DAILTDAVKL
S cerevisiae 2	LLA NHGNVV F V VEVAPAKT PREV DAILTDAVKL
R etli	ILG QYGNIL Y I VEIAPATH DPQL SRLTSDSVKL
P pastoris	LLA NYGNVI P V ARNCSAKT PVEV NAILNDAVKL
M musculus	ILG QYGNIL Y V VEIAPATH DPQL SRLTSDSVKL
H sapiens	ILG QYGNIL Y I VEIAPAAH DPQL TRLTSDSVKL
C elegans	LLG HHGNIV Y V VEIAPAPA PEGV EKILADALRL
B stearothermophilus	ILG YEGNIV Y V VEVAPSVS SDEL QRICEAAVQL
M tuberculosis	ILA NLGDVI Y V IELAPAPH DAEL YKMCVDAVAF
C glutamicum	ILG HTGEVV L VEIAPAQH DPEL DRICADAVKF
Consensus	D HL EKDCE QRRHQKV L R

S pombe
S cerevisiae 1
S cerevisiae 2
R etli
P pastoris
M musculus
H sapiens
C elegans
B stearothermophilus
M tuberculosis
C glutamicum
Consensus

```
S pombe                  AKEVKYCN   A     LDQKGRHY    INP I       I  EI  GV I  SA
S cerevisiae 1           AKECGYRN   A     VDNQNRHY    INP I       I  EI  GI I  AA
S cerevisiae 2           AKVCGYRN   A     VDNQNRHY    INP I       I  EI  GI I  SA
R etli                   AKQVGYEN   V     VDKHGKHY    VNS L       V  EI  DV L  HA
P pastoris               AKTANYRN   A     VDSQNRHY    INP I       I  EI  GV I  AA
M musculus               AKQVGYEN   V     VDKHGKHY    VNS L       V  EI  DV L  HA
H sapiens                AKQVGYEN   V     VDPHGKHY    VNS L       V  EI  DV L  HA
C elegans                ARHVGYQN   V     VDQKGNYY    VNA L       V  EI  GV L  QA
B stearothermophilus     MRSVGYVN   V     VSGDE-FY    VNP I       I  MI  GI I  QS
M tuberculosis           ARHIGYSC   V     LDEPGEYV    MNP I       V  EI  DV L  AS
C glutamicum             CRSIGYQG   V     VDEYGNHV    MNP I       V  EV  EV L  KA
Consensus                   AGT  EFL           YIE    R  QVEHT  TL  T    D V    Q
```

```
S pombe                  LHVAA FT PEIGL---T DK STR FAI C V        NNGFA  I KIE
S cerevisiae 1           IQIAA AS PQLGL---F DK TTR FAI C I        AKNFQ  T RIE
S cerevisiae 2           IQIAA AT TQLGL---L DK TTR FSI C I        SKNFQ  T RLE
R etli                   IHVSE RS PDLGL---R EN RIN CAI C V        ARSFQ  T RIE
P pastoris               IQIAA AS EQLGL---L EK TTR FAI C I        TKNFQ  T KIE
M musculus               IHVSE RS PDLGL---R EN RIN CAI C V        APSFQ  T RIE
H sapiens                IHVAE RS PDLGL---R EN RIN CAI C V        APTFQ  T RIE
C elegans                IRIAE KS DDLKL---S ET QTT SAI C V        AKGFQ  S RIE
B stearothermophilus     ILIAD CS HSHEVGIPK ED RIN YAI S V        LNNFM  T KIM
M tuberculosis           LRIAA ET EQLGLR--- ED APH AAL C I        ANGFR  T RIS
C glutamicum             MPLAA AT KTLGLT--- DK KTH AAL C I        NNGFR  T TIT
Consensus                   G L      Q  I  G     Q R TTEDP    PD G
```

```
S pombe                  VY SAG N V   GANGFAGSV TPHY ML CTCHDATYEYTRPKMI
S cerevisiae 1           VY SAG N V   GGNAYAGTI SPHY ML CSCSGSTYEIVRPKMI
S cerevisiae 2           VY SAG N V   GGNAYAGAT SPHY ML CSCSGSTYEIVPRKMI
R etli                   VF SGE M I   NASAFQGAV SPHY LL VIAHGKDHPTAATKMS
P pastoris               VY SSG N V   GGNGFAGAV SPHY ML CSTSGSNYEIPRPKMI
M musculus               VF SGE M I   NASAFQGAV SPHY LL VIAHGKDHPTAATKMS
H sapiens                VF SGE M I   NASAFQGAV SPHY LL VIAHGKDHPTAATKMS
C elegans                VF SGE M I   SASAFAGSV SPHY LM VIASARNHPNAAAKMI
B stearothermophilus     AY SGG F V   AGNGFQGAV TPYY LL LSTWALTFEQAARKML
M tuberculosis           AL TAG A V   GSTNLG-AE SPYF ML LTCRGRDLPTAVSPAP
C glutamicum             AY SPG A V   GAAQLG-GE TAHF ML MTCPGSDFETAVAPAC
Consensus                   R    G GRLD     I    DS  VK
```

```
S pombe                  SLIE RVRGVK   P VLRLLMHDT IQGNCWTT  DDT ELFQLYPSPN
S cerevisiae 1           ALIE RIFGVK   P LLTLLLTNPV IEGTYWTT  DDT QLFQMVSSQN
S cerevisiae 2           ALIE RIFGVK   P LLTLLTNPV IEGTYWTT  DDT QLFQMVSSQN
R etli                   ALAE RVPGVK   P LQNVLNNQQ LAGIVDTQ  DEN ELFQLRPAQN
P pastoris               ALVE RIRGVK   P LLALLTHPV MTSECWTT  DDT ELFKILTSQN
M musculus               ALAE RVPGVK   P LQNVLNNQQ LAGTVDTQ  DEN ELFQLPAQN
H sapiens                ALAE RVPGVK   A LQNVLNNQQ LAGTVDTQ  DEN DVFQLPPAQN
C elegans                ALKK RIPGVK   P LLNVLRQPS LDASVDTY  DEH ELFQFYPSQN
B stearothermophilus     NLRT FIPVFK   P LENVVQHPK LSGEYDTS  DTT ELFVFPPPFD
M tuberculosis           AIAE RIPGVS   P LQAVLDDPD PAGRVTTS  DER CLLTAPASAD
C glutamicum             ALAE TVSGVA   G LRALLREED TSKRIATG  ADH HLLCAPPACD
Consensus                   F      TNI     F              FI   P
```

```
S pombe                  RAQKL AYLGDLA   GSSIKGQNGEPALKSEIVMPVLLDSTGNCIDVSHP
S cerevisiae 1           RAQKL HYLADVA   GSSIKGQIGLPKLKSNFSVPHLHDACGNVINVTKS
S cerevisiae 2           RAQKL HYLADLA   GSSIKGQIGLPKLKSNPSVPHLHDACGNVINVTFS
R etli                   RAQKL HYLGHVM   GPTTPIPYKVSPSPDPIVPVVPIGP---  ---
P pastoris               RAQKL AYLGDLA   GSSIKGQIGLPKLHYEADIPSITDINGDVIDVSIP
M musculus               RAQKL HYLGHVM   GPTTPIPVNVSPSPVDPAVPVVPIGP---------
H sapiens                RAQKL HYLGHVM   GPTTPIPVKASPSPTDPVVPAVPIGP --------
C elegans                RAQKL NYLGEVK   GPTTPLATDLKPAVVSPPIPYIPAGAK---------
B stearothermophilus     RGTKM TYIGTVT   GFPGIGFKK-KPVFDKPRLPKLSEAEP---------
M tuberculosis           PGTKI NFLADVT   NPYGSPPSTIYPDDKLPDLDLPAAPF---------
C glutamicum             EQGPI DYLADVT   FPHGVPPKDVAAPIDFLPNIFDLPLP---------
Consensus                   L             VN
```

```
S pombe                  SE-K WRKL LDNGPAAFAKAVPNHKPGLIM  W              TI
S cerevisiae 1           APPS WPQV LEKGPAEFAPQVRQFNGTLLM  W              TH
S cerevisiae 2           APPS WRQV LEKGPSEFAKQVRQFNGTLLM  W              TH
R etli                   -PPA FRDI LREGPEGFARAVRNHQGLLLM  F              TH
P pastoris               PP-D WRQF LEKGPEQFAQQVPAFPGLLM   W              TH
M musculus               -PPA FRDI LPEGPEGFAPAVPNHQGLLLM  F              TH
H sapiens                -PPA FRDI LPEGPEGFARAVRNHPGLLLM  F              TH
C elegans                -PPT IPDV VQPGPTEFAYEVRSPPGCMIT  F              TY
B stearothermophilus     -IPA TKQI DKHGPEGLVPWIQEQPPVLLT  F              TV
M tuberculosis           ---A SKQP VFLGFEGFAPWLRESAAVGVT  F              TS
C glutamicum             --R SRDR KQLGPAAFAPDLPEQDALAVT   F              SF
Consensus                   Q    L                     DTT RDAHQSLLATRVR
```

71

	Sequence
S pombe	D L V N I A P Y T S H A L A S A Y S L ■ M ■ ■ F ■ S M ■ H ■ C ■ D ■ R R ■ K L V
S cerevisiae 1	D L A T I A P T T A H A L A G R F A L ■ C ■ ■ F ■ A M ■ H ■ D ■ E ■ R K ■ S L V
S cerevisiae 2	D L A T I A P T T A H A L A G A F A L ■ C ■ ■ F ■ A M ■ H ■ D ■ E ■ R K ■ S L V
R etli	D L K K I A P Y V A H N F N N L F S I ■ N ■ ■ F ■ A M ■ Y ■ C ■ R ■ Q E ■ E L I
P pastoris	D L L N I A P A T S Y A L H H A F A L ■ C ■ ■ F ■ S M ■ H ■ D ■ Q ■ R K ■ K A V
M musculus	D L K K I A P Y V A H N F N K L F S M ■ N ■ ■ F ■ A M ■ Y ■ C ■ R ■ Q E ■ E L I
H sapiens	D L K K I A P Y V A H N F S K L F S M ■ N ■ ■ F ■ A M ■ Y ■ C ■ R ■ Q E ■ E L I
C elegans	D M A A I S P F V A Q S F N G L F S L ■ N ■ ■ F ■ S M ■ H ■ C ■ E ■ Q T ■ K L I
B stearothermophilus	D L V R A A E P S A R L L P N L F S L ■ N ■ ■ F ■ A Y ■ K ■ D ■ D ■ L K ■ D A F
M tuberculosis	G L S R V A P Y L A R T M P Q L L S V ■ C ■ ■ Y ■ A L ■ K ■ D ■ E ■ A T ■ A A M
C glutamicum	A L K P A A E A V A K L T P E L L S V ■ A ■ ■ Y ■ A M ■ F ■ D ■ D ■ D E ■ E A M
Consensus	E WGGAT DV RFL E PW RL LR P

	Sequence
S pombe	■ I P F Q M ■ ■ ■ G A ■ G L C ■ S S L ■ D N V I Y P ■ C E Q A K K N ■ I D I ■ V ■ A ■ D V N N
S cerevisiae 1	■ I P F Q M ■ ■ ■ G A ■ G V A ■ S S L ■ D N A I D H ■ V K Q A K D N ■ V D I ■ V ■ A ■ D L E Q
S cerevisiae 2	■ I P F Q M ■ ■ ■ G A ■ G V A ■ S S L ■ D N A I D H ■ V K Q A K D N ■ V D I ■ V ■ A ■ D L E Q
R etli	■ I P F Q M ■ ■ ■ G A ■ A V G ■ T N Y ■ D N V V F K ■ C E V A K E N ■ M D V ■ V ■ S ■ Y L P N
P pastoris	■ I P F S M ■ ■ ■ G G ■ G V A ■ Y S L ■ D N A I D H ■ L K Q A K D T ■ V D V ■ V ■ A ■ D I E Q
M musculus	■ I P F Q M ■ ■ ■ G A ■ A V G ■ T N Y ■ D N V V F K ■ C E V A K E N ■ M D V ■ V ■ S ■ Y L P N
H sapiens	■ I R F Q M ■ ■ ■ G A ■ A V G ■ T N Y ■ D N V V F K ■ C E V A K E N ■ M D V ■ V ■ S ■ Y L P N
C elegans	■ I P F Q C ■ ■ ■ G A ■ A M G ■ S N Y ■ D N V I Y K ■ C E L A V K N ■ M D V ■ V ■ S ■ Y L P N
B stearothermophilus	■ V L F Q M ■ ■ ■ S A ■ A V G ■ K N Y ■ D N V I R E ■ V E K S A H A ■ I H V ■ I ■ S ■ W V K G
M tuberculosis	■ I C L Q M ■ ■ ■ G R ■ T V G ■ T P Y ■ E I V T S A ■ V Q E A T A T ■ I D I ■ I ■ A ■ N I E S
C glutamicum	■ V N I Q M ■ ■ ■ G R ■ T V G ■ T P Y ■ D S V C R A ■ V K E A A S S ■ V D I ■ I ■ A ■ D V S Q
Consensus	N LLR N N P P G FK FD LN

	Sequence
S pombe	L S L G I D ■ A K R A - G G V V ■ A T M C Y S ■ M L N P K K - K ■ N ■ D ■ V N L V D K M V E M G
S cerevisiae 1	L K V G V D ■ V K K A - G G V V ■ A T V C F S ■ M L Q P G K - K ■ N ■ D ■ L E I A E K I V Q M G
S cerevisiae 2	L K V G V N ■ V K K A - G G V V ■ A T V C Y S ■ M L Q P G K - K ■ N ■ D ■ L E V V E K I V Q M G
R etli	M L L G M E ■ A G S A - G G V V ■ A A I S Y T ■ V A D P S R T K ■ S ■ E ■ M G L A E E L V R A G
P pastoris	L K V G V D ■ V K K A - G G V V ■ A T M C Y S ■ M L K P G K - K ■ N ■ E ■ I N L A T E I V E M G
M musculus	M L L G M E ■ A G S A - G G V V ■ A A I S Y T ■ V A D P S R T K ■ S ■ E ■ M G L A E E L V R A G
H sapiens	M L L G M E ■ A G S A - G G V V ■ A A I S Y T ■ V A D P S R T K ■ S ■ Q ■ M G L A E A L V R A G
C elegans	L L V G M E ■ V G K A - G G V V ■ A A I A Y T ■ V T D K S P D K ■ D ■ K ■ L N L A D Q L V K A Q
B stearothermophilus	M T V A I D ■ V R Q S - G K I A ■ A A I C Y T ■ I L D P S P S K ■ N ■ D ■ K A L A K E L E Q A G
M tuberculosis	M R P A I D ■ V R E T G S A I A ■ V A M C Y T ■ L T D P G E Q L ■ T ■ D ■ L K L A E Q I V D A G
C glutamicum	M R P A I D ■ V L E T N T A V A ■ V A M A Y S ■ L S D P N E K L ■ T ■ D ■ L K M A E E I V F S G
Consensus	A E GD Y L YI

	Sequence
S pombe	I H I ■ G I ■ ■ ■ V M K ■ K ■ A P L ■ I S A I R E K H P E L ■ I H V H T H D S A ■ T A V A S M A
S cerevisiae 1	T H I ■ G I ■ ■ ■ T M K ■ A ■ A K L ■ I G S L R A K Y P D L ■ I H V H T H D S A ■ T A V A S M T
S cerevisiae 2	T H I ■ G I ■ ■ ■ T M K ■ A ■ A K L ■ I G S L R T R Y P D L ■ I H V H S H D S A ■ T A V A S M T
R etli	T H I ■ C I ■ ■ ■ L L K ■ A ■ C T M ■ V S S L R D R F P D L ■ L H I H T H D T S ■ S G V A A M L
P pastoris	T H I ■ A V ■ ■ ■ T L K ■ T ■ A K Q ■ I S A L R R K F P S L ■ I H V H T H D S A ■ T G V A S M V
M musculus	T H I ■ C I ■ ■ ■ L L K ■ A ■ C T M ■ V S S L R D R F P D L ■ L H I H T H D T S ■ A G V A A M L
H sapiens	T H I ■ C I ■ ■ ■ L L K ■ T ■ C T M ■ V S S L R D R F P D L ■ L H I H T H A P S ■ A G V A A M L
C elegans	A H I ■ S I ■ ■ ■ V L K ■ E ■ A K L ■ I G A L R D K F P D I ■ I H V H T H D T S ■ A G V A A M L
B stearothermophilus	A H I ■ A I ■ ■ ■ L L K ■ Q ■ A H V ■ I S G L K E T - V D I ■ I Y L H T H D T S ■ N G I Y T A A
M tuberculosis	A H V ■ A I ■ ■ ■ L L R ■ P ■ A Q R ■ V S A L R S P - F D L ■ V H L H T H D T P ■ G Q L A S Y V
C glutamicum	A H I ■ A I ■ ■ ■ L L R ■ A ■ V T K ■ V T A L R R E - F D L ■ V H V H T H D T A ■ G Q L A T Y F
Consensus	L KDRAG P A L F Q

	Sequence
S pombe	A A L E ■ ■ A ■ V ■ D V A T D S M S ■ L ■ S F G A V L A S V D G T D K Q L E F D N N Q L P E I
S cerevisiae 1	A C A L ■ ■ A ■ V ■ D V A I N S M S ■ L ■ S I N A L L A S L E G - N I D T G I N V E H V P E L
S cerevisiae 2	A C A L ■ ■ A ■ V ■ D V A I N S M S ■ L ■ S I N A L L A S L E G - N I D T G I N V E H V P E L
R etli	A C A Q ■ ■ A ■ V ■ D V A V D S M S ■ M ■ S M G A L V A C T K G T P L D T E V P L E R V F D Y
P pastoris	A C A R ■ ■ A ■ V ■ T V R V N S M S ■ M ■ S M S A F I A S L D G - E I E T G I P E A N A R E I
M musculus	A C A Q ■ ■ A ■ V ■ D V A V D S M S ■ M ■ S M G A L V A C T K G T P L D T E V P L E R V F D Y
H sapiens	A C A Q ■ ■ A ■ V ■ D V A A D S M S ■ M ■ S M G A L V A C T R G T P L D T E V P M E R V F D Y
C elegans	E C A K ■ ■ A ■ V ■ D A A V D S M S ■ M ■ S M G A I V A S L Q G T K H D T G L S L D D I S F Y
B stearothermophilus	K A I E ■ ■ V ■ I ■ D V A I S S M A ■ L ■ S A N T L Y Y A L E G T E R A P E V D I Y G L E Q L
M tuberculosis	A A W H ■ ■ A ■ A ■ D G A A A P L A ■ T ■ A L S S I V A A A A H T E Y D T G L S L S A V C A L
C glutamicum	A A A Q ■ ■ A ■ A ■ D G A S A P L S ■ Q ■ S L S A I V A A F A H T P R D T G L S L E A V S D L
Consensus	AG D V G TSQP

	Sequence
S pombe	D S ■ A Q M ■ L L ■ S P ■ E S - - E I K G T D S D ■ N H ■ I ■ L T ■ K F ■ T S L ■ G
S cerevisiae 1	D A ■ A E M ■ L L ■ S C ■ E A - - D L K G P D P E ■ Q H ■ I ■ L T ■ L F ■ Q Q L ■ G
S cerevisiae 2	D A ■ A E M ■ L L ■ S C ■ E A - - D L K G P D P E ■ Q H ■ I ■ L T ■ L F ■ Q Q L ■ G
R etli	S E ■ E G A ■ G L ■ A A ■ D C T A T M K S G N S D ■ E N ■ D ■ Y T ■ H F ■ H S M ■ G
P pastoris	D A ■ A E M ■ L L ■ S C ■ E A - - D L K G P D P E ■ Q H ■ I ■ L T ■ L F ■ Q Q V ■ G
M musculus	S E ■ E G A ■ G L ■ A A ■ D C T A T M K S G N S D ■ E N ■ I ■ Y T ■ H F ■ H S M ■ G
H sapiens	S E ■ E G A ■ G L ■ A A ■ D C T A T M K S G N S D ■ E N ■ I ■ Y T ■ H F ■ H S M ■ G
C elegans	S A ■ E S T ■ Q L ■ A P ■ E C A T T M K S G N A D ■ K H ■ I ■ Y T ■ Q F ■ F S L ■ G
B stearothermophilus	A R ■ E D V ■ K F ■ Q E ■ E S - - G M N A P H T E ■ M H ■ M ■ Y S ■ Q Q ■ K A V ■ G
M tuberculosis	E P ■ E A L ■ K V ■ A P ■ E S - - G L P G P T G R ■ H H ■ I ■ L S ■ P Q ■ I A L ■ G
C glutamicum	E P ■ E A V ■ G L ■ L P ■ E S - - G T P G P T G R ■ R H ■ I ■ L S ■ R A ■ T A L ■ A
Consensus	YW R Y F VY E PGGQ NL QA GL

S pombe	TQWAETKKA	IEANK.L	DII	T V	L	QFM	QNKLSAEDVENP		
S cerevisiae 1	EQWAETKRA	REANYLL	DIV	T V	L	QFM	SNKLTSDDVRRL		
S cerevisiae 2	EQWAETKRA	REANYLL	DIV	T V	L	QFM	SNKLTSDDIRPL		
R etli	SKFKEVKKA	VEANQML	DLI	S I	L	QFM	QNGLSRAEAEAQ		
P pastoris	EKWVETKKA	EAANRLL	DIV	T V	L	QFM	SNKLSSEDVEPL		
M musculus	SKFKEVKKA	VEANQML	DLI	S I	L	QFM	QNGLSRAEAEAQ		
H sapiens	SKFKEVKKA	VEANQML	DLI	S I	L	QFM	QNGLSRAEAEAQ		
C elegans	PQFDEVKRM	REANLVL	DII	S I	L	QFM	QNNLTRETLVDP		
B stearothermophilus	DRWDEVKEM	RRVNDLF	DIV	S V	M	LYM	QNNLTEQDIFER		
M tuberculosis	DRFEEIEEA	AGADRVL	RLV	T V	L	LAL	GAGVSADEFASD		
C glutamicum	DRFELIEDN	AAVNEML	RPT	S V	L	LHL	GAGVDPADFAAD		
Consensus	Y	G	KVTP	SK VGD	A	V			

S pombe	ATTLDF A	LDFFQ	LM Q	YG F	EPLRTNVLKGRRQPLTDPPGIFLP
S cerevisiae 1	ANSLDF D	MDFFE	LI Q	YG F	EPFRSDVLRNKRPFYLTCPPGLELE
S cerevisiae 2	ANSLDF D	VMDFFE	LI Q	YG F	EPLPSDVLRNKRRKLTCPPGLELE
R etli	AEELSF R	VEFLQ	YI I	HG F	EPFRSKVLKDLPP-IEGGPGASLP
P pastoris	ASELDF D	VEFFE	LM T	YG F	EPLRTNVISGKRRKLTSPPGLTLE
M musculus	AEELSF R	VEFLQ	YI I	HG F	EPFRSKVLKDLPR-IEGRPGASLP
H sapiens	AEELSF R	VEFLQ	YI V	HG F	EPFRSKVLKDLPP-VEGRPGASLP
C elegans	ADDLSF K	VDFMQ	NV Q	PY F	EPLPTKVLRGKPK-VDGPPGENAF
B stearothermophilus	GETLNF D	VELFE	YL Q	HG F	KELQRIILKGREP-ITVRPGELLE
M tuberculosis	PAPFGI E	LGFLR	EL D	PG W	EPLRTAALAGRGA------AR-P
C glutamicum	PQKYDI D	IAFLR	EL N	PG W	EPLRTRALEGPSE-------GK-A
Consensus	P SV	G	GP	GF	

S pombe	AADFDAIRKLLSEKFG-VSSDCDIAAYTQF	GVFEEYRQFVDRY	DLTTV
S cerevisiae 1	PFDLEKIREDLQNPFG-DVDECDVASYNMY	RVYEDFQFMPETY	DLSVL
S cerevisiae 2	PFDLEKIPEDLQNRFG-DIDECDVASYNMY	RVYEDFQKIRETY	DLSVL
R etli	PLNLKELEKDLIDRHGEEVTPEDVLSAAMY	DVFAQFKDFTATF	PLDSL
P pastoris	PYNIPAIREDLEARFS-KVTENDVASYNMY	KVYEAYKKQQELY	DLSVL
M musculus	PLNLKELEKDLIDPHGEEVTPEDVLSAAMY	DVFAQFKDFTATF	PLDSL
H sapiens	PLDLQALEKELVDPHGEEVTPEDVLSAAMY	DVFAHFKDFTATF	PLDSL
C elegans	PVDLDAVKVELEEKHGRTLSEEDVMSYSMF	TVFDEFETFPQQY	PVDFL
B stearothermophilus	PVDFEQMKKELYDKLGPEVTDFDAIAYALY	KVFLEYAETVEKY	DVSVL
M tuberculosis	TAQLAADDEIALSSVG--AKPQATLNRLLF	SPTKEFNEHPEAY	DTSQL
C glutamicum	PLTEVPEEEQAHLDADDSKEPRNSLNRLLF	FPTEEFLEHPRPF	NTSAL
Consensus			G

S pombe	PTKF LSPPEMNE	---MHVEIDQGFTLIVKFVALGPLNPPT	Q E	YFE
S cerevisiae 1	PTPS LSPLETDE	---IEVVIEQGKTLIIKLQAVGDLNKFT	E E	YFD
S cerevisiae 2	PTKN LAPAEPDE	---IEVTIEQGKTLIIKLQAVGDLNFKT	Q E	YFE
R etli	NTPL LQGPKIAE	---FEVELERGKTLHIFALAVSDLN-PA	Q Q	FFE
P pastoris	PTRN LSPPKIDE	PHVTIVTIETPKTLIIKCMAEGELSQSS	T E	YFE
M musculus	NTPL LQGPKIAE	---FEVELEPGKTLHIKALAVSDLN-RA	Q Q	FFE
H sapiens	NTRL LQGPKIAE	---FEVELEPGKTLHIKALAVSDLN-RA	Q Q	FFE
C elegans	PTPL LTGLEIAE	---VDVEIESGKTLAIQLLAEGFLN-KR	E E	FFD
B stearothermophilus	DTPT LYGMRLGE	---IEVEIERGKTLIVKLVSIGQPQ-AD	T V	YFE
M tuberculosis	SANQ FYGLRQGE	---HPVFLEPG-VELLIGLEAISEPDER	M T	MCI
C glutamicum	CDPE FYGLVEGR	---TLIPLPDVRTPLLVPLDAISEPDDK	M N	VAN
Consensus	G			

S pombe	L EN HVTVE	KKAAIETVFPPR	DPGNPGHVA	PMS	TIVEIPVKEGA
S cerevisiae 1	L EM KIPVA	PSQKVETVTKSK	DMHDPLHIG	PMA	VIVEVKVHKGS
S cerevisiae 2	L EL KIPVA	FSQNIQSVAFKS	DVHDTHQIG	PMA	VIIEVKVHKGS
R etli	L QL LVK	TQAMKEMHFHPK	LKDVKGQIG	PMP	KVIDVKVAAGA
P pastoris	L EM KVTVE	KNGAVETITPPK	DAHNPNEIG	PMA	VVVEVPVHENG
M musculus	L QL SILVK	TQAMKEMHFHPK	LKDVKGQIG	PMP	KVIDIKVAAGD
H sapiens	L QL SILVK	TQAMKEMHFHPK	LKDVKGQIG	PMP	KVIDIKVVAGA
C elegans	L QM SILVK	FEASFEIVTRPR	LPGLVPGHIG	PMP	DVLELKIKEGE
B stearothermophilus	L QP EVVIR	ESIKTAVVEHIK	DRTNPNHIA	TMP	TVVFVLVEFGE
M tuberculosis	L QL PVLVR	PSIASAVPAAEK	DPGNPGHIA	PFA	-VVTVGVCVGE
C glutamicum	V QI PMPVR	PSVESVTATAEK	DSSNFGHVA	PFA	-VVTVTVAEGD
Consensus	N G	R	D	A	A G

S pombe	FVKKGDIIAVLS	IVISAPHS	VLKSLAVVQGDSVNGG	CAVLEH
S cerevisiae 1	LIKKGQPVAVIS	MIISSPSD	QVKEVFVSDGENVDSS	LVLLED
S cerevisiae 2	LVKKGESIAVLS	JVSSPAD	QVKDVFIKDGESVDAS	LVVLER
R etli	KVVKGDQPLCVLS	TVVTSPME	TIPKVHVTKDMTLEGD	ILEIE-
P pastoris	EVKFGDPIAVLS	MVISSPVA	PIGQIAVKENDSVDAS	IPKSSP
M musculus	KVAKGQPLCVLS	TVVTSPME	TIPKVHVTFDMTLEGD	ILEIE-
H sapiens	KVAFGQPLCVLS	TVVTSPME	TVPKVHVTKDMTLEGD	ILEIF-
C elegans	FVTKFQPLFVLS	MVIDSPIA	TVFAIHAPQGTFCSAG	VVEVEP
B stearothermophilus	FVDFGDHLMVTE	TTVQAPFA	VVKDIYVKSGDAICAG	LMELSP
M tuberculosis	RVGAGQTIATIE	APITAPVA	TVEPVAVSDTACVEGG	LVVVS-
C glutamicum	EVKAGDAVAIIE	ATITASVD	FIDPVVVPAATFVEGG	IVVVS
Consensus	AHKML	G		DL

S pombe E-------------

S. cerevisiae 1	Q V P V E T K A - - - - -
S. cerevisiae 2	E T L P P S Q K K - - - - -
R. etli	- - - - - - - - - - - - -
P. pastoris	L S K L L M F I I L I I L Y
M musculus	- - - - - - - - - - - - -
H sapiens	- - - - - - - - - - - - -
C. elegans	- - - - - - - - - - - - -
B. stearothermophilus	- - - - - - - - - - - - -
M. tuberculosis	- - - - - - - - - - - - -
C glutamicum	- - - - - - - - - - - - -
Consensus	

CHAPTER 3: PYRUVATE CARBOXYLASE OVEREXPRESSION

3.A Introduction

3.A.1 Chapter Overview

In the previous chapter the discovery of the gene of pyruvate carboxylase was described. The sequence of the gene of pyruvate carboxylase came as a conclusion to a long experimental effort that focused on establishing the presence of this enzyme in *C. glutamicum*, but also created the need to further elucidate its role in the physiology of this organism. This is the central idea behind the research topic presented in this chapter.

The effect of pyruvate carboxylase on amino acid production and growth rates were studied using the lysine overproducers *C. glutamicum* strain 21253 and *C. lactofermentum* strain 21799. The first strain produces lysine after threonine has been completely depleted from the extracellular medium while the production of lysine from the second strain is independent of the presence of threonine. This chapter describes the construction of four recombinant strains that overexpress pyruvate carboxylase as well as a pyruvate carboxylase deficient strain. From the four recombinant strains one was able to give a significantly higher translation level of pyruvate carboxylase while the other three gave intermediary or no overexpression of this enzyme. The strain with the significantly higher pyruvate carboxylase overexpression was used to study the physiological effect that this enzyme has on *C. glutamicum* in combination with the physiological evidence from the pyruvate carboxylase deficient strain.

In the second half of this chapter, productivity results are presented from batch cultures of the parental, control, recombinant and mutant strains. More specifically the physiological data presented include growth, carbon source uptake, threonine depletion and lysine production. Other fermentation parameters are then calculated in order to assess the final differences between the different strains. In the discussion, conclusions and speculations are made on the internal cell physiology.

3.A.2 Transconjugation

Gene disruption and replacement techniques are well-established tools for the genetic engineering of Gram-negative bacteria. The manipulation of the genome of Gram-positive strains however is often hampered by the lack of efficient gene transfer systems, since many strains can not be readily transformed by routine transformation or electroporation procedures (Schafer, Schwarzer et al. 1994). Studies have shown that bacterial conjugation works throughout the microbial world (Davies 1990) (Mazodier and Davies 1991). The essential ingredients for the conjugal plasmid transfer are the broad host range transfer and mobilization functions of plasmid RP4, transposable elements, mainly Tn5, and usual *Escherichia coli* specific vector plasmids (Simon, O'Connell et al. 1986). RP4 is able to promote mobilization of plasmids from *E. coli* at high efficiency into several Gram$^+$ genera, like *Streptococcus, Bacillus, Staphylococcus, Mycobacterium, Streptomyces, Corynebacterium, Brevibacterium, Arthrobacter* or *Rhodococcus* (Trieu-Cuot, Carlier et al. 1987) (Mazodier, Petter et al. 1989) (Schafer, Kalinowski et al. 1990).

The most useful mobilizing donor is strain S17-1. It is a *rec*A- deficient derivative of *E. coli* 294 (*thi, pro, hsd*R⁻) (Simon, O'Connell et al. 1986). The natural antibiotic resistance markers of the chromosomally integrated RP4 in S17-1 have been inactivated and therefore all the resistance markers of usual *E. coli* vector plasmids can be used in this strain.

3.A.3 *Corynebacterium glutamicum* replication origins

The presence of small plasmids has been reported from several strains of corynebacteria, including *B. lactofermentum, Corynebacterium callunae, Corynebacterium lilium, C. glutamicum, Brevibacterium linens* and *Corynebacterium renale*. These smaller plasmids are generally cryptic, since they do not possess any known phenotypic markers (Deb and Nath 1999). In the past decade, two naturally occurring plasmids have been successfully characterized and used to construct cloning vectors: the

4.4 kb *Brevibacterium lactofermentum* plasmid pBL1 (Miwa, Matsui et al. 1985) (Santamaria, Gil et al. 1985) and the 3 kb plasmid pSR1 from *C. glutamicum* (Archer *et al.* 1993; Yoshihama *et al.*, 1985). The nucleotide sequences of both pBL1 and pSR1 have been reported (Archer and Sinskey 1993) (Yamaguichi, Terabe et al. 1986) (Fipula, Ally et al. 1986). These two plasmids have been used extensively in the past to construct artificial chimeras encoding two complete sets of replication functions, one for *C. glutamicum* and one for *E. coli* where all the cloning procedures are performed.

Only a few naturally occurring plasmids that have an inherent ability to replicate in diverse hosts have been described to date. One of these is pNG2, a 14.4 kb plasmid isolated from *Corynebacterium diphtheriae* (Schiller, Growman et al. 1980) (Serwold-Davis and Growman 1986). Moreover a 1.85 kb derivative of pNG2, pEP2, retains the broad host range of its parental plasmid, having been shown to replicate in corynebacteria, both fast- and slow- growing mycobacteria, and *E. coli* (Radford and Hodgson 1991). Because of its small size and its broad host range, the replication site of pEP2 appears to be much more convenient to be used for *C. glutamicum* vector construction.

The construction of new *C. glutamicum* vectors and its use for cloning purposes is described further. Moreover the cloning and transconjugation procedures for the generation of the *pc⁻* mutant are detailed.

3.B   Materials and Methods

3.B.1 Bacterial strains and plasmids
The strains employed for this work have been described in the previous chaprter. The bacterial strains were stored on LB5G agar plates (*E. coli* strains), or BHI plates (*C. glutamicum* strains). For long-term storage, frozen liquid stocks were maintained at -85°C in 80% (v/v) LB5G medium (*E. coli*) or BHI (*C. glutamicum*) and 20% glycerol.

C. glutamicum ATCC 21253 (U.S. Patent 3,708,395) from the American Type Culture Collection (ATCC) was the primary parent-type strain. This strain is auxotrophic for L- homoserine (or L-methionine and L-threonine) and L-leucine; therefore it requires

these substrates for growth on defined medium. Optimal growth temperature for this strain is 30°C.

E. coli strain TOP10 purchased from Invitrogen Inc. was used as a host strain for plasmid preparation [Genotype: F⁻, *mcrA* Δ (*mrr-hsd*RMS-*mcr*BC) φ80*lacZ*ΔM15 Δ*lac*X74 *deo*R *rec*A1 *ara*D139 Δ*(ara-leu)* 7697 *gal*U *gal*K *rps*L(Str^R) *end*A1 *nup*G].

3.B.2 Media Preparation

Luria-Bertani (LB) broth supplemented with glucose was the rich medium used for *E. coli* cultivation. This medium which will thereafter been referred to as LB5G contains 5 g/l yeast extract, 10 g/l tryptone, 5 g/l NaCl and 5 g/l glucose. Ampicillin, chloramphenicol and kanamycin were the antibiotic markers used for selection purposes of plasmids in *E. coli* (at a concentration level of 50 µg/ml) and kanamycin and chloramphenicol were used as selection markers for *C. glutamicum*.

The rich medium used for *C. glutamicum* cultures was Brain Heart Infusion, referred to as BHI thereafter. It contains 40 g/l BHI, 40 g/l sorbitol and 10 g/l sucrose. A mixture of 80 g/l sorbitol and 20 g/l sucrose is first autoclaved separately from a BHI solution of 80 g/l and the two parts are then mixed at equal portions to give the final complex medium.

Another complex medium used for culturing both *E. coli* and *C. glutamicum* was referred to as MB and contains 5 g/l of yeast extract, 15 g/l of tryptone, 5 g/l of soytone and 5 g/l of NaCl.

The defined medium used for *C. glutamicum* cell culture was adopted from Shaw-Reid (1997). Medium composition adopted for small shake flask cultures is listed in Table 2.3. Portions A, B and C were sterized by autoclaving separately and then mixed at equal portions. Part D was sterilized by filtration and added to the appropriate final concentration.

3.B.3 Quantification of Extracellular Metabolites

Organic compound detection

Organic compounds used for carbon source, such as glucose and lactate were measured by high performance liquid chromatography (HPLC) on a Bio-Rad Aminex HPX-87H reverse phase column (Bio-Rad Laboratories, Hercules, CA) operating at 45 °C with a 5 mM H_2SO_4 mobile phase. The Waters HPLC instrument (Bedford, MA) was outfitted with a 501 pump, 680 pump controller, WISP 710 automatic injector and an R401 differential refractometer. The flow rate through the HPX-87H column was 0.6 ml/min and the typical sample injection volume was 30 µl.

Amino Acid Detection

All amino acids were analyzed as *ortho*-phthaldialdehyde (OPA) derivatives using a Hewlett-Packard reverse phase AminoQuant column on a series 1050 HPLC system (Hewlett-Packard, Wilmington, DE). The column ran in a gradient mode at 40 °C with a method developed by Hewlett Packard, and data were obtained by monitoring UV absorbance at 338 nm. Buffer flow rate was 0.45 ml/min, and the composition of the filtered buffers was:

Buffer A: 20 mM Na acetate, 50 mM tetrahydrofuran, 2 mM triethylamine

Buffer B: 20 mM Na acetate (20% by volume), methanol (40%), acetonitrile (40%).

3.B.4 Molecular Biology Methods

Escherichia coli **plasmid minipreparation-DNA extraction**

Plasmid isolation from *E. coli* was performed using the standard QIAGEN miniprep, maxiprep and gigaprep kits. DNA fragments were isolated from 1% agarose gels using the QIAEX II Gel Extraction Kit for DNA purification.

Corynebacterium glutamicum **minipreparation of plasmid DNA**

Recombinant strains of *C. glutamicum* are grown in 5 ml rich medium (MB or BHI) overnight. Cells are resuspended in 0.25 ml of solution E (40 mM Tris; 2 mM EDTA, pH 8.0) and 0.25 ml of 1M sorbitol containing 3 mg/ml lysozyme. Mutanolysin is then added from a stock of 1 mg/ml and the suspension is incubated at 37 °C for 1 hour. To that solution, 50 µl of 0.5 M EDTA is added and 0.1 ml of 10% SDS. The mixture is boiled for 3 min and immediately returned to ice. 0.1 ml of 5 M NaCl is then added and incubation on ice follows for 10 min. After spinning at 14,000 rpm for 40 min a phenol/chloroform extraction and a double chloroform extraction follow. 0.6 volumes of isopropanol are then added and the plasmid DNA is precipitated by centrifugation at 13,000 rpm for 25 min. The pellets are air dried and dissolved is 0.2 ml of Tris/EDTA buffer. 1/10 volume Na acetate is added together with two volumes of room temperature ethanol. The DNA is precipitated by cooling on ice for at least 30 min and then spinning at 13 K for 25 min. The precipitated DNA is then dissolved in a minimal solvent volume (Tris/EDTA or water) and electroporated into *E. coli* competent cells. The restriction and PCR profile of the plasmid is then analyzed after it has been isolated from the *E. coli* host strain.

Preparation of *C. glutamicum* competent cells for electroporation

C. glutamicum cells are grown overnight (5 ml culture) and inoculated the next day in 50 ml MB broth (initial absorbance at 600 nm of 0.1). The cells are then grown for about four to five hours to a final absorbance of 0.6-0.8. At that point, 4 µl of ampicillin stock solution of 50 mg/ml is added to the flask and the cells are grown for another hour. The cells are then collected by centrifugation and the pellet is washed three times with buffer EPB1 (20 mM Hepes, 5% glycerol, pH 7.2). After the final centrifugation the cells are resuspended with 1.5 ml buffer EPB2 (5 mM Hepes, 15% glycerol, pH 7.2) and split to 0.15 ml aliquots. The cells are stored at −85 °C.

Electroporation

The necessary aliquots of frozen competent cells are first thawed on ice. The plasmid DNA is then added (1-3 µl) and the mixture is incubated on ice for 5 min before transferred to chilled 2 mm electroporation cuvettes. The electroporation conditions are 200 ohms, 25 µFd, and 2.5 kV voltage (time constant usually between 3.5-5.0). Immediately 0.7 ml of BHI are added to the cuvette and the cells are incubated at 30 °C for 1 hour before platting them on BHI/Km selective plates.

3.B.5 Enzymatic Assays

A pyruvate carboxylase activity was first reported in *B. lactofermentum* crude extracts by disrupting the cells using sonication and then applying the pyruvate carboxylase- malate dehydrogenase or –citrate synthase coupled assay (Tosaka, Morioka et al. 1979). However, using this method, no pyruvate carboxylase activity was detected in *C. glutamicum* (Vallino and Stephanopoulos 2000) (Gubler, Park et al. 1994) (Peters-Wendisch, Wendisch et al. 1997) (Park, Shaw-Reid et al. 1997b). Peters- Wendisch *et al.* (Peters-Wendisch, Wendisch et al. 1997) proposed a new method based on a discontinuous glutamate-oxaloacetate transaminase- (GOT) coupled assay performed with hexadecyltrimethylammonium bromide (CTAB) permeabilized *C. glutamicum* cells. Using one mole of glutamate, one mole of pyruvate carboxylase- synthesized oxaloacetate is transformed into one mole of aspartate. The rate of disappearance of glutamate or increase in aspartate is then correlated to pyruvate carboxylase activity. The quantification of aspartate is done by converting it to its *o* –phthaldialdehyde derivative followed by reversed- phase HPLC (Schrumpf, Eggeling et al. 1992). It was later reported that due to the dilution of the culture samples required for the HPLC determination, the measure of newly synthesized aspartate could be very hazardous (Uy, Delaunay et al. 1999).

This general procedure was followed here for measuring pyruvate carboxylase activity. Cells collected from shake flask cultures are washed twice with 20 ml of 50 mM Tris- HCl buffer, pH 6.3 containing 50 mM NaCl and then resuspended to an optical

density (OD)$_{570}$ of about 150 in 100mM 4-(2-hydroxyethyl)-1-piperazine-ethanesulphonic acid buffer, pH 7.5 containing 20% glycerol (v/v). This suspension can then be stored at - 20°C.

For the permeabilisation step, the frozen cells were slowly thawed on ice and then mixed with a solution of 2.5% CTAB. The final concentration of CTAB is 0.3% and the duration time for permeabilisation is 1 min.

The permeabilised cell suspension was immediately used to assay pyruvate carboxylase activity by measuring transformation rate of pyruvate into oxaloacetate. 3 mg of cells are added to 1 ml of reaction mixture containing 100 mM Tris- HCl, pH 7.3, 25 mM NaHCO$_3$, 5 mM MgCl$_2$, 4 mM ATP and 0.5-3 mM pyruvate. The mixture is left to incubate at 30 °C for four different times: 0.5min, 1 min, 1.5 min and 2 min. The reaction is terminated by the addition of 80 μl of 30% (w/v) orthophosphoric acid, and cell debris is removed by centrifugation (25,000 g, 4°C, 15 min).

Pyruvate carboxylase activity is determined based on the concentration of the remaining pyruvate in the reaction mixture. This is done by using the pyruvate analysis kit (Sigma Aldrich Co). In a preweighed vial containing 1 mg of NADH, 2.2 ml of TRIZMA (Trishydroxymethylaminomethane 1.5 mol/l, 0.05% sodium azide) base solution is added. In a 1-cm lightpath cuvet 1.0 ml of sample solution and 0.25 ml of TRIZMA Base solution are pipeted and mixing is performed by inversion. Mixing at this step is essential, as the mixture must be brought to proper pH by thorough mixing before adding NADH. The absorbance at 340 nm vs water as reference is read and recorded. This is the initial absorbance. The value of this number should always be greater than 0.6 using a 1 cm cuvet. To the cuvet 0.025 ml of Lactate Dehydrogenase (1000 U/ml) are quickly added and the solution is mixed by several inversions. After approximately 2-5 minutes the new absorbance at 340 nm is read and recorded. This is the final absorbance. After 5 minutes the absorbance is reread in order to assure that a constant minimum value has been reached. Reaction is considered complete when absorbance decrease is less than 0.001/min. The absorbance change is determined by the difference between the initial and the final A. Pyruvate concentration is then calculated using the form

Pyruvate concentration (mmol/l)$= \Delta A_{340} \times 0.723$

This form is derived by generating a standard curve using the following pyruvate concentrations (mmol/l): 0.0225, 0.045, 0.0675, 0.09 and 0.1125. For generating the standard curve the pyruvate assay was performed in exactly the same way as described above. The plot of ΔA_{340} detremined by this method versus pyruvate concentration (mmol/l) is linear.

One Unit of pyruvate carboxylase activity is defined as the amount of enzyme converting 1 nmol of pyruvate per minute.

3.B.6 Experimental Results

Construction of *C. glutamicum* vectors pMAGK(-) and pMAGK(+)

C. diphtheriae plasmid pNG2 (14.4 kb) is a broad host range plasmid and has an origin of replication contained in a 2.6 kb fragment. This origin of replication functions in *Corynebacterium* sp. *Rhodococcus* sp. and in *E. coli.* Sequences essential for replication of this plasmid were analyzed in an 1.85 kb derivative, pEP2 (Radford and Hodgson 1991), which encodes a 483 amino acid protein, RepA, essential for its replication. PEP2 replicates by the rolling circle mechanism, as confirmed by accumulation of single stranded DNA after rifampicin treatment of *E. coli* cell (Zhang, Praszkier et al. 1994).

For the generation of a pEP2 derivative plasmid that contains the lacIQ repressor gene, *E. coli* vector pMAL-p2X was digested with restiction enzymes *Pfl*M I and *Bgl* I. The DNA fragment that contained the repressor was *Pfu* polished and inserted into the *Pst* I site of pEP2 that had previously been *Pfu* polished. Depending on the orientation of the lacIQ repressor, two vectors were thus generated: pMAGK(-) and pMAGK(+) (Figure 3.1).

Construction of PCR-based pyruvate carboxylase vectors

The polymerase chain reaction (PCR) was used to amplify the pyruvate carboxylase gene (*pc*) as a 3.6 kb DNA fragment. Two restriction recognition sites where

introduced to the PCR primers, an *Nhe* I at the upstream (beginning) primer and an *Xba* I site at the downstream (end) primer. The sequence of the beginning primer was 5'-CCCGCTAGCATTGAGAGGAAAACAAAAACCGATGT-3' and that of the end primer 5'-CCAATCTAGACTCACCCATCTCCCATCTTGGCCC-3'. PCR was performed using *PfuTurbo* as the DNA polymerase enzyme (Stratagene Cloning Systems, La Jolla, CA) in 1x cloned *Pfu* DNA polymerase reaction buffer under the following conditions: 1.25 mM dNTPs (each), 100 ng of cosmid IIIF10, 250 ng of each PCR primer and 2.5 U of *PfuTurbo*. The total PCR volume was 20 µl. A total of 25 PCR cycles were performed during the amplification with an annealing temperature of 58 °C.

The 3.6 kb DNA fragment that corresponded to the pyruvate carboxylase gene was visualized on a 0.8% agarose gel by performing standard nucleic acid electrophoresis. After its extraction from the gel the fragment was inserted into the *Srf* I restriction site of plasmid pCR-Script Amp according to standard procedures provided by Stratagene (Stratagene Cloning Systems, La Jolla, CA) (Figure 3.2). The resulting plasmid, pCR-Script (PC) was transformed into *E. coli* strain TOP10. Positive clones were identified initially by color selection derived from the *lacZ* gene activity present in the pCR-Script Amp vector and then by restriction digests.

Cosmid IIIF10 that contains the pyruvate carboxylase gene was digested with enzymes *Sal* I- *Pvu* I. An approximately 3.4 kb DNA band that corresponds to the internal part of the pyruvate carboxylase gene was extracted and ligated to plasmid pBC KS (-) digested with the same enzymes (Stratagene, La Jolla, CA) producing plasmid pBC KS (-)+ PCf. This plasmid was then digested with *Bgl* II- *Bsu36* I and a DNA fragment of approximately 2.9 kb that is internal to the *pc* gene was excised and introduced to a 3.6 kb DNA fragment that was produced by digesting pCR-Script (PC) with the same enzymes (Figure 3.3). The final construct, pCRPC-Script was then sequenced at its beginning and end in order to detect for the introduction of point mutations during the PCR and ligation procedures. No such mutations were detected.

The chloramphenicol resistance gene was removed from vector pMG110 as a 1.1 kb DNA fragment after digesting this plasmid with *Sal* I- *Sma* I and polished by *Pfu* treatment. This plasmid was then introduced into the *Srf* I site of pCR-Script (Amp)

producing plasmid pCRCm-Script. From here the chloramphenicol resistance gene was removed by double digest with *Not* I- *EcoR* V, polished and introduced next to the *pc* gene on plasmid pCRPC-Script that had previously been digested with *EcoR* V (Figure 3.4). The resulting plasmid pCRPCCm-Script was then digested with *Nhe* I- *EcoR* V and an approximately 5 kb DNA fragment that was produced was polished and introduced into the previously polished *Xho* I site of the *C. glutamicum- E. coli* vector pMG108. Selection for positive clones was done based on chloramphenicol resistance and further verified by restriction and PCR analysis. The derived plasmid was named pMG108 (Cm)+PC (Figure 3.5).

Cosmid based pyruvate carboxylase constructs

Cosmid IIIF10 was digested with restriction enzyme *Hind* III that does not cut the pyruvate carboxylase gene and three DNA fragments with molecular weights higher than 3.6 kb were generated, one of approximately 15 kb, one of approximately 9 kb and one of approximately 4 kb. It was shown by PCR analysis that the 9 kb DNA fragment was actually the one containing the pyruvate carboxylase gene and this was further inserted into the *Hind* III site of pCR-Script, generating plasmid pCR-Script+9kb. This plasmid was then digested again with *Hind* III and the 9 kb DNA fragement was inserted after polishing into the *Xmn* I site of plasmid pMAGK(-), generating plasmid pMAGK(-)+9kb (Figure 3.6). The presence of the pyruvate carboxylase gene on this plasmid was verified both by PCR as well as restriction digest experiments.

Plasmid pCR-Script+ 9kb was digested with the following eighteen enzymes: *Afl* II, *Avr* II, *BspH* I, *Dra* I, *EcoR* V, *Hind* III, *Kas* I, *Nco* I, *Nde* I, *Nhe* I, *Nsp* I, *Spe* I, *Sph* I, *Ssp* I, *Xba* I, *Xmn* I, *Apa* I, *Sac* I, none of which cuts the pyruvate carboxylase gene. Only one DNA fragment of about 4 kb in size had a molecular weight higher than 3.6 kb and was isolated from this digestion. It was also shown by PCR analysis that indeed this DNA fragment contained the pyruvate carboxylase gene. This DNA fragment was further polished by *Pfu* treatment and was introduced into the *Xmn* I site of plasmid pMAGK (-), generating construct pMAGK (-)+4kb that contains the complete gene of pyruvate

carboxylase (Figure 3.7). Again restriction digests and PCR analysis showed that this plasmid contains the pyruvate carboxylase gene.

Cosmid IIIG7 that also contains the pyruvate carboxylase gene was digested with enzyme *Hind* III. Of the two DNA fragments that were produced by this digest and had molecular weights above 3.5 kb it was shown that the larger band of about 12 kb contained the pyruvate carboxylase enzyme, again by PCR analysis. This fragment was inserted into the *Hind* III site of pCR-Script and produced plasmid pCR-Script+IIIG7pc. This plasmid was further digested with *Sca* I- *Ssp* I and a 7 kb DNA fragment that was produced and contained the pyruvate carboxylase gene was introduced into the *Xmn* I site of plasmid pMAGK (-), producing plasmid pMAGK (-)+7kb or pKD7 (Figure 3.8). The presence of the *pc* gene on this plasmid was proved again both by restriction as well as PCR analysis.

Construction of plasmids for generating pc deletion mutant of *C. glutamicum*

The polymerase chain reaction (PCR) was used to amplify a 1.8 kb DNA fragment of the pyruvate carboxylase gene. The sequence of the beginning primer was 5'-GAGGGGCCCGCTAGCTTGACGGTGCAGCTCAGCTCGGTGGCGAAATC-3' and that of the end (downstream, reverse) primer 5'-AGGGAGCTCGCTAGCGGCGACGGTGCTCGAGGAACTCTTCGGTTG-3'. An artificial *Apa* I and *Nhe* I site were introduced at the beginning (forward) primer while an artificial *Sac* I and *Nhe* I sites were introduced at the end (reverse) primer. Cosmid IIIF10 was used as the template for the PCR and *Taq* polymerase as the polymerase enzyme. PCR was performed for 35 cycles. The resulting PCR fragment was next inserted into the *Srf* I site of pCR-Script, after polishing with *Pfu*, generating plasmid pCRPCf-Script.

Plasmid pMG110 that contains the chloramphenicol acetyl transferase (Cm; *cat*) gene from Tn9 under the *neo* promoter from Tn5 was digested with *Sma* I- *Sal* I. The resulting 1.1 kb fragment that contains the *cat* gene was then polished with *Pfu* and introduced into the *Srf* I site of pCR-Script, generating plasmid pCRCm-Script.

Plasmid pCRPCf-Script was subsequently digested with *Not* I- *Pst* I and a 0.45 kb fragment that is internal to the pyruvate carboxylase gene fragment was removed and

replaced by the 1.1 kb DNA fragment that contains the *cat* gene, produced by a *Not* I- *Pst* I digest of pCRCm-Script. The generated plasmid was named pCRPCfCm-Script (Figure 3.9).

This plasmid (pCRPCfCm-Script) was then digested with *Nhe* I, polished with *Pfu* and introduced into the *Sca*I site of the β- lactamase (*amp*, ampicillin resistance determinant) of the mobilization vector pSUP301 (5.1 kb) (Figure 3.10). The resulting plasmid, pSUPCCfCm was first transformed into TOPF10 *E. coli* cells and verified by antibiotic selection and restriction analysis and then it was transformed into the mating donor strain *E. coli* S17-1 in preparation for transconjugation and disruption of the *pc* gene in *C. glutamicum* strain 21253.

Construction of pyruvate carboxylase deficient strain

Disruption of the pyruvate carboxylase gene in the chromosomal DNA of *C. glutamicum* was performed by transposon mutagenesis (transconjugation), a type of marker- exchange mutagenesis (Simon *et al*. 1983, Swarzer and Puhler, 1991, Schafer *et al.*, 1990). Clearly, transposon mutagenesis has great advantages compared to random mutagenesis, the most important of which is the targeting of specific physical location within the genome. This is because the exchange of the complete (functional) gene in the chromosome with the disrupted version of this gene happens through a homologous recombination event.

Transconjugation experiments were performed as described extensively previously, with two major changes:

1.   *C. glutamicum* recipient strains were grown in Epo medium (van der Rest, Lange et al. 1999) (Schafer, Schwarzer et al. 1994)

2.   The selection marker for the recombination event was kanamycin and not chloramphenicol.

The Epo medium used contains isonicotinic acid hydrazide (isoniazid), which acts as an inhibitor of the synthesis of mycolic acid, a component of the cell wall of Mycobacteria and close relatives. It probably affects the integrity of the cell wall in *C. glutamicum* and does not necessarily work in other organisms. Also for reasons that are

unclear at this point, we were unable to obtain chloramphenicol resistance expression in *C. glutamicum* even though it has been reported in the past. Due to the fact that a fragment and not the complete gene of pyruvate carboxylase was cloned on pSUP301 disrupted by the *cat* gene, we did not have to distinguish between single and double crossover events.

After the transconjugation experiment, about 48 colonies were obtained that were kanamycin resistant. The expected phenotype of a pyruvate carboxylase mutant is a minimal or very small growth on lactate minimal medium. When those colonies were inoculated in such a liquid medium, none was able to grow during the first 24 hours and about 12 of them showed a minimal growth after 48 hours.

From the colonies that were completely unable to grow, one was selected and inoculated into 5 ml BHI/Km liquid medium and left to grow overnight. The cell pellet that was collected was then washed with water and inoculated into glucose minimal medium. After a 24-hour growth cells were harvested and western blot analysis was performed with the derived protein crude extract. No pyruvate carboxylase band was detected while the acetyl CoA carboxylase band was evidently present, thus proving the disruption of the *pc* gene in the 253pc(-) mutant strain (Figure 3.13).

3.C Results and Discussion

3.C.1 Pyruvate carboxylase overexpression

A great effort has been devoted in the past in order to elucidate the role of anaplerosis on lysine production, as this is defined by the formation of oxaloacetate from various carbon sources. This effort initiates from physiological data using *C. glutamicum* strains as well as carbon flux distribution measurements at the pyruvate principal node in *C. glutamicum* during lysine overproduction.

It has been shown (Vallino and Stephanopoulos 1994) that lysine yield is not pyruvate limited, based on fermentation evidence using pyruvate dehydrogenase attenuated mutants. It was later shown using [13]C- labeled carbon substrates that the

pyruvate carboxylating pathway is probably nine times more active than the PEP carboxylating pathway (Park 1996). The combination of these observations suggested the attenuation of the pyruvate carboxylase enzyme as an interesting research topic to be investigated, both by overexpressing and by completely removing it from lysine producing *C. glutamicum* strains.

The experimental evidence that we have obtained on the subject of pyruvate carboxylase overexpression with this study is really intriguing. It appears that the overexpression of pyruvate carboxylase is related to the presence of certain DNA sequences upstream of the gene of this enzyme. As it has already been mentioned plasmid pMG108 (Cm)+PC that included a PCR generated pyruvate carboxylase gene gave no significant overexpression based on western blot evidence. This lack of amplification may stem from base-pair mismatch errors in the PCR. This led to the construction of the second PCR based plasmid (pMG108 (Cm)+PC) that has the central part of the pyruvate carboxylase gene (from *Bgl* II to *Bsu36* I) completely derived from the cosmid library. No significant overexpression was detected either, indicating that an upstream or downstream DNA region may be critical for the overexpression of this enzyme.

To further study the effect of such regions in the case of pyruvate carboxylase of *C. glutamicum* another two additional constructs were made, one with the pc gene within a 4 kb DNA fragment and another with the gene within a 9 kb DNA fragment. Both these fragments contained different lengths of DNA regions upstream and downstream of the pyruvate carboxylase gene and none yielded any significant enzymatic overexpression. In the case of the 9 kb fragment the region upstream of the start codon of the gene was about 0.7 kb in length, as estimated from restriction analysis.

The lack of significant overexpression of pyruvate carboxylase either by PCR or by directly isolating the gene from cosmid IIIF10 Suggested that an alternative approach should be used. To this end we switch our efforts to the use of a different cosmid, and as such cosmid IIIG7 was selected. This selection was based on the fact that a *Hind* III digest of IIIG7 yielded a larger pc containing DNA fragment of about 12 kb compared to cosmid IIIF10. This fragment was first ligated into pCR-Script and, because we were unable to further insert it into vector pMAGK (-), its size was first decreased by a *Sca* I- *Ssp* I

digest. The resulting blunt ended 7 kb fragment was then introduced to the *Xmn* I site of pMAGK (-) and the final construct, after introduction into *C. glutamicum* and *C. lactofermentum*, gave a significant pyruvate carboxylase overexpression (Figures 3.11 and 3.12). Restriction analysis showed that this 7 kb fragment contains an approximately 2.3 kb DNA fragment upstream of the *pc* gene, much larger than the upstream region of the 9 kb DNA fragment. This finding suggests that perhaps an upstream or downstream region is critical for pyruvate carboxylase overexpression. This possibility is supported by evidence from other organisms. In *Saccharomyces cerevisiae*, such regions have also been identified that are involved in the regulation of the pyruvate carboxylase genes of this organism under different culture conditions. For example, in the case of the second isozyme of pyruvate carboxylase it was shown that elimination of a fragment from –417 to –291 from the start codon yielded a two-fold decrease in the expression under repressed conditions (Menendez and Gancedo 1998).

The possible regulatory regions of the construct originating from cosmid IIIG7 obviously merit further study starting with the sequencing of the upstream region of the pyruvate carboxylase gene.

3.C.2 Effect of pyruvate carboxylase overexpression on *C. glutamicum* cell physiology

The role of pyruvate carboxylase in lysine biosynthesis has been the center of speculation for a long time. Based on a plethora of phenotypic evidence, that could not be verified further due to the lack of the gene of this enzyme, a favorable influence of this enzyme on lysine production was anticipated. Furthermore, based on mathematical models and fermentation data, the enhancement of the anaplerotic reaction catalyzed by this enzyme was expected to cause a significant increase of the carbon flux that leads to lysine production. The experimental testing of these predictions has been one of the major goals of this work.

In the case of *C. glutamicum* (strain 21253) our findings seem to be pointing towards the direction of a special role for pyruvate carboxylase for growth and lysine production. The overexpression of this enzyme seems to have no effect on growth when

glucose is used as a sole carbon source. The specific growth rate of the control strain was almost identical to that of the recombinant (0.35 hr^{-1} versus 0.34 hr^{-1}, Table 3.2). It has been shown by enzymatic assays using permeabilized cells that the level of pyruvate carboxylase expression is minimal with glucose as carbon substrate (Peters-Wendisch, Kreutzer et al. 1998). This we were also able to verify by similar methods (Table 3.1). So it is possible that the absence of any physiological effect by glucose is due to the low level of pyruvate carboxylase expression during the fermentation. No enhancement in activity indicates that some kind of inhibition might exist in *C. glutamicum*. One possible explanation might be the inhibition effect of aspartate on the pyrubate carboxylase activity. Because strain 21253 has aspartokinase regulated by threonine and lysine, aspartate should be accumulated in the strain 21253 until the depletion of threonine and accumulated aspartate might, then, inhibit pyruvate carboxylase. In the case of the recombinant strain 253pKD7, higher level of accumulated aspartate might be resulting in higher inhibition level. For this reason, the resulting activity of pyruvate carboxylase seemed to be the same in the wild type and recombinant *C. glutamicum*. Despite the fact that there are no reported inhibition effects of aspartate on pyruvate carboxylase activity in *Corynebacterium* sp., various kinds of prokaryotes, including *Pseudomonas aeruginosa, Streptococcus faecalis, S. faecium, S. lactis, Micrococcus cerificans, Bacillus megatrium, B. sutilis, B. lichenformis* and *B. stearothermophilus*, have been shown to have pyruvate carboxylase inhibited by aspartate (Al-ssum and White 1977), (Hartman 1970), (Hiller and Jago 1978), (Lachica and Hartman 1969), (Libor, Sundaram et al. 1978), (Phibbs, Feary et al. 1974), (Renner and Bernlohr 1972).

On the other hand when lactate was employed as a sole carbon source the physiological effect was much more dramatic. First, the level of pyruvate carboxylase in the recombinant strain was much higher than that of the parental strain (Table 3.1) something that has also been reported in the literature (Peters-Wendisch, Wendisch et al. 1997). Both the parental (21253) and the recombinant strain showed similar specific growth rate- 0.35 hr^{-1} for the control versus 0.41 hr^{-1} for the recombinant strain. Even though they were both able to grow almost identically during the exponential face, the major difference was observed during the stationary phase when lysine excretion is

initiated from the cell to the medium. The parental strain was only able to reach a final cell concentration of 2.5 g/l with 20g/l lactate in the medium while the recombinant strain reach a final level of 4 g/l (Table 3.2). It is obvious that the induction of the transcription or translation of pyruvate carboxylase by lactate is the reason why such an extensive difference is observed between the two strains.

These first phenotypic results based on our shake flask experiments prove that in *C. glutamicum* the carbon source used has a major effect on the physiology of the recombinant cells that overexpress pyruvate carboxylase. This does not come as a surprise to us. Such an effect has been observed in the past in other organisms such as *Pseudomonas citronellolis* and *Saccharomyces cerevisiae.*

In the case of *S. cerevisiae*, the expression of both isoenzymes (*pyc1* and *pyc2*) of pyruvate carboxylase is influenced by both the growth phase and the type of carbon source. On glucose minimal medium *pyc1* has a constant level of expression throughout the main growth phase compared to a high level of expression of *pyc2* only in the early growth phase. On ethanol minimal medium, the growth related pattern of *pyc1* and *pyc2* expression was similar and showed a decline from early to mid log phase. The expression of *pyc1* plays an important anaplerotic role in maintaining fermentative growth and more notably for the establishment of gluconeogenic growth. On the other hand *pyc2* expression seems to support growth on a glycolytic carbon source (Brewster, Val et al. 1994).

In the case of *P. citronellolis* it has been shown that the activity of pyruvate carboxylase is controlled by the carbon source of the growth medium (Taylor, Routman et al. 1975). The activity of the enzyme is highest in cells grown on lactate or glucose and virtually absent in cells grown on malate or aspartate. This study also showed that coordinated regulation occurs at the level of the synthesis of the two polypeptides, which make up pyruvate carboxylase in this strain rather than at the stages of their assembly into protomers or the biotinylation of the apoenzyme (Taylor, Routman et al. 1975). What is even more interesting is the fact that this pyruvate carboxylase shows no control of its catalytic activity via effectors, such as, for example, acetyl –CoA carboxylase, aspartate or palmityl-CoA carboxylase. In most varieties of pyruvate carboxylases examined so far,

the enzyme appears to be constitutive with regulation accomplished either through effector modulation of holoenzyme activity (pyruvate carboxylase from animal sources, yeast, several species of bacteria) or through control of the biotinylation of the apoenzyme by biotin ligase (*Bacillus stearothermophilus,* yeast), (Cazzulo, Sundaram et al. 1971),(Sundaram, Cazzulo et al. 1971), (Cazzulo, Sundaram et al. 1969).

It is interesting to note that a case similar to that observed with *P. citronellolis* is reported here with pyruvate carboxylase of *C. glutamicum*: the *C. glutamicum* pyruvate carboxylase does not seem to be influenced on the enzymatic level by allosteric effectors such as acetyl-CoA (a common *pc* activator), similarly to the pyruvate carboxylase of *P. citronellolis*. Further studies will be needed in order to elucidate whether the formation of pyruvate carboxylase in *C. glutamicum* is also controlled at the level of the synthesis of the polypeptides.

Other cases of control of pyruvate carboxylase synthesis have been reported in the literature. In *Azotobacter vinelandii* pyruvate carboxylase expression is dependent on the carbon source used for cell growth (Scrutton and Taylor 1974). In *Rhodobacter capsulatus* the synthesis of pyruvate carboxylase is (i) induced by pyruvate, (ii) repressed by TCA cycle intermediates, (iii) stimulated by aerobiosis, and (iv) directly correlated with growth rate. The sysnthesis of the *R. capsulatus* pyruvate carboxylase appears to be controlled by both positive (pyruvate) and negative (C_4 dicarboxylates) effectors. Thus in *R. capsulatus,* pyruvate carboxylase is potentially subjected to two levels of control: allosteric regulation of enzyme activity (Modak and Kelly 1995), and regulation of protein synthesis (Yakunin and Hallenbeck 1997).

In *C. glutamicum* the action of pyruvate carboxylase is opposed by oxaloacetate decarboxylase, that is, pyruvate is converted to oxaloacetate by pyruvate carboxylase and oxaloacetate is converted to pyruvate by oxaloacetate decarboxylase. Together these two reactions constitute a potential futile cycle with ATPase activity, in light also of the presence of another four enzymes on the PEP-pyruvate-oxaloacetate triangle. The induction of pyruvate carboxylase in this organism provides one mechanism for regulating the futile cycle. It would seem likely that the inducing compound would be pyruvate or a closely related metabolite. Thus in the presence of excess of lactate and so

pyruvate also, oxaloacetate formation may be stimulated by the induction of pyruvate carboxylase, thereby facilitating the oxidation of pyruvate. On the other hand, when glucose is used as a carbon source, phosphoenolpyruvate carboxylase provides an alternative route for carbon to enter the TCA cycle and thus does not favor to the same extend the expression of pyruvate carboxylase. These data help us classify the *C. glutamicum* pyruvate carboxylase together with the other two such enzymes from *P. citronellolis* and *Azotobacter vinlandii*. These two enzymes are not constitutive and exhibit no known activators or inhibitors. The main difference derives from the fact that the *C. glutamicum* enzyme appears to be composed of a single subunit arranged as a tetramer. This of course remains a speculation and it will require more *in vitro* enzymatic studies in order to be proved experimentally.

The most surprising result from the shake flask experiments with the recombinant and the parental (21253) strain has to do with the lysine production itself. A reasonable expectation for many years has been that an enhancement of the anaplerotic activity would lead to an enhancement of lysine production. The focus on pyruvate carboxylase became even more intense as soon as the expectations from the first anaplerotic enzyme, phosphoenolpyruvate carboxylase, proved fruitless. The finding that pyruvate carboxylation controls 90% of the carbon flux towards lysine was another indication for the importance of this enzyme.

With the discovery of the gene of pyruvate carboxylase, the validation of these predictions and the fulfillment of the expectations became an experimental possibility. The first carbon source to be tested, glucose, gave us an almost identical lysine production in comparison to the control strain. The final lysine concentration obtained is about 2.8 g/l, almost identical to the 3 g/l of the control production. This result, together with the fact that the growth profile of the two strains is identical shows that indeed there is no physiological effect of the pyruvate carboxylase overexpression on *C. glutamicum*. This can be a well-expected result since apparently the pyruvate carboxylase expression is not high under glucose growth.

On the other hand the results obtained with lactate as sole carbon source were quite more interesting: a final lysine concentration of 2.14 g/l was achieved with the

recombinant strain in comparison with 2.58 g/l of the control strain. The productivity of the recombinant strain was also dramatically lower, 23.0 mg lys/g cell-hr comparing to 63.0 mg lys/g cell-hr for the control strain (Table 3.2). It is obvious from these data that the anaplerotic reaction enhancement not only did not improve the lysine production as expected but led to a significant drop of the productivity.

In our effort to explain this type of physiological behavior the following metabolic model is proposed: the overexpression of the pyruvate carboxylase reaction leads to an increase of the carbon flux from glycolytic pathway to the TCA cycle. This increase means that more carbon from the TCA cycle is available to form biosynthetic precursors. The citric acid cycle is one of the major degradative pathways for the generation of ATP. But is also provides intermediates for biosynthesis. For example, a majority of the carbon atoms in porphyrins come from succinyl CoA. Other amino acids, besides the aspartate family, are derived from α-ketoglutarate (for example glutamate). So, a recombinant strain with an enhanced carbon flux in the TCA cycle means that the cell is provided with more energy and more biological precursors and that is why it is capable of growing better than the control strain (Figure 3.16).

Of course such a metabolic prediction would not stand if the enhanced metabolic flux were able to exit the TCA cycle at the oxaloacetate node and enter the lysine biosynthesis pathway. The finding that this is not the case with the pyruvate carboxylase shows that a potential bottleneck exists in this linear pathway that leads from aspartate to lysine. Such bottlenecks have been identified in the past and a further discussion and experimental evidence will be provided to that respect in the following chapters.

3.C.3 Effect of pyruvate carboxylase overexpression on *Brevibacterium lactofermentum* cell physiology

In order to elucidate further the effect of pyruvate carboxylase on cell physiology, plasmid pKD7 was also introduced into the *B. lactofermentum* strain ATCC 21799. This is a *Corynebacterium* AEC resistant mutant that possesses an aspartate kinase insensitive to feedback inhibition by the combined presence of lysine and threonine in the

fermentation medium. It has acquired this insensitivity and thus converted into high-lysine producer after extensive rounds of UV mutagenesis (Shiio, Yoshino et al. 1990) (Schrumpf, Eggeling et al. 1992).

The results that were obtained with this strain were predictable to an extend based on the previous results using *C. glutamicum* strain ATCC 21253. A major difference between the two strains is the pyruvate carboxylase activity when glucose is utilized as the sole carbon source: the recombinant strain showed an activity of 53 U/mg CDW compared to 30 U/mg CDW for the parental (21799) strain (Table 3.1). As a result of the higher activity of pyruvate carboxylase the recombinant and parental strains also showed different physiological profiles. Clearly the 21799(pKD7) strain grew faster ($\mu=0.23$ h^{-1}) and reached a higher biomass concentration, 8.0 g/l of cell dry weight (CDW) in minimal medium with glucose as carbon source compared to the parental strain, in contrast to the results obtained with strain 21253. The same numbers for the control strain (21799) were $\mu=0.18$ h^{-1} and maximum biomass concentration of about 4.2 g/l CDW (Table 3.2 and Figure 3.14).

For cells grown on lactate, again a much higher specific activity was shown for the recombinant strain (83 U/mg CDW) comparing to the control strain (36 U/mg CDW) (Figure 3.15 and Table 3.1). The recombinant strain again showed a slightly higher specific growth rate of 0.12 h^{-1} compared to 0.16 h^{-1} for the control strain and a maximum cell concentration of about 6.3 g/l compared to 4.1 g/l for the control strain (Figure 3.14&3.15). This is in agreement with the results previously obtained for strain 21253.

In terms of lysine production using glucose as carbon source, the control strain showed a slow and constant increment of extracellular lysine concentration with respect to the recombinant strain. The final lysine concentration obtained by both strains was almost the same, i.e. approximately 2 g/l. Cells grown on lactate showed lower maximal levels of extracellular lysine than those produced by cells grown on glucose. Unlike glucose- grown cell, the control cells elicited an abrupt release of L-lysine to the medium between 20 and 25 h of incubation in the presence of lactate. The unexpected result was the low lysine concentration (1 g/l) produced by the recombinant strain.

In *B. lactofermentum* as in *C. glutamicum* the overexpression of pyruvate carboxylase increases the biomass production and growth yields in general, while it has no or even negative effect on lysine production. Again two kind of perturbations were performed: (*i*) amplification of the pyruvate carboxylase activity by transformation of strain 21799 with a multicopy plasmid carrying the pyruvate carboxylase gene, and (*ii*) environmental perturbation, such as change in carbon source, e.g. glucose or lactate.

The increase of biomass concentration both in the case of glucose as well as lactate indicates the presence of a kinetic limitation in the pyruvate carboxylase reaction that had not been evaluated before. The overexpression of pyruvate carboxylase releases the cell growth from that limitation.

All four amino acids derived from aspartate require carbon skeletons that are derived from oxaloacetate. There are four reactions in *Corynebacteria* contributing to the formation of oxaloacetate, namely PEP carboxylase, PEP carboxykinase, oxaloacetate decarboxylase and pyruvate carboxylase. Also the enzymes of the glyoxylate shunt that produce one molecule of malate from two molecules of acetyl-coenzyme A are present in this strain. It is already known that in *Corynebacteria* species the enzymes of the glyoxylate cycle are subjected to glucose repression (Gubler, Jetten et al. 1994). It has been shown in the past that overexpression of the enzymes participating in the lysine biosynthesis pathway results in a decrease of growth. For example overexpression of dihydrodipicolinate synthase results in an increase of extracellular L-lysine accumulation from 220 to 270 mM and a simultaneous growth limitation (Eggeling, Oberle et al. 1998) while similar physiological effects have been reported in the case of the aspartate kinase overexpression in *B. lactofermentum* (Jetten, Follettie et al. 1995). Apparently, once biosynthetic fluxes are redirected to produce more lysine, a partial deprivation of oxaloacetate from the TCA cycle occurs accompanied by a similar limitation of other key anabolic intermediates. Thus, deregulation of the pathway directly associated with the synthesis of L-lysine induces an uncoupling between anabolic and catabolic processes. According to the data presented here, overexpression of pyruvate carboxylase produces a more coupled metabolism, improving the assimilation efficiency of carbon substrates for anabolic purposes. As mentioned earlier for the case of *C. glutamicum*, the increase of the

carbon flux through the pyruvate carboxylase pathway leads to new flux distributions at the key branchpoints of the PEP, pyruvate and oxaloacetate nodes. This new distribution, not counterbalanced by various regulatory effects, leads to a higher supply of carbon intermediates for the TCA cycle that provides major building blocks to the cell, such as porphyrins and amino acids.

The final question to be addressed is the phenomenon of lower lysine production in the case of *B. lactofermentum* grown on lactate minimal medium. As it has already been documented, the activity of pyruvate carboxylase is much higher when lactate is used as carbon substrate for cell growth. It is only natural then that the phenotypic effects are more apparent in the case of lactate than in the case of glucose. During lysine production and at the step of synthesis of L-piperideine-2,6-dicarboxylate, a molecule of pyruvate condenses with a molecule of L-2,3-Dihydodipicolinate by dihydrodipicolinate reductase. It is possible that the overexpression of pyruvate carboxylase decreases the pool of pyruvate available for this condensation reaction yielding a lower synthesis rate fc⁻ H$_4$D compared to the control strain. That eventually leads to a lower lysine production and it can explain to an extend the decrease in lysine production observed with lactate as carbon source. Again, this phenotypic behavior is not observable when glucose is utilized simply because of the much lower extend of overexpression of the pyruvate carboxylase enzyme when glucose is used as a sole carbon source.

In summary, in order to preserve a regular biomass composition, intermediary metabolism has evolved a coordination of pathway's function, such that building-block metabolites, energy in the form of ATP and reducing power in the form of NAD(P)H are synthesized in approximate stoichiometric ratios during coupled growth. Consequently, if overproduction of L-lysine demands significant flux alterations in central carbon metabolism compared to flux distributions under balanced growth, it results in product yields that are significantly less than the theoretically expected. The results derived from the single pyruvate carboxylase overexpression demonstrate that an overflow of carbon skeletons as oxaloacetate to biosynthetic pathways downstream pyruvate induces high biomass yields without significant changes in lysine yields. This is a clear indication of the complexity of cell metabolism and gives one more example of the value of metabolic

engineering: the ability to view metabolism in its entirety and derive metabolic models based on phenotypic evidence.
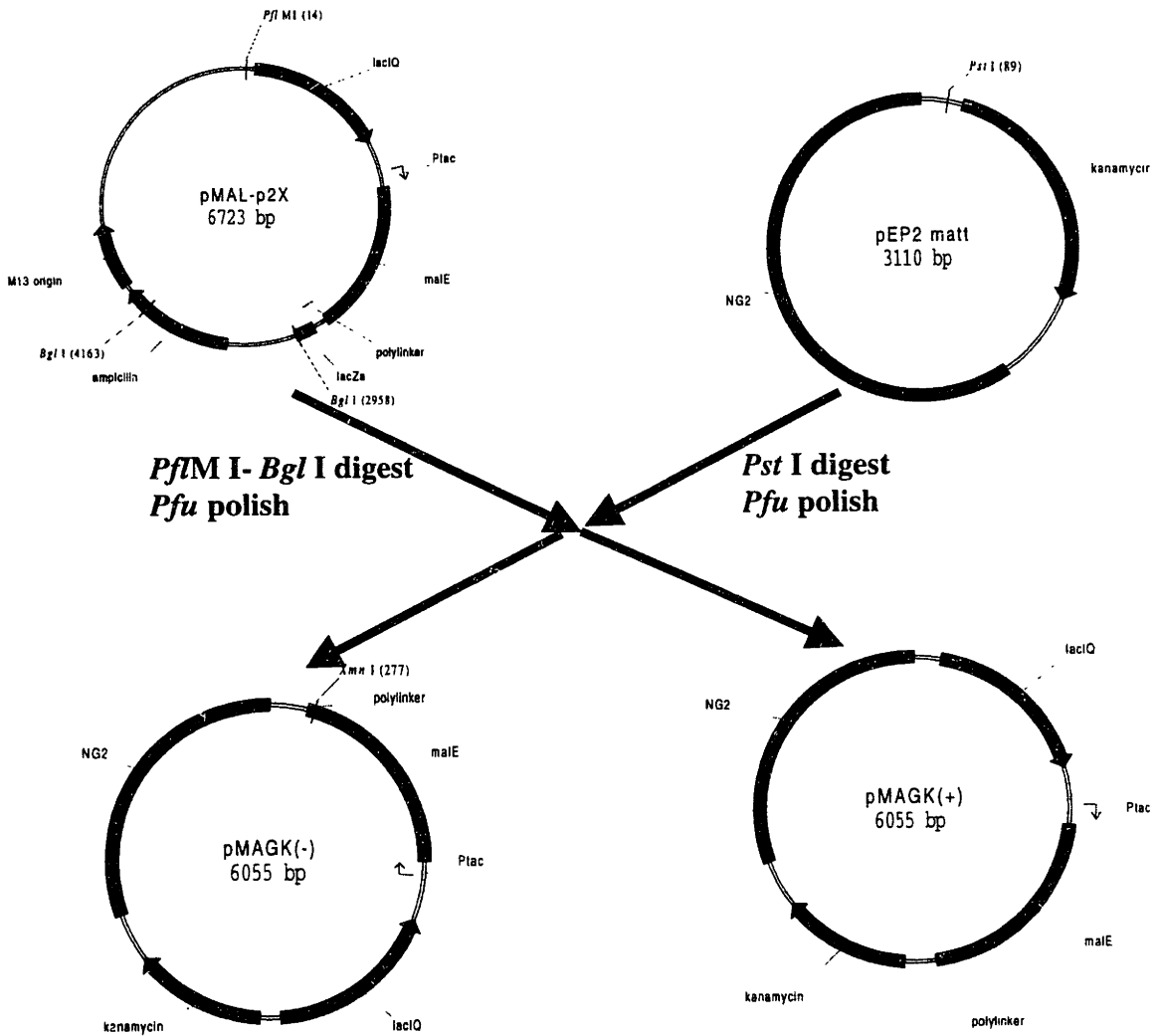
Figure 3.1 Construction of *C. glutamicum* vectors pMAGK (-) and pMAGK (+). Drawings are not to scale.
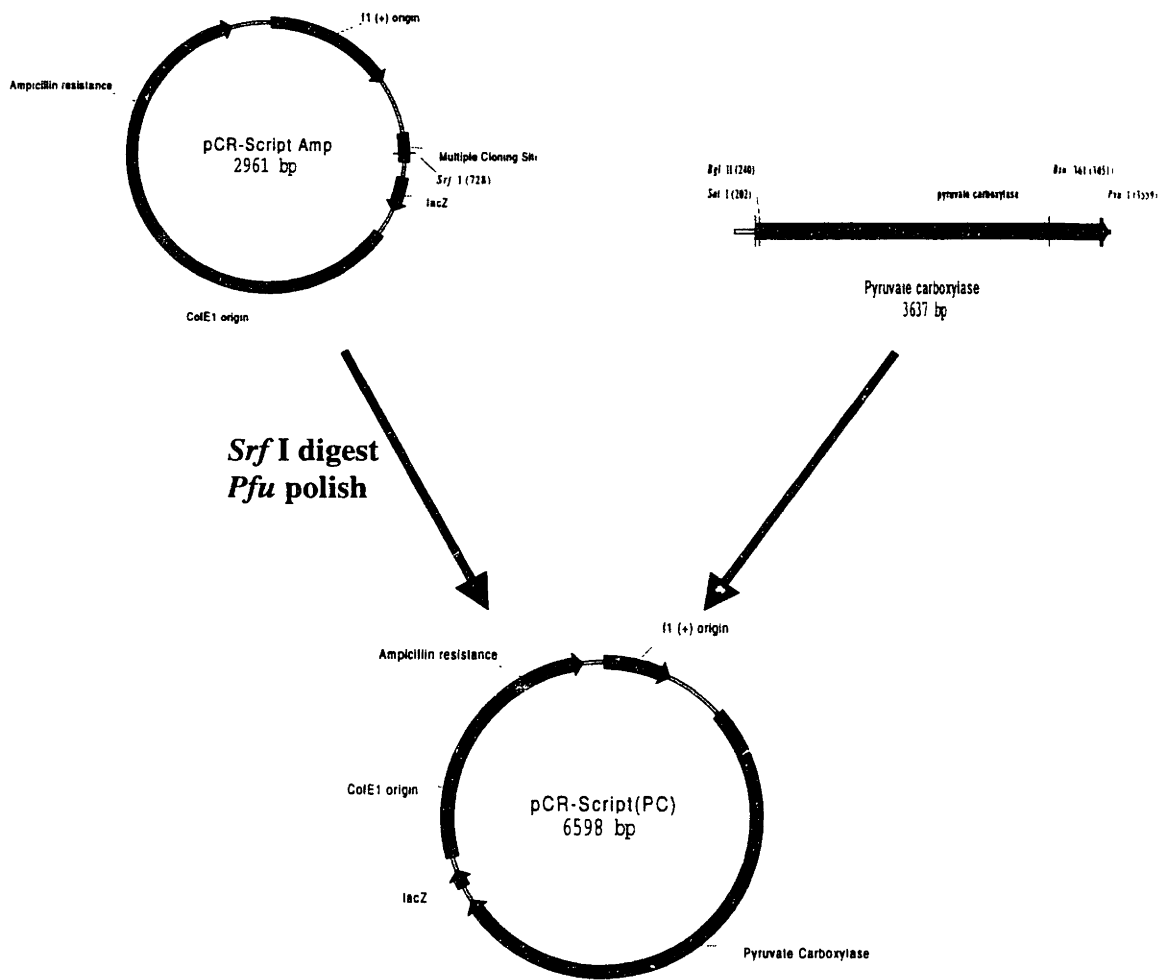
Figure 3.2 Construction of the *E. coli* vector pCR-Script(PC) that carries the *C. glutamicum* pyruvate carboxylase gene generated as a PCR fragment. Drawings are not to scale.

Pvu I (20018)
ampicillin resistance
neomycin resistance
SV40 ORI
cosmid IIIF10
20639 bp
Sal I (3638)
pyruvate carboxylase
Pvu I (6995)

I1(+) origin
Pvu I (501)
Multiple Cloning Site
pBC KS(-)
3400 bp
Sal I (735)
LacZ promoter
chloramphenicol gene
ColE1 ori

Sal I- Pvu I digest

Bsu 361 (6117)
I1(+) origin
Bgl II (539)
chloramphenicol gene
pBC KS(-)+PCf
6523 bp
ColE1 ori
LacZ promoter
pc fragment
Bsu 361 (3350)

I1 (+) origin
Ampicillin resistance
ColE1 origin
pCR-Script(PC)
6598 bp
lacZ
Pyruvate Carboxylase

Bgl II- Bsu36 I digest

I1 (+) origin
Ampicillin resistance
Bgl II (967)
ColE1 origin
pCRPC-Script Amp
6598 bp
lacZ
Pyruvate Carboxylas
Bsu 361 (3778)

Figure 3.3 Construction of the *E. coli* vector pCRPC-Script that carries the *C. glutamicum* pyruvate carboxylase gene. Drawings are not to scale.

102
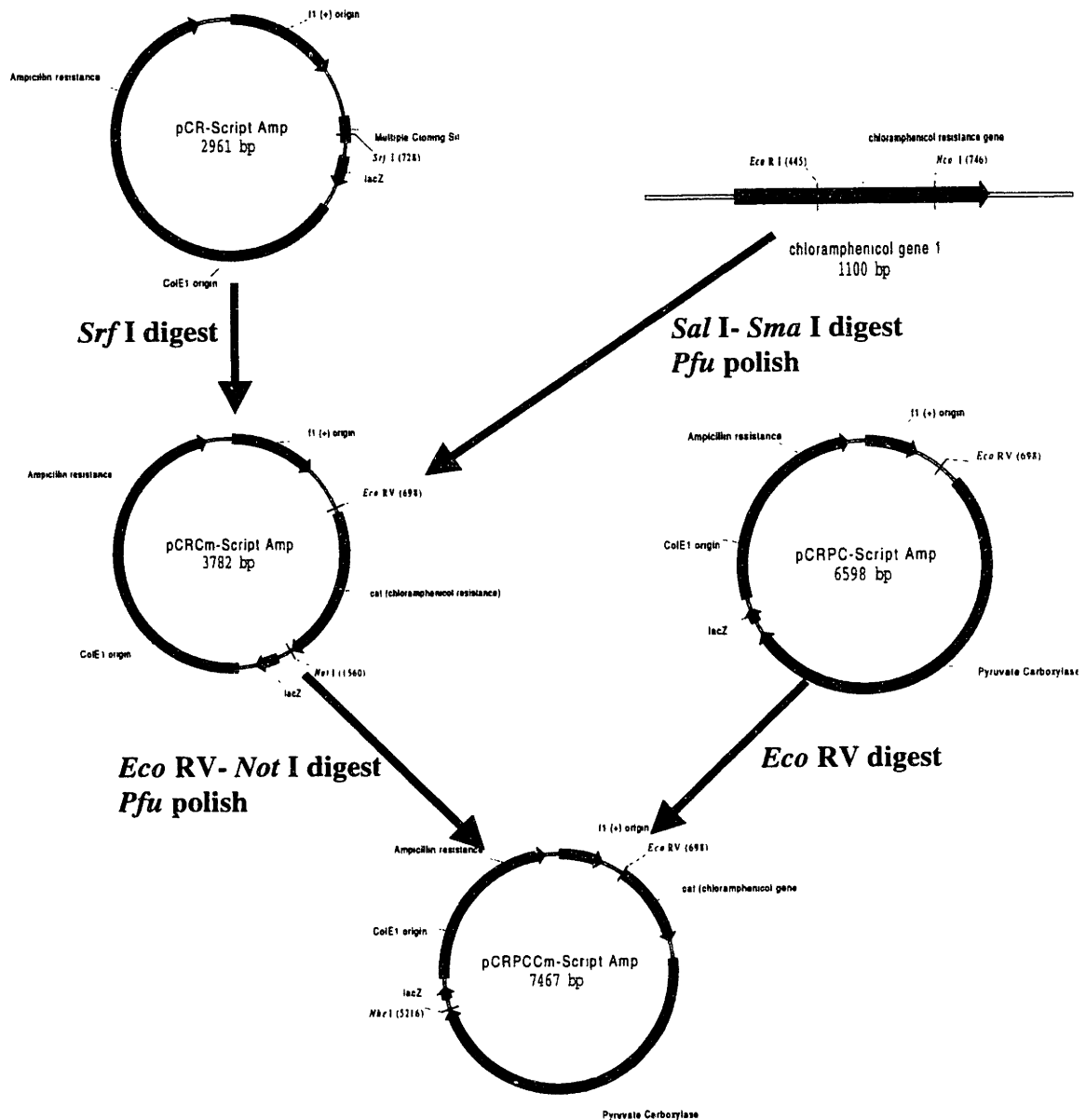
Figure 3.4 Construction of the *E. coli* vector pCRPCCm-Script that carries the chloramphenicol resistance determinant adjacent to the *C. glutamicum* pyruvate carboxylase gene. Drawings are not to scale.
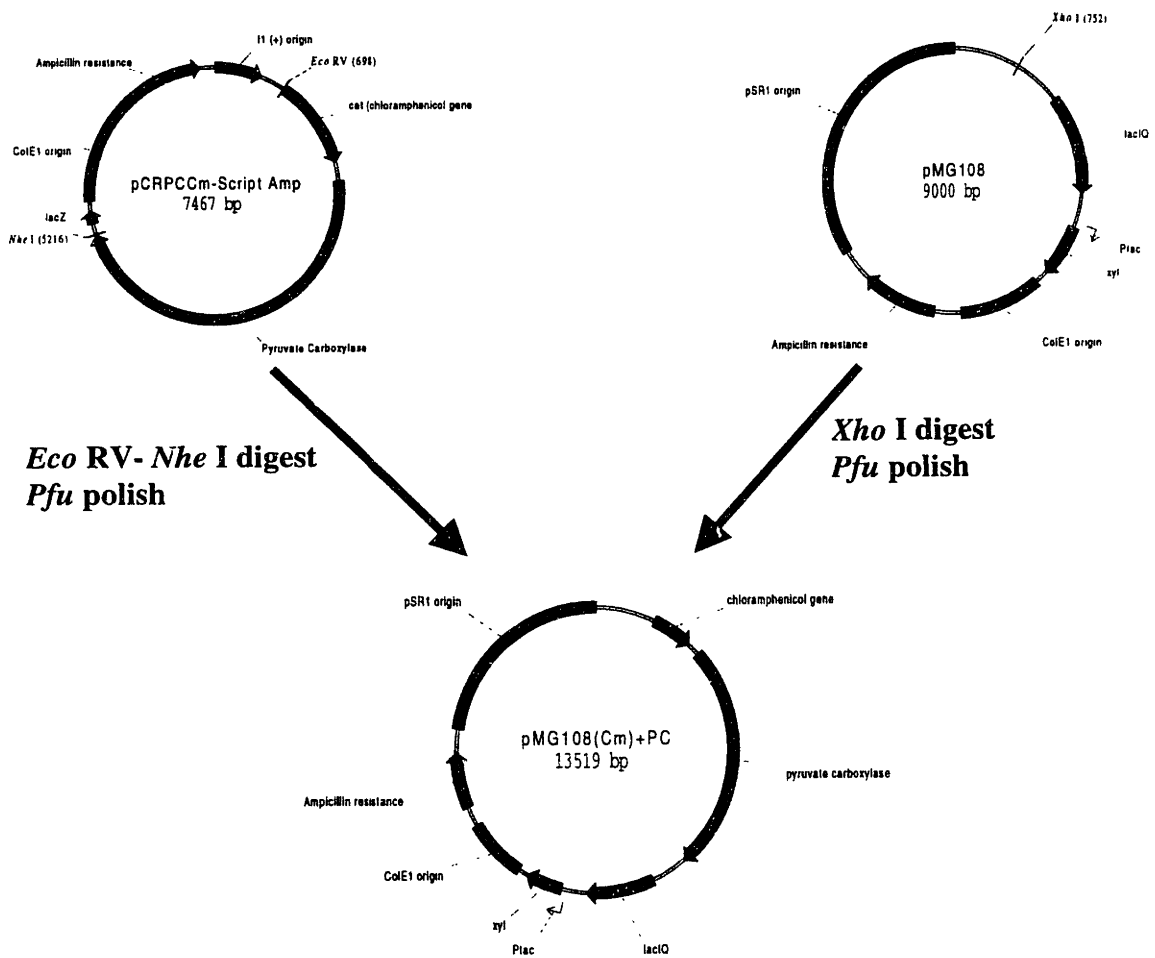
Figure 3.5 Construction of the *C. glutamicum- E. coli* shuttle vector pMG108 (Cm)+PC that carries the *C. glutamicum* pyruvate carboxylase gene. Drawings are not to scale.
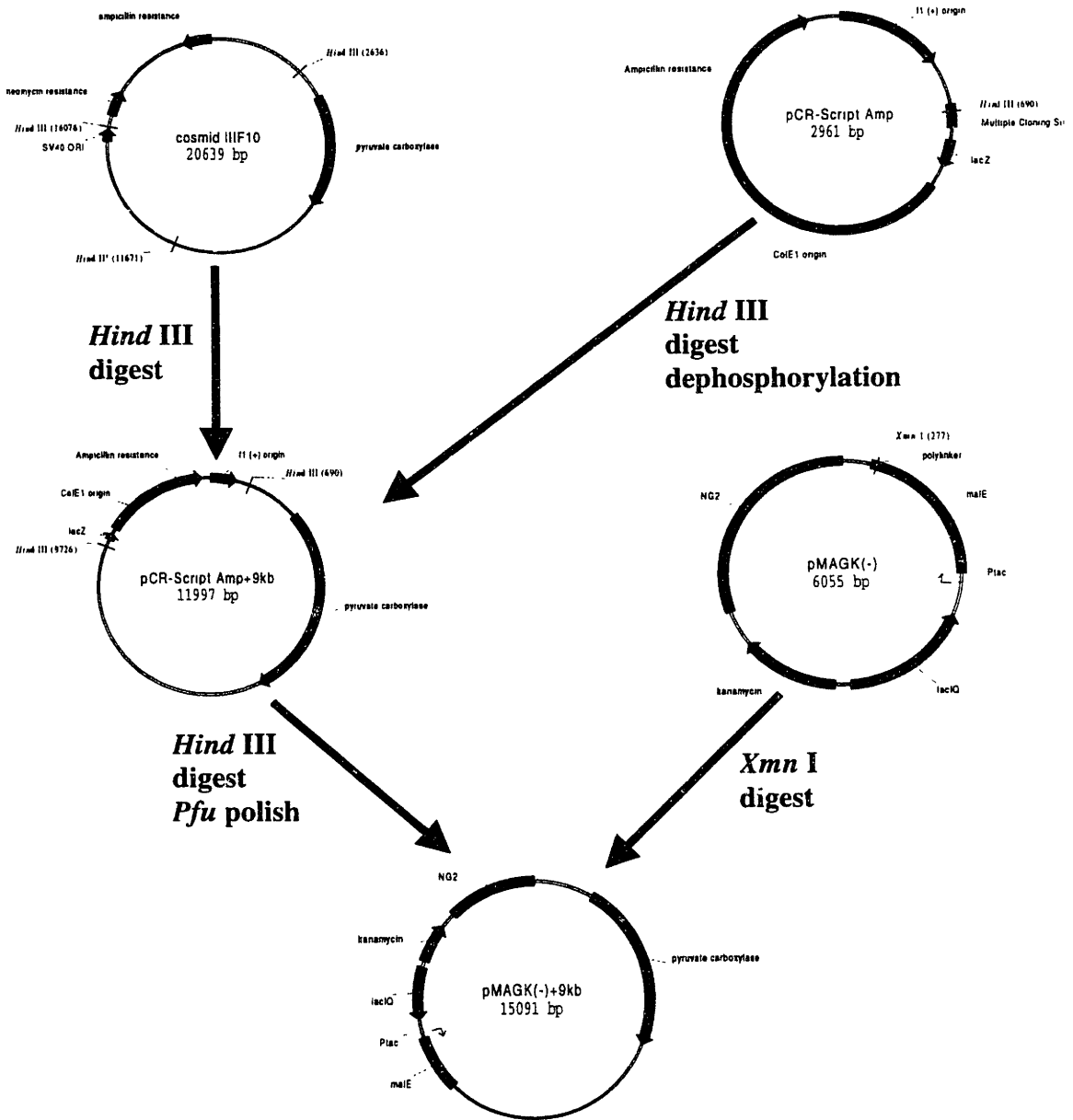
Figure 3.6 Construction of the *C. glutamicum-E. coli* vector pMAGK (-)+9 kb that carries the *C. glutamicum* pyruvate carboxylase gene as a 9 kb DNA fragment. Drawings are not to scale.
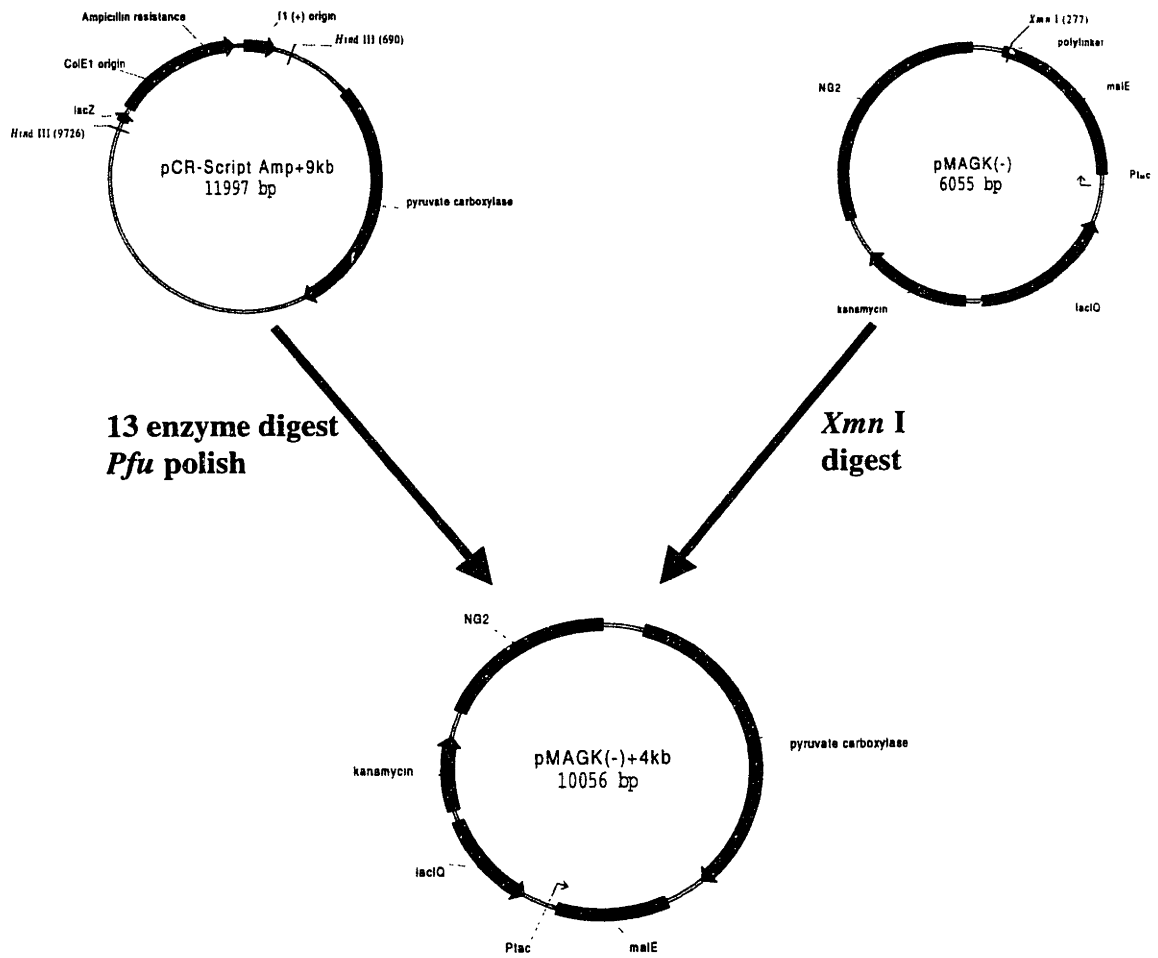
Figure 3.7 Construction of the *C. glutamicum-E. coli* shuttle vector pMAGK (-)+4 kb that carries the *C. glutamicum* pyruvate carboxylase gene as a 4 kb DNA fragment. Drawings are not to scale.
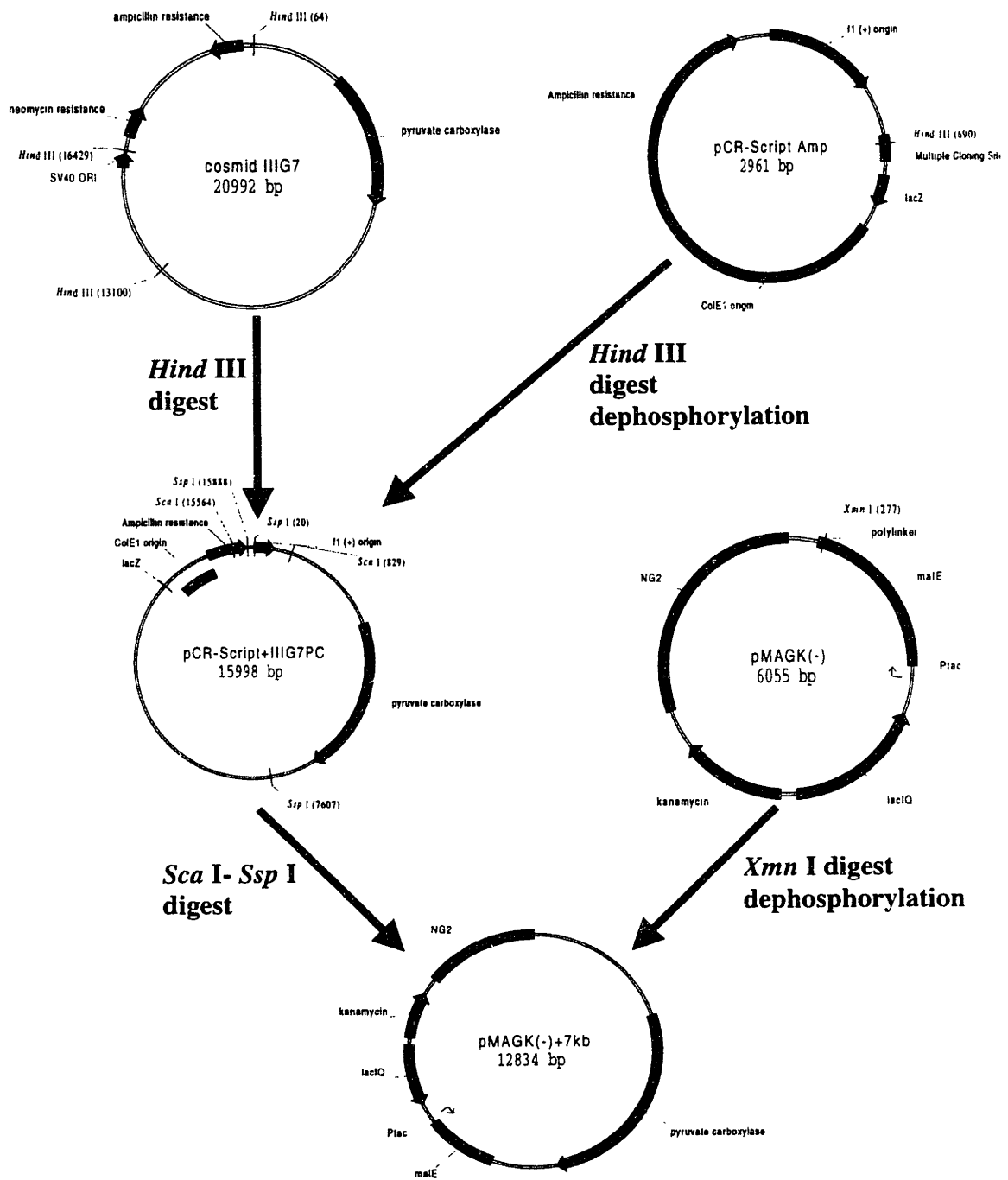
Figure 3.8 Construction of the *C. glutamicum-E.coli* shuttle vector pKD7 (pMAGK (-)+7 kb) that carries the pyruvate carboxylase gene as a 7 kb DNA fragment. Drawings are not to scale.
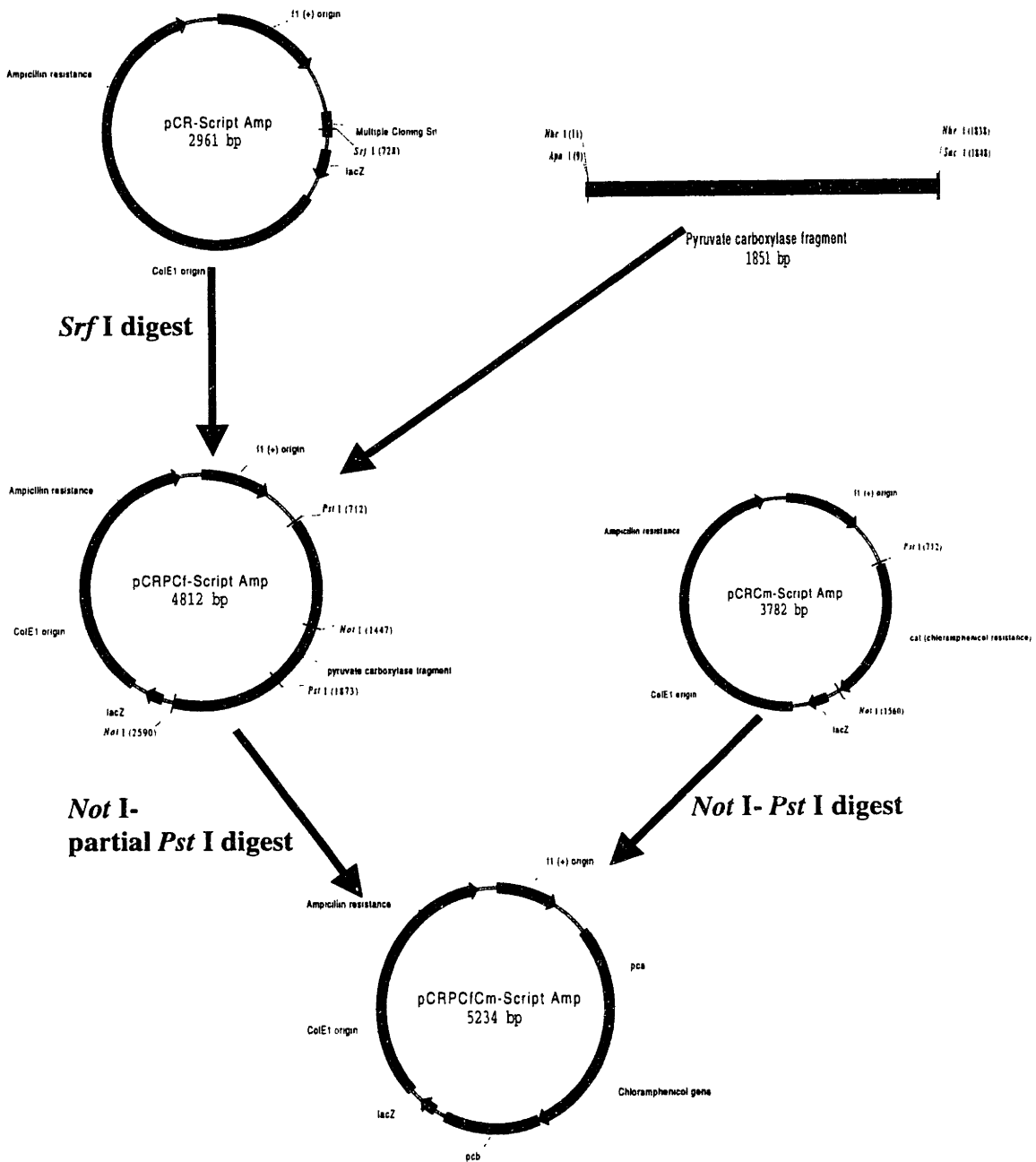
Figure 3.9 Construction of the *E. coli* vector pCRPCfCm-Script that contains the chloramphenicol resistance determinant adjacent to a fragment of the *C. glutamicum* pyruvate carboxylase gene. Drawings are not to scale.
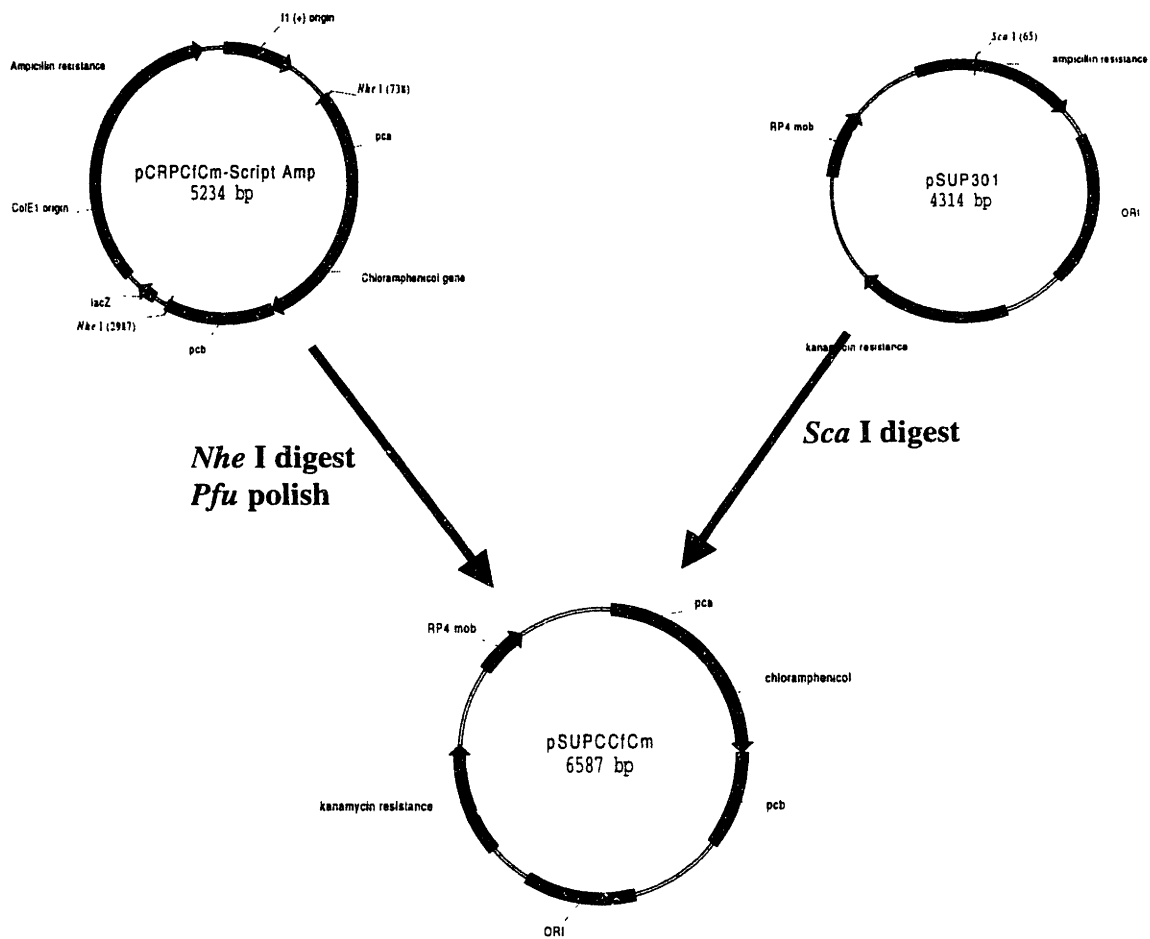
Figure 3.10 Construction of the *E. coli* mobilizable vector pSUPCfCm for marker exchange mutagenesis of pyruvate carboxylase. RP4 mob: mobilization fragment. Drawings are not to scale.
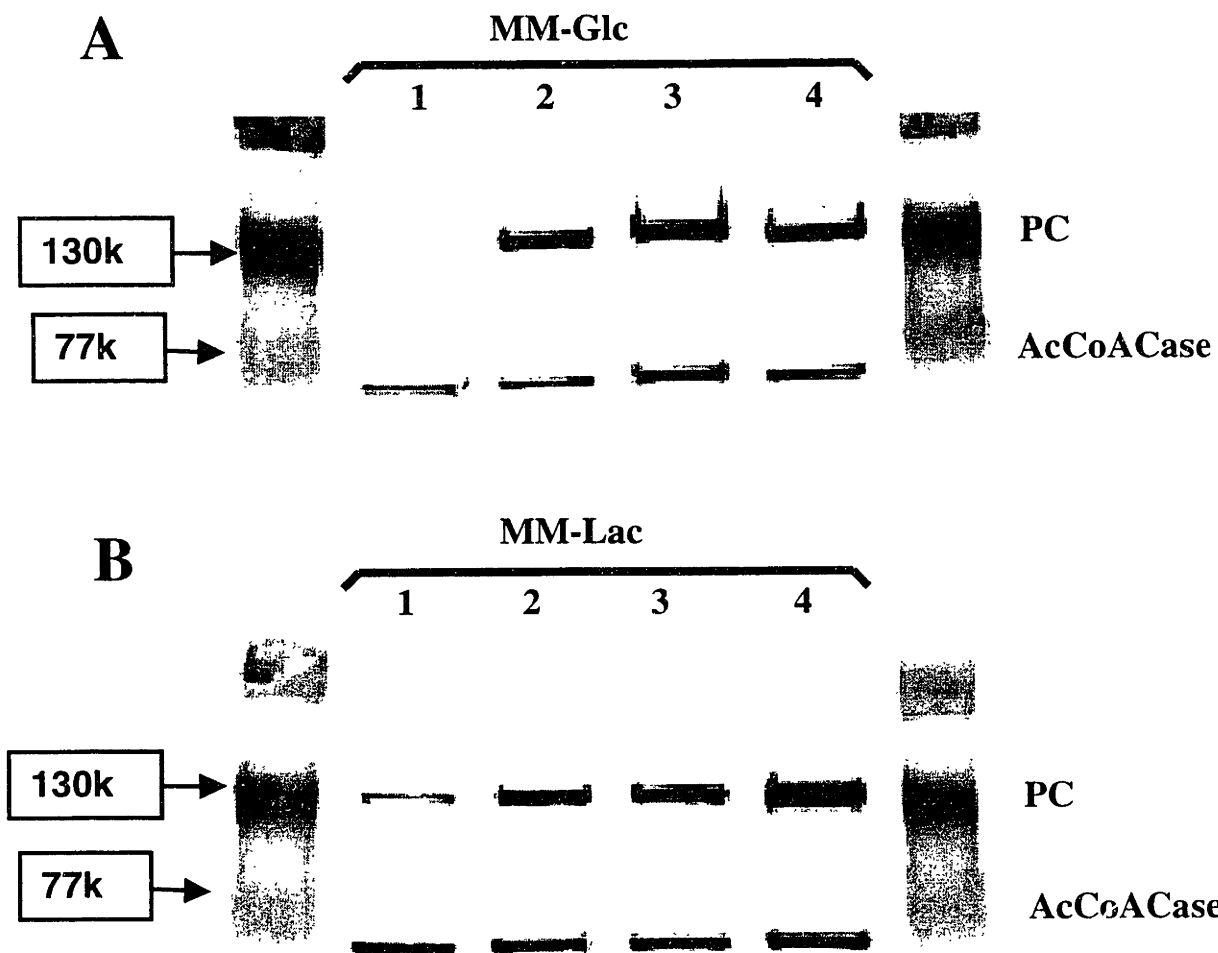
Figure 3.11 Wetern blot of pyruvate carboxylase synthesized in the control and recombinant 21799 strains harboring three different constructs of *pc* genes grown on (a) glucose and (b) lactate. Lane 1, control ; Lane 2, pMAGK (-) +4pc; Lane 3, pMAGK (-) +9pc; Lane 4, pKD7.
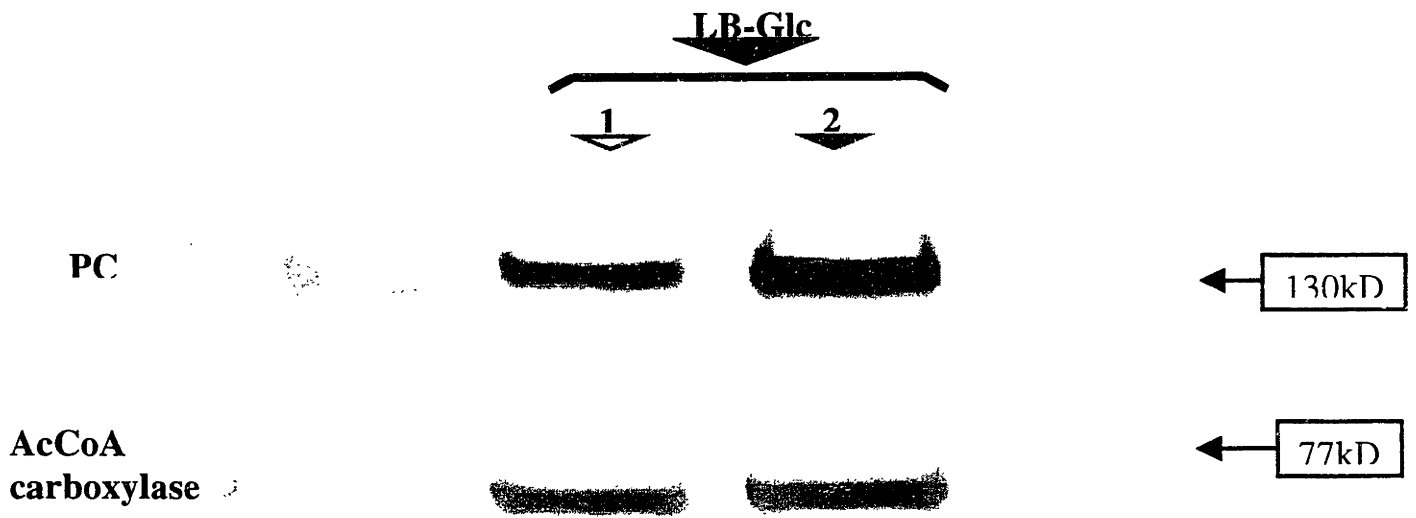
Figure 3.12 Western blot of pyruvate carboxylase synthesized in the control and recombinant 21253 strain harboring pKD7 grown on (a) glucose and (b) lactate. Lane 1, control; Lane 2, pKD7.
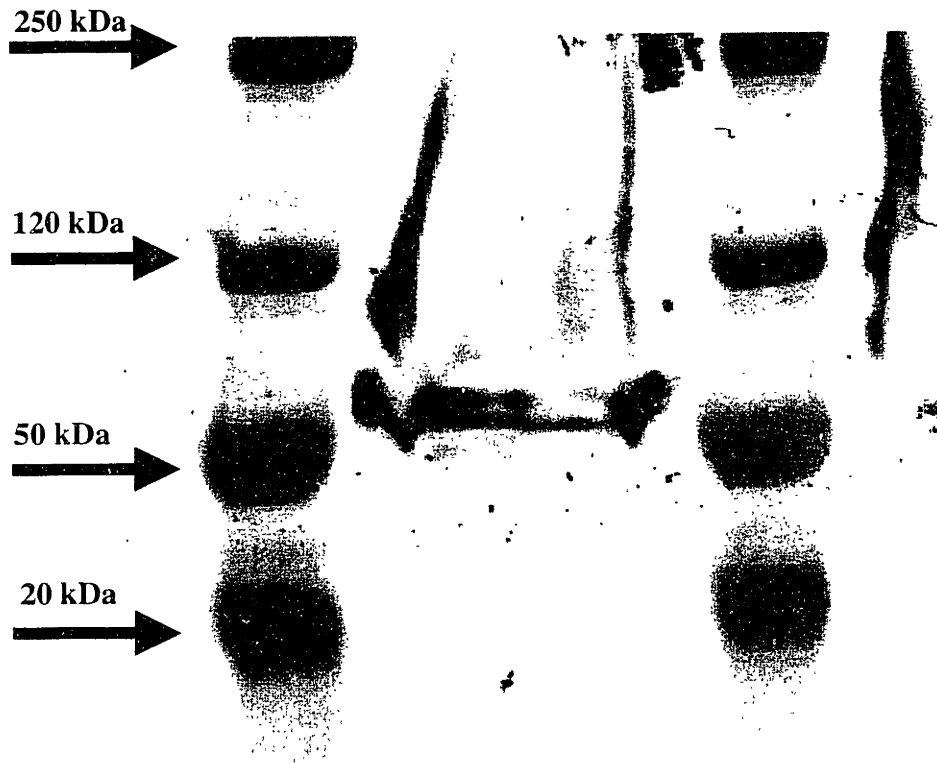
Figure 3.13 Western blot analysis for the *C. glutamicum*pyruvate carboxylase mutant strain. Protein crude extract was prepared and western blot was performed as described in the previous chapter from the mutant strain grown on glucose minimal medium. Lane 1&3, protein standards; lane 2, mutat strain.

112

Strain	Pyruvate carboxylase activity (U/mg DCW ± SD)	
	Glucose	Lactate
21253	58 ± 5	97 ± 7
253pKD7	58 ± 6	161 ± 10
21799	30 ± 4	36 ± 3
799pKD7	53 ± 7	89 ± 8

Table 3.1 Comparison of pyruvate carboxylase activities in *C. glutamicum* ATCC 21253 with those in *C. lactofermentum* ATCC 21799 on different carbon sources.

Strain	Glucose			Lactate		
	μ_m $(h^{-1})^a$	q_m (mg lysine/g cell·h)^b	X_f $(g/l)^c$	μ_m (h^{-1})	q_m (mg lysine/g cell·h)	$X_f(g/l)$
21253	0.35	46	4.2	0.34	63	2.5
253pKD7	0.34	46	4.2	0.41	23	4.0
21799	0.18	32	4.2	0.16	53	4.1
799pKD7	0.23	20	8.0	0.2	20	6.3

Table 3.2 Comparison of growth and lysine production of *C. glutamicum* ATCC 21253 with *C. lactofermentum* ATCC 21799
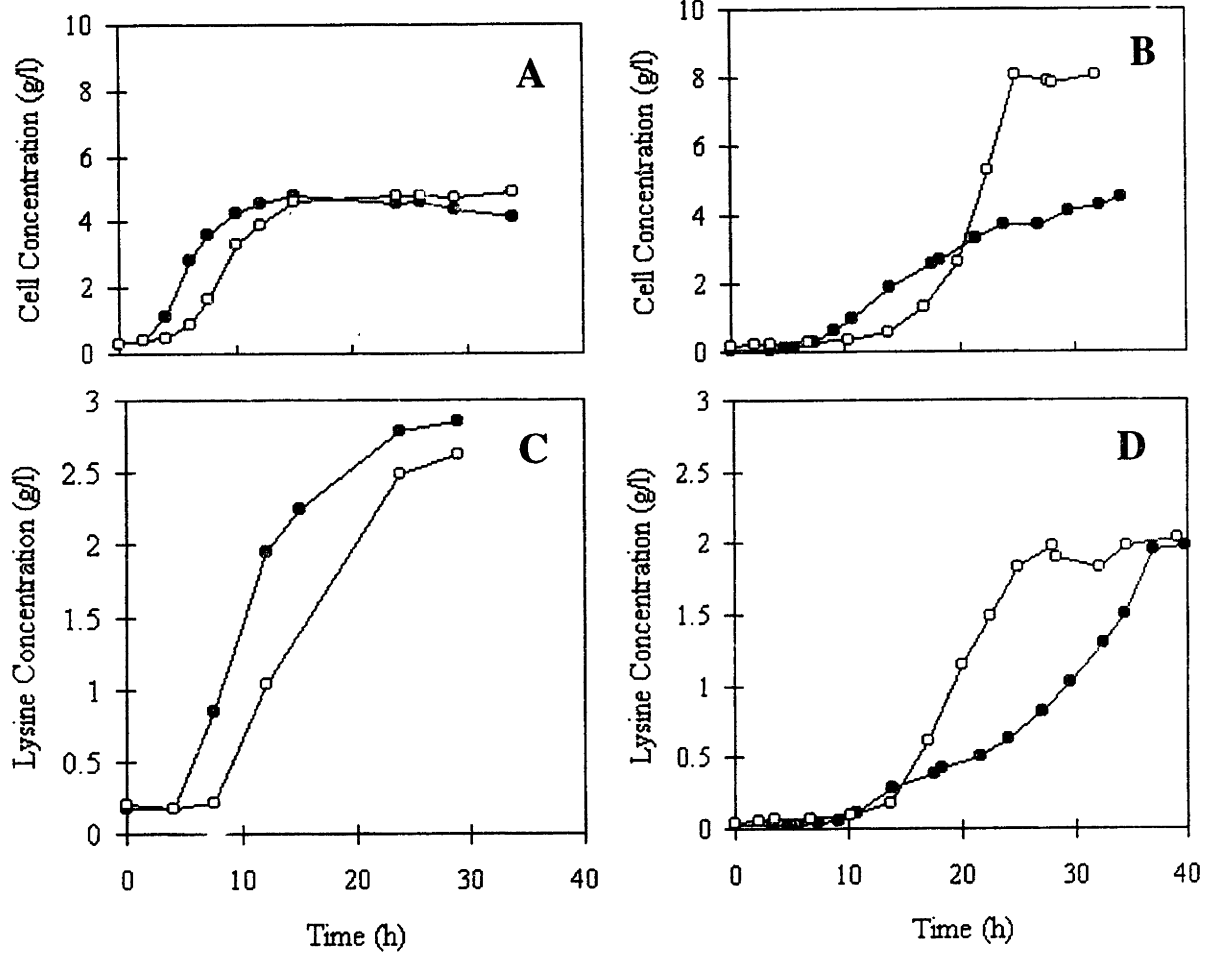
Figure 3.14 Growth and lysine production of (A,C) *C. glutamicum* ATCC 21253& respective recombinant strain and (B,D) *C. lactofermentum* ATCC 21799 and respective recombinant strain grown on glucose. Closed and open symbols represent the control and recombinant strains, respectively.
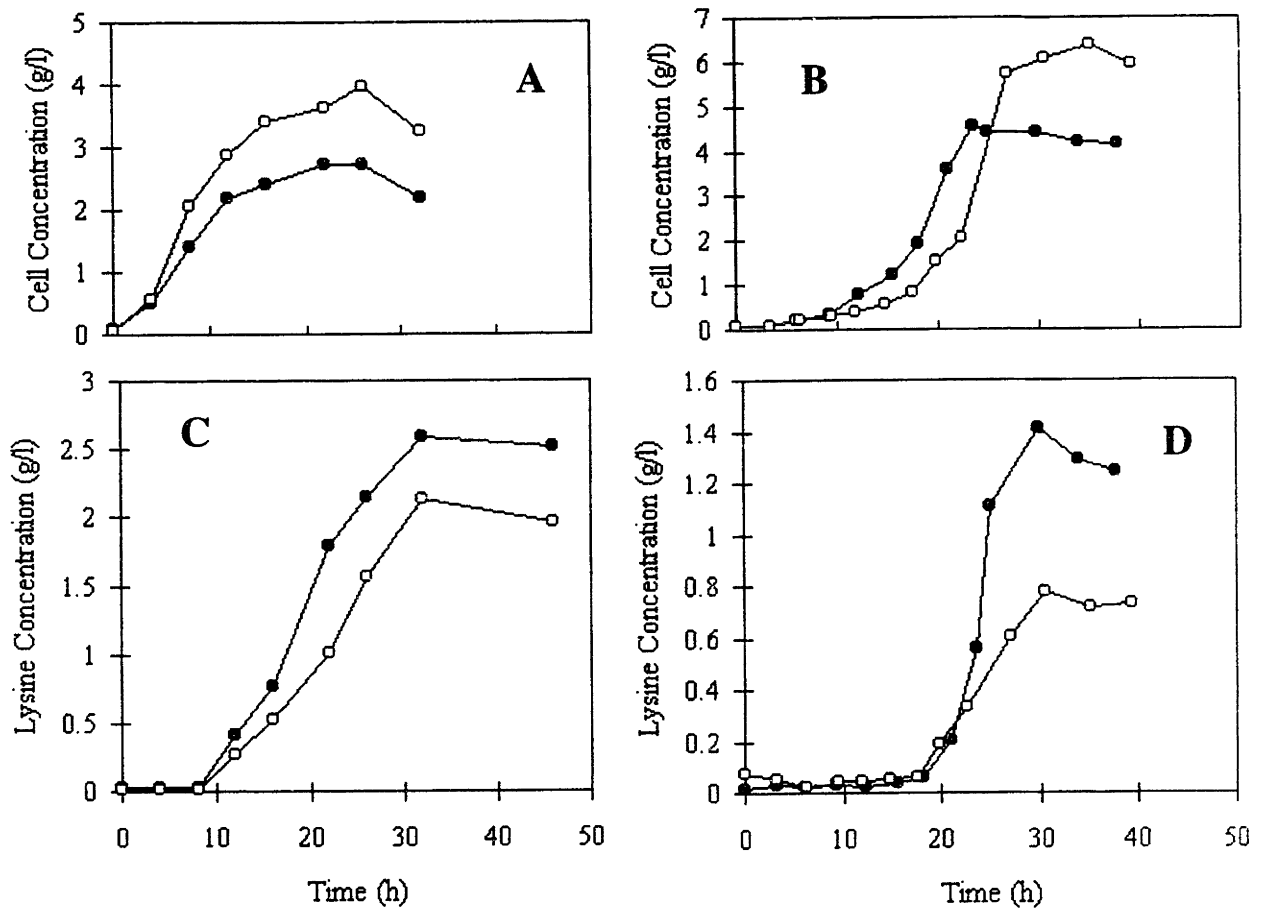
Figure 3.15 Growth and lysine production of (A,C) *C. glutamicum* ATCC 21253&
respective recombinant strain and (B,D) *C. lactofermentum* ATCC 21799 and respective
recombinant strain grown on lactate. Closed and open symbols represent the control and
recombinant strains, respectively.
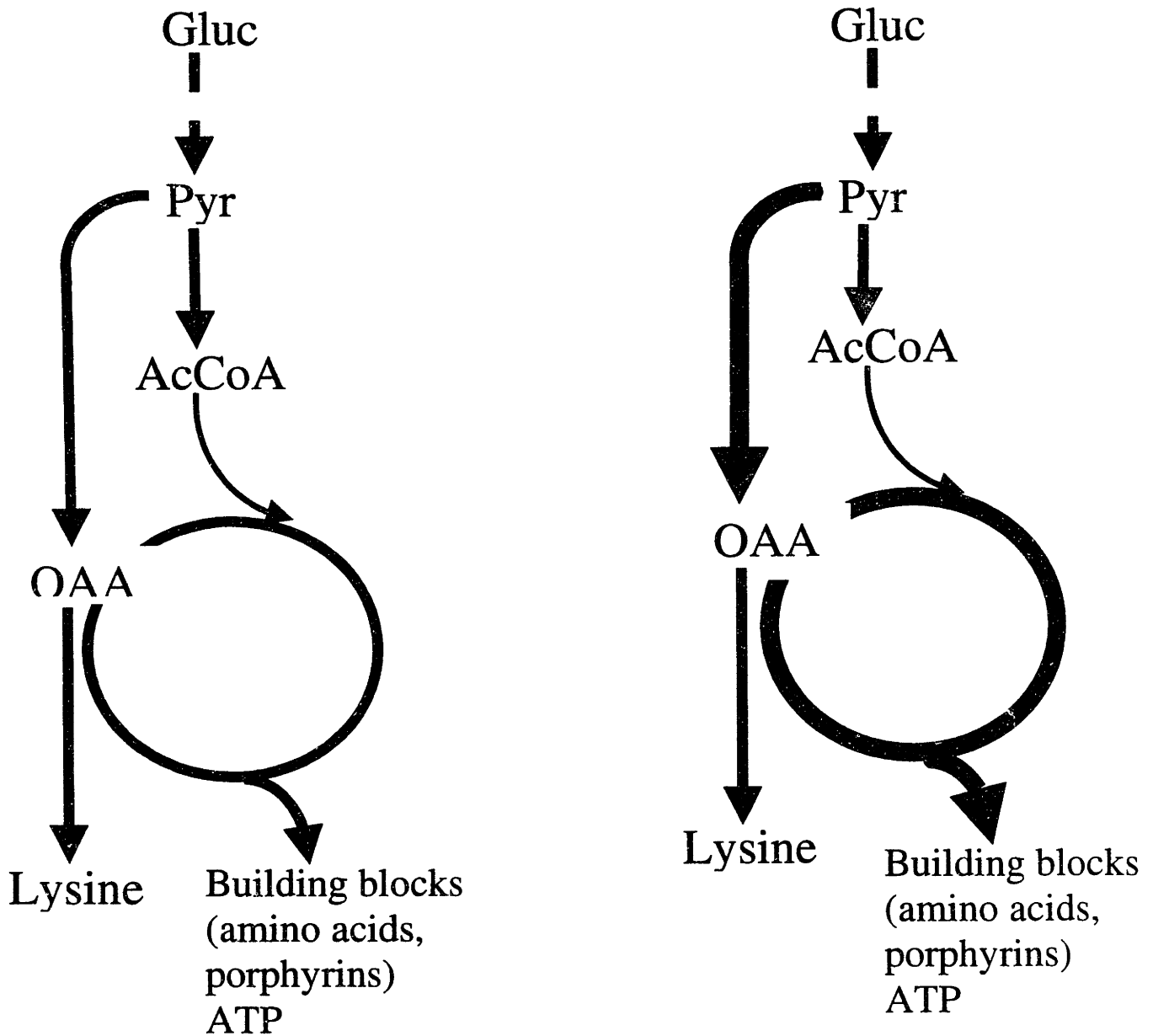
Figure 3.16 A metabolic model:the overexpression of pyruvate carboxylase increases carbon flux towards TCA cycle. Due to the presence of a bottleneck in the lysine biosynthesis pathway, that does not translate into more lysine. Instead carbon flux that exits the TCA cycle for other product formation (other amino acids and porphyrins) increases and that translates into better growth.

CHAPTER 4: LYSINE PRODUCTION IMPROVEMENT BY SIMULTANEOUS GENE OVEREXPRESSION

4. A. Introduction

4.A.1 Chapter Overview

As described in the previous chapter, the overexpression of pyruvate carboxylase removed a growth bottleneck but did not accomplish a redirection of carbon flux towards lysine production. This led to the hypothesis that the higher flux generated by pyruvate carboxylase cannot be matched by an equal flux in the subsequent stes in the lysine biosynthesis pathway. Previous work has led to the identification of three flux-controlling steps in this pathway. First control is exerted at the entry of aspartate into the lysine pathway (aspartate kinase), second at the branching point where aspartate semialdehyde is converted to either lysine or threonine (Cremer, Eggeling et al. 1991), and third control is exerted by the export of L-lysine from the cell (Schrumpf, Eggeling et al. 1992), (Broer, Eggeling et al. 1993).

In this chapter we present the experimental investigation of the first controlling step, namely aspartate kinase (aspartokinase) that performs the phosphorylation of aspartate. First we present the cloning and overexpression of two aspartokinase genes. The first gene encodes the wildtype enzyme and the second a deregulated, feedback non-inhibited enzyme. Batch culture results are then presented that show the physiological effect that the overexpression of this gene has on the physiology of *C. glutamicum*.

In the second part of this chapter the simultaneous overexpression of pyruvate carboxylase and aspartokinase is investigated, based on the assessment of the results that were obtained from all the previous single gene overexpressions. Again the physiological effect of such bicistronic overexpression is assessed based on growth and lysine production profiles. All the fermentation parameters are clearly presented in order to verify some of the speculations that have been developed based on those results on the

carbon flux distribution at the pyruvate node and the lysine biosynthesis pathway. Most importantly the generation of a new recombinant strain with higher lysine productivity is reported.

4.A.2 Background on Aspartate Kinase

Biosynthesis of lysine proceeds via the diaminopimelic acid (DAP) pathway (Jetten and Sinskey 1995), (Jetten, Follettie et al. 1995), (Tosaka, Takinami et al. 1978). Aspartokinase and aspartate-semialdehyde dehydrogenase catalyze the initial reactions of this pathway (Follettie, Peoples et al. 1993), (Kalinowski, Cremer et al. 1991), (Shiio and Miyajima 1969)

$$\text{Aspartate} + \text{ATP} \rightarrow \text{Aspartyl-}P + \text{ADP}$$

$$\text{Aspartyl-}P + \text{NADPH} \rightarrow \text{Aspartate semialdehyde} + \text{Pi} + \text{NADP}$$
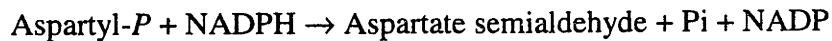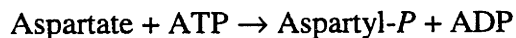
Aspartate kinase (aspartokinase) is the only enzyme involved in the lysine biosynthesis of *C. glutamicum* which is known to be controlled in its activity by feedback inhibition by the end products (lysine and threonine) (Nakayama, Tanaka et al. 1966), (Shiio and Miyajima 1969), (Tosaka, Takinami et al. 1978). The aspartokinase activity is severely inhibited in the presence of lysine together with threonine in low concentrations. In contrast, the *E. coli* aspartate kinase is inhibited by threonine alone and aspartate kinase III by lysine alone (Eikmanns, Eggeling et al. 1993), (Fazel, Guillou et al. 1983), (Cohen and Saint- Geront 1987). The *C. glutamicum* aspartate kinase also differs with respect to its insensitivity to diaminopimelate or methionine, which are known to inhibit some of the respective isoenzymes from *B. subtilis* and *E. coli* (Bondaryk and Paulus 1985).

Aspartate kinase is composed of two distinct subunits α and β that are overlapping in the genome of *C. glutamicum* and are immediately followed by the gene of aspartate semialdehyde dehydrogenase. Two promoters are required for the expression of

the genes of these two subunits. *ask*P1 and *ask*P2. Transcription from *ask*P1 initiates 35 and 38 bp upstream of the aspartokinase structural gene. The second promoter *ask*P2 lies within the aspartokinase coding region, upstream of the translation start site of the β subunit and can direct the expression of both this subunit as well as aspartate semialdehyde dehydrogenase (Fig 4.1) (Follettie, Peoples et al. 1993).

In the past, mutants have been generated with feedback-resistant aspartokinase (Tosaka, Takinami et al. 1978), (Sano and Shiio 1970), (Cremer, Eggeling et al. 1991). One of the mutants isolated, namely ATCC strain 21799, has a partially deregulated aspartokinase attributed to a mutation (G345D) in the β-subunit of the gene (Jetten, Follettie et al. 1995), (Follettie, Peoples et al. 1993).

Characterization of a defined aspartokinase mutant of *C. lactofermentum* obtained via marker exchange mutagenesis showed that the aspartokinase activity can be dispensable for lysine and DAP synthesis required for growth. This indicates either that threonine in some way can be converted to lysine, or that other isozymes of aspartokinase exist in lysine producing *Corynebacterium* strains (Jetten, Follettie et al. 1995).

C. lactofermentum strains with elevated levels of the feedback resistant aspartokinase activity produced more lysine, demonstrating that aspartokinase activity is still a limiting step in lysine production by this strain. A 6- fold increase of aspartokinase activity is detrimental for growth in production medium, and the specific productivity of the recombinant strain is lower (Schrumpf, Eggeling et al. 1992), (Thierbach, Kalinowski et al. 1990). Cells harboring a multicopy plasmid containing the aspartokinase gene did not produce more lysine than cells with two chromosomal copies of this gene. This indicates that introduction of an additional aspartokinase copy is sufficient to remove the aspartokinase as a limiting step in lysine production (Jetten, Follettie et al. 1995).

Simultaneous overexpression of aspartate kinase and dihydrodipicolinate synthase (*dapA*) convert the wildtype strain to a lysine overproducer (Eikmanns, Eggeling et al. 1993). Overexpression of the β subunit of aspartokinase also results in the excretion of more lysine. This is surprising since the total catalytic activity of the aspartate kinase in the respective recombinant strain is low, i.e. comparable to the wildtype. This result shows that in the case of aspartate kinase the overexpression itself is not important. A

specific activity according to only one copy of the aspartokinase β subunit gene results already in a severe flux alteration and thus in lysine excretion.

4.B Materials and Methods

4.B.1 Strains, plasmids and analytical techniques

We define *C. glutamicum* ATCC strain 21253 as our control strain. As feedback resistant strain *Corynebacterium lactofermentum* we define ATCC strain 21799. This is a *Corynebacterium* strain which is resistant to the lysine analogue S-(2-aminoethyl)-L-cysteine (AEC) and which has been generated by UV irradiation of wildtype strain.

The media, plasmids and the analytical techniques used for organic acid detection and amino acid quantification have all been listed extensively in the previous chapter.

The techniques used for plasmid isolation from both *Corynebacteria* and *E. coli* strains as well as chromosomal DNA isolation and electroporation techniques have also been described extensively in the previous chapter.

4.B.2 Cloning of the aspartokinase gene

Polymerase Chain Reaction was used in order to isolate the gene of aspartate kinase as a 1.3 kb DNA fragment. Flanking primers were selected based on the sequence published by Puhler *et al.* (1991), and were purchased from Gibco BRL (Life Technologies, Gaithersburg, MD). The GenBank accession number for the aspartokinase gene is L16848 (National Center for Biotechnology Information, NIH, Bethesda, MD). The upstream primer was 5'-TTTATACCCGGGGAGTTGAGCGGGTAACTG-3', and the downstream 5'-CCTGGTCGACCTGGCCGGTTGCACCAACAAC-3'. An artificial *Sma* I site was inserted at the 5'-terminal end of the beginning primer and an artificial *Sal* I site at the 5'-terminal end of the end primer in order to facilitate subsequent cloning steps. The PCR amplification was performed in a way that minimizes the possibility of error during the amplification process. *PfuTurbo* was the polymerase enzyme, the nucleotide concentration was kept at a low level and 25 cycles were performed in total.

Two separate PCR reactions were set, one using genomic DNA isolated from strain 21253 as a template (regulated apsrtokinase) and the other genomic DNA isolated from the strain 21799 (deregulated aspartokinase).

The 1.3 kb DNA fragment corresponding to the aspartokinase gene from both strains was visualized under UV light on a 0.8% agarose gel using electrophoresis. The genes were then inserted separately into the *Srf* I blund site of plasmid pCR-Script according to the ligation protocol outlined in the pCR-Script Amp Cloning kit (Stratagene Cloning Systems, La Jolla, CA). The resulting 4.2 kb plasmids were named pCR253ask and pCR799ask respectively, with the 21253 *ask* gene inserted in the first one and the 21799 *ask* gene inserted in the second. The transformation candidates were first isolated based on the color detection and later by restriction digests and PCR analysis.

The PCR products corresponding to the genes of the regulated (21253) and deregulated (21799) aspartokinase were then sequenced using six primers that are "walking" along the gene. The templates used for the sequence were plasmids pCR253ask and pCR799ask that contain these genes. From the sequences obtained it was shown that the PCR product corresponding to the aspartate kinase gene of strain 21253 was identical with the one reported in the literature. In the PCR product derived using chromosomal DNA of strain 21799 two mutations were detected, present in the β subunit of the aspartokinase gene.

Both aspartokinase genes were further removed from the *E. coli* vector pCR-Script by double digest with *Sma* I- *Sal* I and were inserted into the *C.glutamicum-E.coli* shuttle vector pMG108 that was also digested with the same enzymes. This cloning strategy introduced the aspartokinase genes immediately downstream of the *E. coli-C. glutamicum tac* promoter that plasmid pMG108 carries (Gubler and Sinskey 1993), (Colon, Jetten et al. 1995). The plasmids produced with this cloning strategy were assigned the names pMG253ask and pMG799ask. Figure 4.2 gives a detailed flow chart for the construction of plasmid pMG253ask. The same flow chart applies for plasmid pMG799ask.

For the introduction of the aspartokinase genes into vector pMAGK(-), plasmids pCR253ask and pCR799ask were digested with *Sma* I and *EcoR* V and the blunt ended

DNA fragments that contain the aspartokinase genes were inserted separately into the *Xmn* I site of vector pMAGK(-). The plasmids derived from this cloning procedure were assigned the names pMAGK253ask and pMAGK799ask (Figure 4.3).

4.B.3 Construction of pyruvate carboxylase- aspartokinase containing plasmid

For the simultaneous overexpression of pyruvate carboxylase and aspartate kinase, plasmid pKD7 containing the gene of pyruvate carboxylase was digested with restiction enzyme *Sac* I, which we found that cuts the construct at only one site. The product of this restriction digest was further polished with *Pfu* and dephosphorylated. Plasmid pMG253ask was digested with restriction enzymes *Kpn* I- *Sal* I. *Kpn* I is a restriction enzyme that cuts at least twice vector pMG108, one of the cuts being right before the beginning of the *tac* promoter region. The approximately 2 kb DNA fragment that contains the aspartokinase gene was polished by *Pfu* and inserted into the previously digested, polished and dephosphorylated plasmid pKD7. This cloning strategy yielded the new bicistronic construct, pKDask253 (Figure 4.4). The introduction of the aspartokinase gene was further verified by PCR analysis and restriction digests.

4.B.4 RNA isolation from *C. glutamicum*

The RNA isolation from *C. glutamicum* is performed using the Qiagen RNA isolation kit with a modification of the protocol provided by the company at the initial lysis step.

Cells are grown overnight in 3 ml cultures using glucose or lactate minimal medium (10^8-10^9 cells). They are harvested by centrifugation and resuspended in solution E (40 mM Tris; 2 mM EDTA, pH adjusted to 7.9 with acetic acid) and 250 µl 1 M sorbitol that contains lysozyme at concentration 3 mg/ml (usually the mixture is prepared by adding 7.73 µl of lysozyme solution 100 mg/ml in 250 µl sorbitol 1 M). After the addition of 10 µl mutanolysin the suspension is left to incubate at 37°C for 1 hr.

To the slurry 50 μl of 0.5 M EDTA and 100 μl of 10% SDS are added. The mixture is then boiled for 2 min and immediately returned to ice where it is left to incubate for 10 min. The cell debris is removed by centrifugation at 14,000 rpm for 40 min. To the supernatant an equal volume of 64% ethanol is added and thorough mixing is performed by repeated pipetting. After centrifugation at 15,000 rpm for 15 min the supernatant is discarded and the precipitated RNA is resuspended in 200 μl lysis/binding solution provided by the manufacturer. The resusspension occurs by vigorous vortexing. An equal volume of 64 % ethanol is applied next and vigorous mixing occurs by repeated pipetting.

The lysate/ethanol mixture is then applied to a binding filter provided by the manufacturer as follows: an RNAqueous filter cartridge is inserted into one of the Rnase-free collection tubes supplied by the manufacturer. The lysate/ethanol mixture is then applied onto the RNAqueous filter (maximum volume to be applied 700 μl) and centrifugation follows for 15 sec- 1min. The flow- through is discarded and the filter cartridge is washed with 700 μl wash solution 1 (provided by the manufacturer) again by applying the wash solution onto the filter and centrifuge for 15sec- 1 min. A second wash follows in a similar way by applying 500 μl of a second wash solution #2/3 (provided by the manufacturer). This wash is repeated one more time and after removing the flow-through the filter is microfuged for an additional 2 min in order to remove the last traces of wash solution.

The final step is the RNA elution from the filter. The filter cartridge is transferred to a fresh collection tube. 25-60 μl of elution solution is added onto the cartridge and incubation follows at 65-70 °C for 5-10 min. The eluate is recovered by spinning the tube for about 1 min. This final elution step is repeated one more time in order to assure the complete elution of the RNA from the cartridge. RNA is quantified by measuring absorbance at 260 nm.

4.B.5 Electrophoresis of RNA through gels containing formaldehyde

Before the electrophoresis 5× formaldehyde gel- running buffer is prepared as follows: 0.1 M MOPS solution is prepared by adding 20.6 gr of MOPS to 800 ml of DEPC- treated 50 mM sodium acetate. The pH of the solution is adjusted to 7.0 with 2 N of NaOH and 10 ml DEPC- treated 0.5 M EDTA (pH 8.0). So the final concentrations in the solution are 0.1 M MOPS, 40 mM sodium acetate and 5 mM EDTA.

For preparing the gel, the agarose solution in H_2O is melted as usual by microwaving and the solution is cooled to about 60 °C. To this solution 5× formaldehyde gel running buffer is added to a final concentration of 1× as well as pure formaldehyde to a final concentration of 2.2 M.

For running the gel, the RNA samples need first to be prepared as follows: up to 30 μg of RNA in H_2O (usually 4.5 μl) are mixed with 5× formaldehyde gel-running buffer (2.0 μl), formaldehyde (3.5 μl) and formamide (10.0 μl). To this mixture 2 μl of sterile, DEPC- treated formaldehyde gel-loading buffer is added (50% glycerol, 1 mM DTA pH 8.0, 0.25% bromophenol blue and 0.25% xylene cyanol FF). The gel is initially prerun for 5 min at 5 V/cm. The samples are then immediately loaded into the lanes of the gel and the running is performed by submerging the gel in 1× formaldehyde gel-running buffer at 3-4 V/cm until the desired resolution is obtained.

4.B.6 Northern Blot Hybridization

The procedure for Northern Blot hybridization is based on the protocols provided for Southern Blot hybridization from Roche Molecular Biochemicals (DIG- based procedure, Roche Molecular Biochemicals Inc.) and Schleicher & Schuell (Transblotter, S&S Inc.) with necessary adjustments.

Probe labeling: The probe used for northern blot hybridization is a DNA fragment corresponding to the gene of aspartate kinase and derived by restriction digest of plasmid

pCR253ask. After the digest, the probe is purified from agarose gel and is resuspended (10 ng to 1 µg) in 16 µl TE buffer. A denaturation procedure follows, during which the probe is boiled for 10 min and quickly chilled on ice for 5 min (prevent annealing). 4 µl of DIG- High Prime (label) is then added, mixed and spin briefly. An overnight incubation follows at 37°C, during which the labeling reaction takes place. The reaction is stopped by adding 2 µl of 0.2 M EDTA (pH 8.0). An equal volume (22 µl) of prechilled (-20°C) is added and the mixture is kept at -70°C for 1 hr. After a brief spinning at 14,000 rpm (15 min) the supernatant is carefully removed and the pellet is resuspended with 50 µl of 70% ethanol. Another spinning follows at 14,000 rpm for 15 min and after drying under vacuum the pellet is resuspended in 50 µl of TE with 0.1% SDS. The labeled DNA probe can then be stored at -20°C.

RNA transfer- Blotting: After the RNA gel electrophoresis has been performed, the gel is washed for 15 min in 100 ml transfer buffer (3 M NaCl, 8 mM NaOH and 2 mM sarkosyl) while at the same time the nitrocellulose membrane is soaked for 15 min in distilled H_2O. A "sandwich" between the gel and the membrane is then created as follows: 24 sheets of blotting paper are placed in the stack tray of the transblotter device (S&S Inc.) with the last one of them being prewet in transfer buffer. On top of this stack the nitrocellulose membrane is placed and then the RNA gel, making sure that there are no air bubbles between the gel and the membrane. On top of the gel another three sheets of blotting paper are placed presoaked in transfer buffer and the whole buffer tray is filled with 125 ml of transfer buffer. The tranfer is performed overnight.

After the transfer is complete the membrane is rinsed briefly in 2× SSC buffer (30 mM sodium citrate, 30mM NaCl, pH 7.0) and air dried over Whatman 3 MM paper. It is then wraped in Saran wrap and the RNA is fixed to the membrane by exposure to UV light for 2 min each side. The membrane is then placed is a plastic bag, sealed and left to incubate with 30 ml hybridization buffer (5× SSC, 0.1% N-lauroylsarcosine, 0.02% SDS, 1/10 volume 10X Blocking reagent provided by Roche Molecular Biochemicals) at 55°C for 2 hr.

For the hybridization, the DIG-labeled DNA probe is denatured by boiling for 5 min and then immediately cool on ice. In the mean time, 30 ml of hybridization buffer is

prewarmed at 55 °C and the probe is then added to it. The final mixture is then added to the bad containing the membrane and it is left to incubate at 55 °C overnight. Next day the hybridization solution is removed and the membrane is washed twice, 15 min each time, in 100 ml 2× SSC, 0.1% SDS at room temperature. Another two washes follow, 15 min each, in 100 ml of 0.1× SSC, 0.1 % SDS at 55 °C under constant agitation (0.1 × SSC is used in order to achieve higher strigency; the conenctration can actually be raised to 0.5 ×).

For the detection, the membrane is rinsed in 100 ml washing buffer (0.1 M maleic acid, 0.15 M NaCl, pH 7.5 with solid NaOH, 1.5% Tween 20, sterile) and then left to incubate with 100 ml blocking solution for 30 min. 2 μl of anti- DIG-AP conjugate are then added to 30 ml blocking solution (antibody solution) and the membrane is incubated with this mixture for another 30 min. Two washes follow for 15 min each with 100 ml washing buffer and the membrane is then left to equilibrate for 2-5 min in 20 ml AP-buffer (0.1 M Tris HCl, 0.1 M NaCl, 50 mM MgCl$_2$, pH 9.5 sterile). Twenty drops (about 0.5 ml) of CSPD are then placed on top of the membrane and it is then left to incubate for 10 min at 37 °C in order to enhance the luminescent reaction. Signal is finally obtained by exposing to X- ray B&W film (Kodak BIOMAX) at room temperature for more than two hours.

4.B.7 Aspartate kinase enzymatic assay

The enzymatic assay of aspartokinase was performed as previously described in the literature (Kalinowski, Bachmann et al. 1990) (Thierbach, Kalinowski et al. 1990). It is based on the reaction:

$$\text{L-Aspartate} + \text{ATP} \xrightarrow{\text{MgCl}_2} \text{L- Aspartylphosphate} + \text{ADP}$$
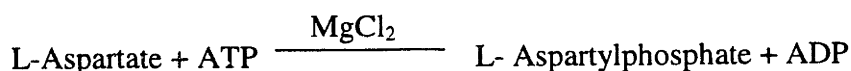
After the protein crude extract has been generated from C. glutamicum cells (wildtype or recombinant), one volume of crude extract is mixed with five volumes of saturated (NH$_4$)$_2$SO$_4$ solution. In this step, the protein is precipitated from the solution by

increasing the salt concentration followed by centrifugation. The pellet that contains the protein is then resuspended carefully (avoid generating bubbles) in 1 ml of assay buffer. A typical assay buffer contains the following: 100 mM Hepes (pH 7.8), 400 mM KCl, 12 mM $MgCl_2$, 500 mM hydroxylamine, 10 mM ATP and 15 mM L-aspartate with the pH adjusted to 7.6. The reaction is then left to proceed for 30 or 60 min at 30 °C and it is terminated by the addition of 0.75 ml of ferric chloride solution (10% $FeCl_3 \cdot 6$ H_2O, 3.3% TCA and 0.7 N HCl). The denatured protein debris is then removed by centrifugation for 5 min at maximum speed (15,000 rpm) and the absorbance of the supernantant left is measured at 540 nm.

The background activity is also estimated by measuring, by the same procedure described above, the extend of the reaction without the presence of aspartate, crude extract and ATP. The estimated background activity is then substracted from the activity estimated previously (complete assay).

The aspartate kinase specific activity is calculated with the following equation:

$$V = \frac{\Delta Abs/min * 10^9}{600 * mg\ of\ total\ protein} \cdots [=] \text{ nmol aspartyl - hydroxymate/min/mg total protein}$$

where V is the enzymatic velocity, $\Delta Abs/min$ is the linear slope of the curve describing the accumulation of aspartyl-hydroxymate at 540 nm over time, and the molar extinction coefficient for aspartyl-hydroxymate is 600 (Jetten, Follettie et al. 1995).

The total amount of protein in the crude extract is measured using the BCA Protein Assay Kit provided by Pierce (Pierce, Rockford IL.) following standard instructions provided by the company.

4.C Results

4.C.1 *C. lactofermentum* aspartokinase sequence

An extensive analysis of the aspartokinase expression in *C. glutamicum* 13032 and in *C. flavum* N13 has shown that this gene encodes two subunits from the same reading frame but with distinct promoters (Figure 4.1). It has also been reported that the reduced feedback inhibition by lysine and threonine of 21799 aspartokinase is due to mutations introduced in the open reading frame of the 18 kDa β subunit of this enzyme (Kalinowski, Cremer et al. 1991) (Follettie, Peoples et al. 1993) (Jetten, Follettie et al. 1995). The sequence of the β subunit of the partially resistant aspartokinase from *C. lactofermentum* and more specifically from strain 21799 has been published in the past and has also been deposited in GenBank with accession number L27125.

Since the aspartokinase genes for the current work have been derived by PCR reaction, it was imperative to obtain the sequences of the DNA fragments that were corresponding to the desired genes in order to verify the absence of any mutations introduced during PCR. Those sequences were compared to the wildtype *C. flavum* N13 aspartokinase gene whose sequence is also available in GenBank (accession number L16848) by the BLAST search program available at the NCBI web site.

The DNA fragment that was corresponding to the 21253 aspartokinase gene showed that no mutations were introduced during the PCR reaction since the sequence was identical to the of the *C. flavum* N13 gene. The most interesting results were obtained from the sequence of the PCR fragment corresponding to the 21799 aspartokinase gene. First, the sequence of the complete gene (both α and β subunits) was obtained. In the α subunit no mutations were detected, in accordance to what has been reported in the literature. However, in the sequence of the β subunit, only two mutations were detected. At nucleotide 834 from the start codon a change from G (N13) to A results in an amino acid change from alanine to threonine. At nucleotide 1033 from the start codon, a change from A to G results in an amino acid change from aspartate to glycine. No other mutations were detected in comparison with the N13 aspartokinase gene. The mutations that were detected with this sequence are different from the ones reported in the literature (Jetten, Follettie et al. 1995) and both of them are non silent i.e. they affect the amino acid encoded by the new nucleotide sequence.

4.C.2 Aspartokinase Northern Blot analysis

Using a DNA probe corresponding to the gene of aspartate kinase derived from *C. glutamicum* the transcription levels of aspartokinase were detected in the wildtype, monocistronic (*ask* single overexpression) and bicistronic (*pc* and *ask* overexpression) strains. Two different growth conditions were selected in both of which glucose was utilized as a carbon source (glucose minimal medium) but in one case IPTG was included in the medium (20 μmole) while in the second case no IPTG induction was introduced. The reason of using IPTG as an inducer has to do with the presence of a *tac* promoter upstream of the gene of aspartokinase when that gene was initially introduced into vector pMG108 and later into pMAGK(-). Even though we were anticipating no effect of IPTG on the transcription level of aspartokinase, since both the α and the β subunit need to be under the control of *tac*, it was of interest to investigate whether this was indeed the case.

During the northern blot analysis, a single signal was detected in all three strains. This signal was almost negligible in the control strain, amplified in the monocistronic strain and even stronger in the monocistronic strain when IPTG was added during the culture. This indicates that the RNA signal detected was indeed under IPTG induction (upper band, Figure 4.5). The signal was also present when no IPTG was present (even though both pMG108 and pMAGK(-) include the lacIQ repressor), showing that the aspartokinase gene was still under native promoter induction. So apparently the northern blot data indicate that at least the α subunit of aspartokinase is under the combined control of both the native (P1) and the *tac* promoter. Since no other band was detectable, we can make no conclusions about the transcription levels of the β subunit even though as mentioned earlier, we would expect no IPTG induction. Whether the whole gene (α and β subunits) was under the combined control of native promoters (P1 and P2) as well as the *tac* promoter was left to be concluded from the enzymatic assay experimental data.

4.C.3 Enzymatic activities

The enzymatic assays performed with both the monocistronic and bicistronic strains for aspartokinase concluded the work previously done with the PCR fragment sequences and the Northern Blot analysis. The two major questions to be answered had to deal with the deregulation of aspartokinase derived from strain 21799 and whether there was a some control excercised from the *tac* promoter on the transcription levels of the genes.

The first result that was evident from the beginning was that there was no IPTG effect on the enzymatic activities as it can be seen in figure 4.6. In the monocistronic aspartokinase strain the activity of aspartokinase was 9.2 U/min/mg of protein when IPTG was not present compared to 10.3 U/min/mg of protein when the inducer was added. In the bicistronic aspartokianse-pyruvate carboxylase recombinant strain, the aspartokianse activities were 6.2 U/min/mg of protein and 6.5 U/min/mg of protein with and without inducer. The presence of IPTG did not affect significantly the aspartokinase activity of either the monocistronic or the bicistronic strain even at inducer levels of the order of 20 μmole. Clearly the β subunit is under the control of only its native (P2) promoter, as it was expected. The band that was detected by the northern blot analysis corresponded most likely to the α subunit which of course can be under the control of both the native (P1) and the *tac* promoters. In the case of the monocistronic strain the aspartokinase activity was about 5- fold higher compared to the control strain as it has also been reported previously in the literature (Follettie, Peoples et al. 1993), (Kalinowski, Cremer et al. 1991), (Kalinowski, Bachmann et al. 1990), (Thierbach, Kalinowski et al. 1990).

The second interesting result from the assays is that indeed the aspartokinase gene cloned both in the monocistronic (containing plasmid pMAGK253ask) and the bicistronic (containing plasmid pKD253ask) constructs codes for a regulated enzyme as expected based on the obtained DNA sequence. In the presence of 25 mM lysine and 25 mM threonine in the reaction mixture the enzymatic activity dropped by 60% in the case of the control strain, almost to zero in the case of the monocistronic strain and by more than 70% in the case of the bicistronic strain. These results are consistent with data obtained previously and are reported in the literature (Jetten, Follettie et al. 1995) (Kalinowski,

Cremer et al. 1991). However, when aspartokinase activity was measured in the recombinant 21799 strain crrying plasmid pMG799ask, no deregulation was detected from the combined presence of lysine and threonine. It is also interesting to note the higher aspartokinase specific activity of the monocistronic strain versus the bicistronic by about 40%. This is most likely due to expression attenuation in the simultaneous overexpression of the pyruvate carboxylase and aspartokinase genes in the case of the bicistronic strain.

Pyruvate carboxylase enzymatic activity in the bicistronic strain was measured to be lower compared to the activity in the monocistronic (containing plasmid pKD7) strain. This comes as a verification to the point made previously. Pyruvate carboxylase activity in cells grown with lactate minimal medium was 1.4 times higher compared to the control (wildtype) strain and it was about 1.2 times lower comparing to the monocistronic strain. As in the case of the monocistronic strain, cells grown on glucose showed the same specific activity (if not less) with control cells. Apparently the overexpression of pyruvate carboxylase is much smaller when glucose is the carbon source and this small activity increase is not detectable with the permeabilized cell assay method.

4.C.4 Effect of simultaneous *ask-pc* overexpression and carbon source on growth

As already mentioned in the previous chapter, carbon source has a major effect on the expression level of pyruvate carboxylase and as such it affects directly the phenotype of the recombinant strain that overexpresses this enzyme. More specifically, it was shown in the previous chapter that when lactate was used as sole carbon source in minimal medium, the recombinant strain derived from the *C. glutamicum* strain 21253 was able to grow to a higher final cell concentration compared to the control strains. On the other hand when glucose was used as carbon source, no effect on the cell physiology was observed. The increase in final growth was also observed in the case of the *B. lactofermentum* strain 21799 when either lactate or glucose was used as sole carbon source, even though to a lower extend in the latter case. So there is a direct positive correlation between the level of pyruvate carboxylase activity and cell growth. On the

other hand, in the case of aspartokinase the only carbon source that has been studied extensively in literature is glucose and it was shown that the overexpression of this enzyme either has no effect on cell growth or it is actually influencing it in a negative way (Jetten, Follettie et al. 1995), (Schrumpf, Eggeling et al. 1992), (Thierbach, Kalinowski et al. 1990).

For the present work, the growth characteristics of recombinant strain 21253 harboring plasmid pKD7253ask were studied using lactate and glucose as two distinct carbon sources.

In the case of glucose as carbon source, the recombinant strain demonstrated a longer lag phase compared to the control, but the growth rates of both strains were calculated to be similar and about 0.52 h^{-1}. In terms of final cell concentration, both strains reached similar values, i.e. the control reached 5.79 g/l while the recombinant strain reached a slightly higher value of about 6.6 g/l. The growth profiles of the control and bicistronic recombinant strain can be seen in the upper panel of figure 4.7.We therefore conclude from these data that the simultaneous overexpression of pyruvate carboxylase and aspartokinase had no effect on the growth characteristics, with glucose as carbon source. In fact these data resemble the results that were obtained from the single pyruvate carboxylase overexpession and match with the data reported in the literature for the single aspartokinase overexpression. So no major flux redistributions most likely take place under those conditions.

In contrast, however, to monocistronic pyruvate carboxylase overexpression, similar were the physiological data when lactate was used as the sole carbon source. Both the recombinant and the control strains showed an identical lag phase and an almost identical growth rate of 0.35 h^{-1}. Also the final cell concentration for both strains was similar i.e. 2.6 g/l. So as in the case of glucose, the bicistronic recombinant strain exhibited a similar phenotype with the control strain. These growth profiles can be seen in the upper panel of figure 4.8. Compared to the recombinant monocistronic pyruvate carboxylase strain, the specific growth rate decreased 1.2 times while the final cell concentration by 1.5 times. It is evident from these data that an increase in the expression

level of pyruvate carboxylase has no counterbalancing effects on the possible negative effect that the overexpression of aspartokinase has on growth.

4.C.5 Effect of simultaneous *ask-pc* overexpression on lysine production

In the previous chapter it was determined that the overexpression of pyruvate carboxylase, contrary to our expectations, does not increase carbon flux towards lysine production, but rather improves the growth of the recombinant strain. This is hypothesized to be due to an increase of the carbon flux into the TCA cycle. This carbon cannot exit the TCA cycle because of the presence of a rate-limiting step and because of that no improvement in lysine production is pbserved.

In our effort to investigate possible bottlenecks in the lysine biosynthesis pathway, asprtokinase became the next focus of interest, mainly because of literature evidence that suggest a major role for that enzyme in lysine biosynthesis (Eikmanns, Eggeling et al. 1993). Even though elevated levels of aspartokinase produce more lysine, it has been shown that a 6-fold overexpression of this enzyme is detrimental for growth in production medium (glucose minimal medium), and the specific productivity of the recombinant strain is lower (Jetten, Follettie et al. 1995). It is evident therefore that the physiological effect of aspartokinase overexpression is exactly the opposite from that of pyruvate carboxylase and the simultaneous overexpression of both these enzymes becomes an interesting target for research study.

The simultaneous overexpression of these enzymes was performed in ATCC strain 21253 (*C. glutamicum*) using the native aspartokinase gene of this strain. As mentioned earlier, despite the presence of the *tac* promoter, upstream of the aspartokinase gene, no effect of IPTG was detected on enzymatic activity of aspartokinase.

The results obtained when glucose was used as the sole carbon source can be summarized as follows: the volumetric productivity of the recombinant (bicistronic) strain was about 0.05 g lysine/l/h which is about 1.5 times higher than that of the control of 0.03 g lysine/l/h. In terms of specific productivities, the recombinant strain reached a value of about 7.9 mg lysine/g cell/h which is about 1.5 times higher than that of the

wildtype strain (5.5 mg lysine/g cell/h). Finally in terms of final lysine concentration in the fermentation medium, the recombinant strain achieved an approximately 1.1 g lysine/lt again about 1.2 times higher than that of the wildtype strain of 0.68 g lysine/l. The lysine production and threonine consumption profiles with glucose as carbon source are presented in the lower panel of figure 4.7. It is obvious from these data that the improvement in terms of lysine production when glucose is used as a carbon source is significant.

The results were also similar when lactate was utilized as the unique carbon source during the fermentation experiments. The volumetric productivity of the bicistronic recombinant strain was 0.24 g lysine/l/h, which was 2.6 times higher than that of the wildtype strain (0.09 g lysine/l/h) and 3.8 times higher than that of the monocistronic pyruvate carboxylase recombinant strain (0.06 g lysine/l/h). In terms of specific productivity, the bicistronic recombinant strain achieved a value of 106.6 mg lysine/g cell/h, 2.5 times higher than that of the wildtype strain of 42.3 mg lysine/g cell/h and 4.8 times higher than that of the recombinant monocistronic pyruvate carboxylase overexpressing strain (22.1 mg lysine/g cell/h). As for the final lysine concentration, the bicistronic recombinant strain reach a value of 3.27 g lysine/lt, while the wildtype strain ended at 2.58 g lysine/l and the monocistronic pyruvate carboxylase recombinant strain at 2.14 g lysine/l. The lysine production and threonine consumption profiles with lactate as carbon source are presented in the lower panel of figure 4.8. Again a very significant increase in terms of lysine productivity can be observed with these data by the simultaneous overexpression of pyruvate carboxylase and aspartokinase.

A summary of all the fermentation parameters that have been calculated for the bicistronic and control strains is given in tables 4.1 and 4.2.

4.D Discussion

In the past, extensive analysis of the aspartokinase expression in *C. flavum* N13 and *C. glutamicum* 13032 demonstrated that the aspartokinase gene encoded two subunits from the same open reading frame. Mutations in the open-reading frame of the 18 kDa β

subunit were shown to result in aspartokinase activity with reduced sensitivity to the combined inhibition effect of lysine and threonine. One such inhibition in the *C. glutamicum* ATCC 13032 aspartokinase (a change of C to A at nucleotide 1118of the recorded DNA sequence results in the S301Y mutation) which confers a loss of sensitivity to feedback inhibition by threonine and lysine and a concominant increase in expression of the downstream aspartate semialdehyde dehydrogenase activity has been reported (Kalinowski, Cremer et al. 1991). Comparison with *E. coli* promoter sequences suggested that the latter effect reflects conversion from a weak to a strong promoter. In a separate study, the β subunit of the deregulated aspartokinase gene of *C. lactofermentum* ATCC strain 21799 was sequenced and studied for the presence of mutations. Between the β aspartokinase sequences of strains 21799 and 13032 a total of 15 base pair differences were observed. Only three of these mutations, T to C at nucleotide 351, C to G at nucleotide 361 and G to A at nucleotide 437, were shown to alter the primary structure of the β subunit, resulting in S118A, R121G and G146D mutations, respectively. The 21799 and N 13 β aspartokinase sequences showed four nucleotide differences. Only the T to C change at nucleotide 351 resulted in the R104G substitution. The N13 and 13032 sequences showed 11 nucleotide differences, resulting in two amino aicd residue changes at S118A and G146D (Jetten, Follettie et al. 1995).

In the work presented in this chapter, we have obtained the complete sequence of a PCR fragment corresponding to the aspartate kinase gene of *C. lactofermentum* ATCC strain 21799. As correctly predicted in the literature, no mutations were detected in the part of the gene encoding for the α subunit. In the β subunit however, the mutations detected were completely different from the ones expected. Comparing to the *C. flavum* N13 strain, two nucleotides were detected to be different: one at position 834 and one at position 1033. Both these changes affect the amino acid encoded by the corresponding codon. None of these mutations is similar to the ones reported in the literature (Jetten, Follettie et al. 1995). It is therefore possible that these two mutations were introduced during the PCR reaction. However they do not seem to be affecting the enzymatic activity of the encoded aspartokinase, which, in a similar way to the *C. flavum* enzyme, is also feedback inhibited. For that reason, for the creation of the bicistronic vector that

contained both the aspartokinase and the pyruvate carboxylase gene, the gene derived from *C. glutamicum* ATCC 21253 strain was utilized.

As already mentioned, elevated levels of aspartokinase activity result in higher lysine production, demonstrating that this enzyme is a limiting step. A six-fold overexpression of this enzyme is detrimental for growth and the specific productivity of the plasmid containing strain is lower (Jetten, Follettie et al. 1995).. In order to explain this type of physiological behavior, the following metabolic model is proposed: an increase in aspartokinase activity, one of the three controlling steps in lysine biosynthesis, results in an increase in carbon flux towards lysine production. This translates in higher lysine production, as observed experimentally. However, this increase results in a decrease of the carbon flux that enters the TCA cycle, and as a result of that less amino acids and porphyrins are produced and smaller amount of ATP is generated. This is exactly the reason why an attenuation in growth is observed in these recombinant strains which also results in a decrease in specific productivity. This proposed metabolic model is outlined in Figure 4.9.

As it can be seen from the above discussion, the effects of overexpression of aspartokinase and of pyruvate carboxylase are exactly the opposite. In the first case, as presented in the previous chapter, pyruvate carboxylase results in less lysine but better growth. This was attributed to a possible increase in carbon flux that enters the TCA cycle and which cannot eventually translate into more lysine because of the presence of a rate limiting step in the lysine pathway. So the question that naturally arises is: is it possible to counterbalance the negative effect on growth that aspartokinase overexpression causes by increasing the anaplerotic activity? In other words, could we expect an improved cell physiology in terms of growth and lysine production if both pyruvate carboxylase and aspartokinase are overexpressed simultaneously?

From the data presented in Tables 4.1 and 4.2 we can conclude the following: the simlultaneous overexpression of pyruvate carboxylase and aspartokinase leads to the creation of recombinant strain with growth characteristics similar to those of the control strain. Specific growth rates are almost identical while final cell concentrations are very close to each other, especially when lactate is used as a sole carbon source in the minimal

medium. On the other hand, both final lysine production as well as lysine productivity increase to significant levels in the case of the recombinant strain. This is especially true when lactate is used as the sole carbon source and in which case, as shown in the previous chapter, the highest pyruvate carboxylase expression occurs.

These shake flask data come as a conclusion to the speculations made earlier in regards to the effect of pyruvate carboxylase and aspartokinase on carbon flux distributions. It is obvious that the two enzymes have counterbalancing effects that eventually lead to a strain optimization. In order however to further prove the validity of the metabolic models proposed, further experimental work needs to be accomplished, especially by making intracellular and extracellular measurements of critical metabolites, such as aspartate and glutamate. The fact though remains, that this is the first report in the literature where improvement in primary metabolite synthesis, such as lysine, is accomplished through means of metabolic engineering. Further studies need also to be directed towards characterization of recombinant strains with simultaneous overexpression of pyruvate carboxylase and other potential bottlenecks in the lysine biosynthesis, such as diaminopimelate synthase and the lysine export system.
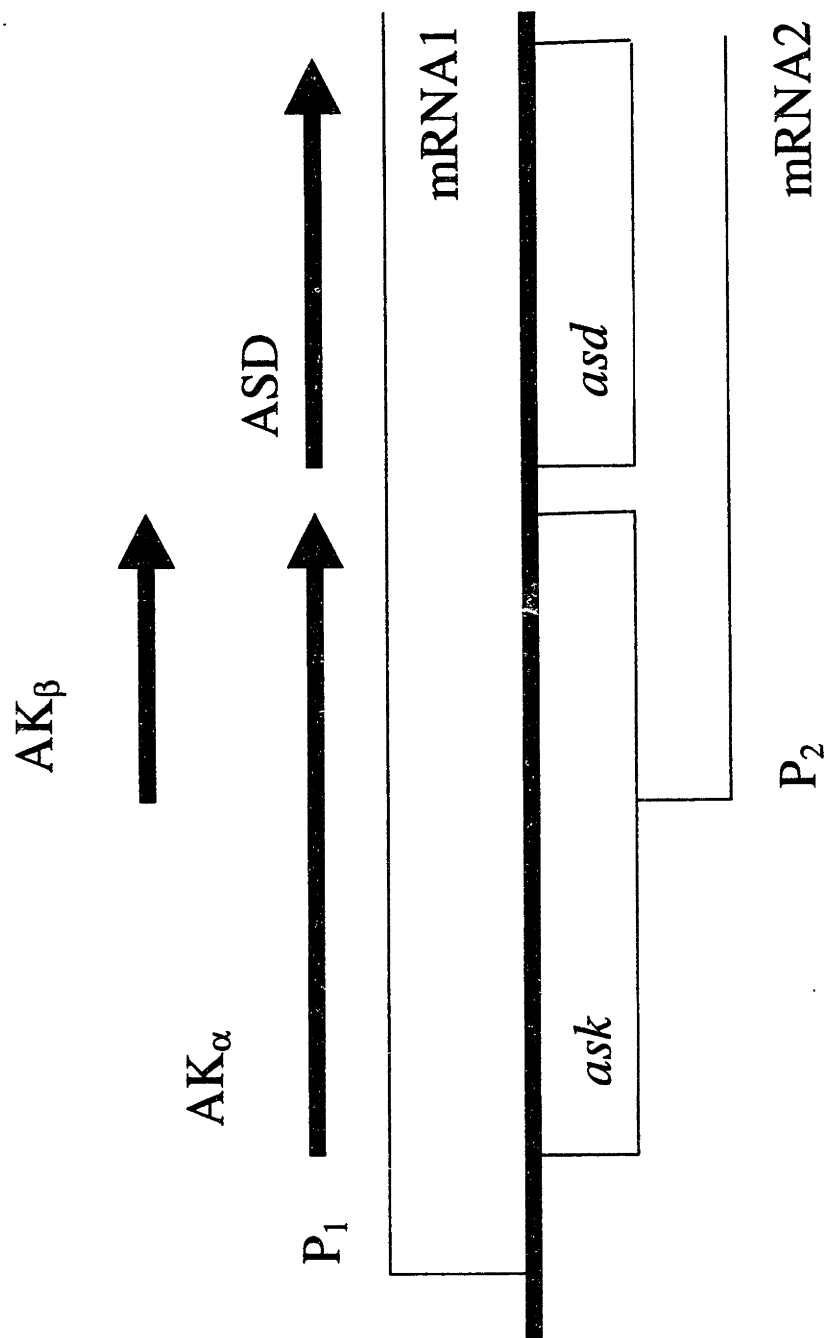
Figure 4.1 Model of the *C. flavum* N13 *ask-asd* operon, showing the structural genes and mRNA transcripts initiated at promoters *ask*P1 and *ask*P2.
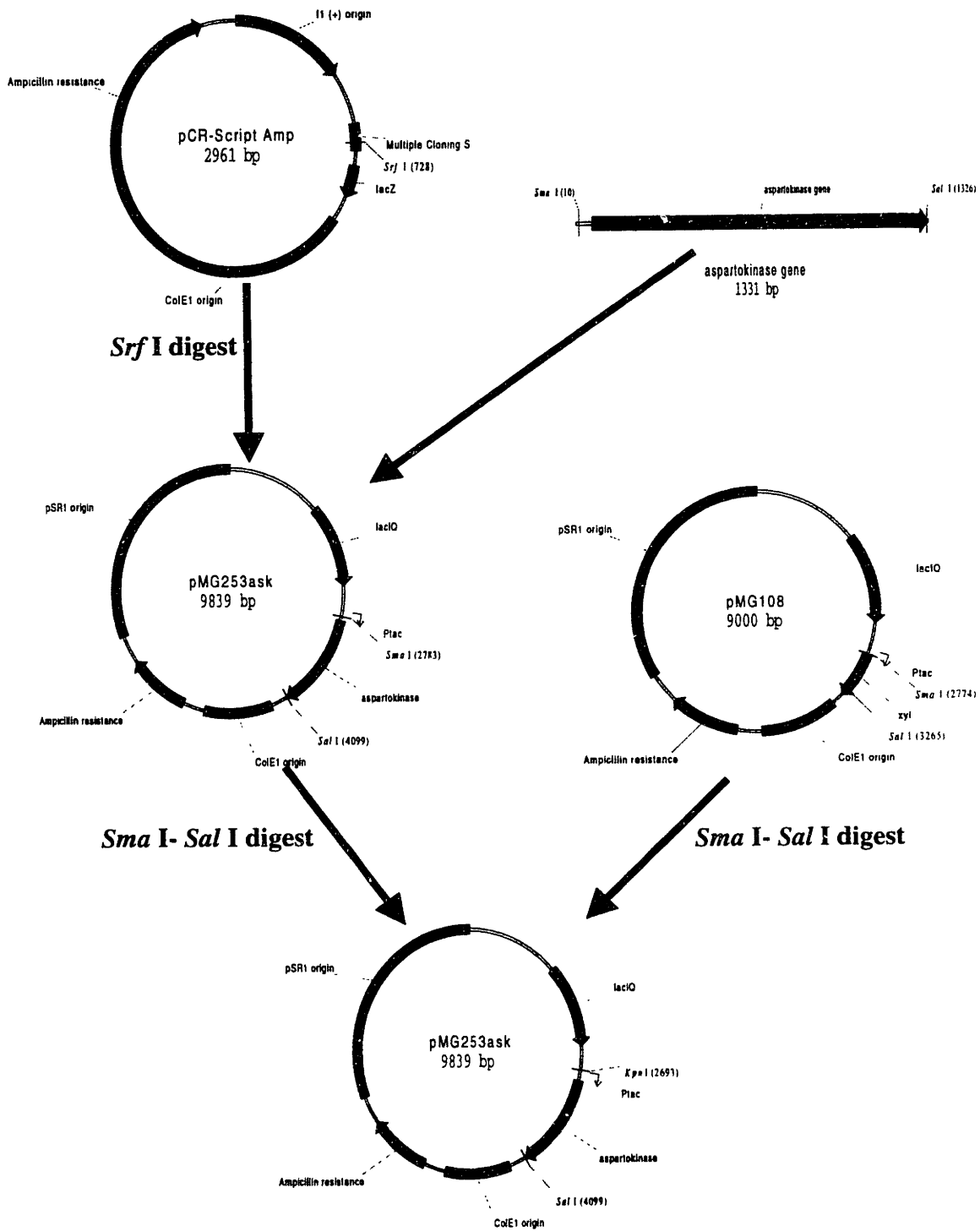
Figure 4.2 Construction of the *E. coli- C. glutamicum* vector pMG253ask that carries the *C. glutamicum* aspartokinase gene next to the *tac* promoter. Drawings are not to scale.
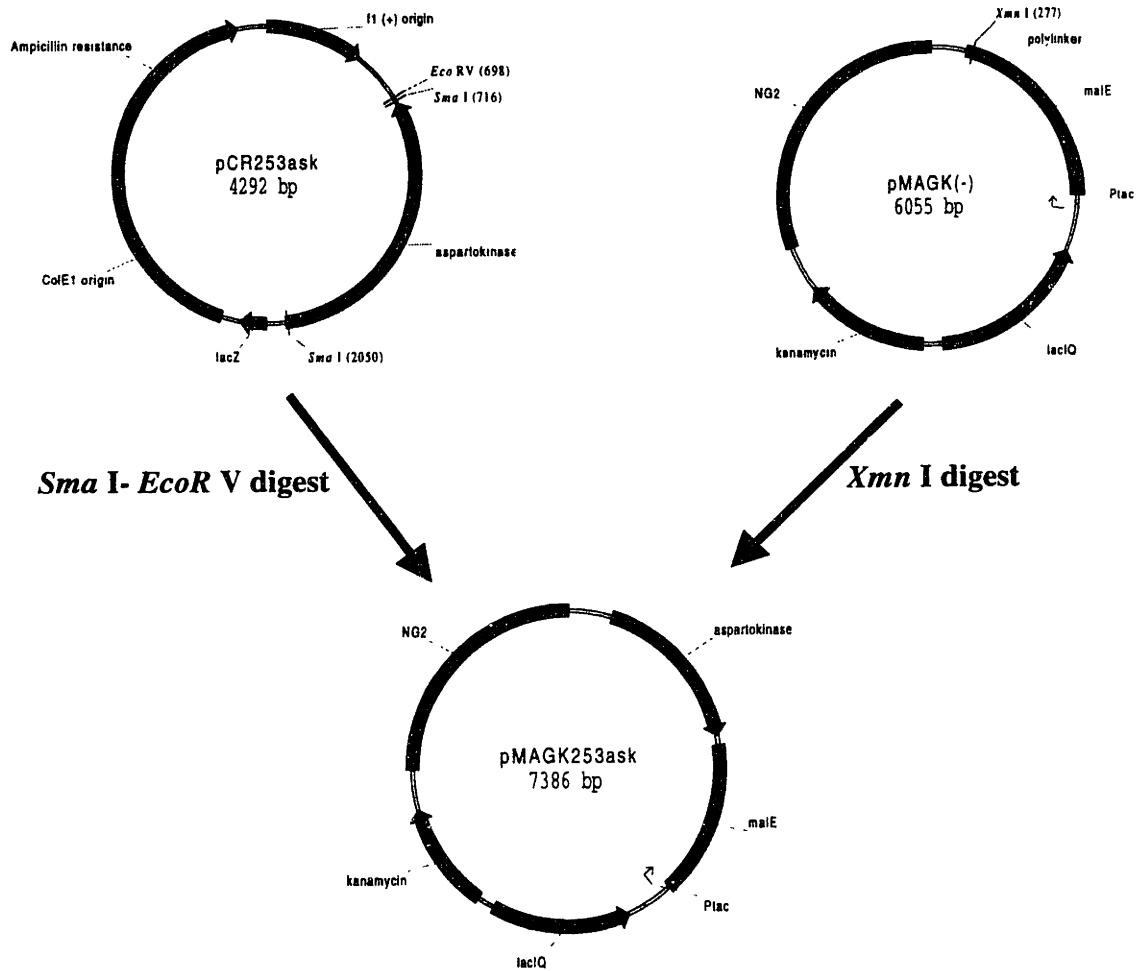
Figure 4.3 Construction of the *E. coli*- *C. glutamicum* vector pMAGK253ask that carries the *C. glutamicum* aspartokinase gene next to the *tac* promoter. Drawings are not to scale.
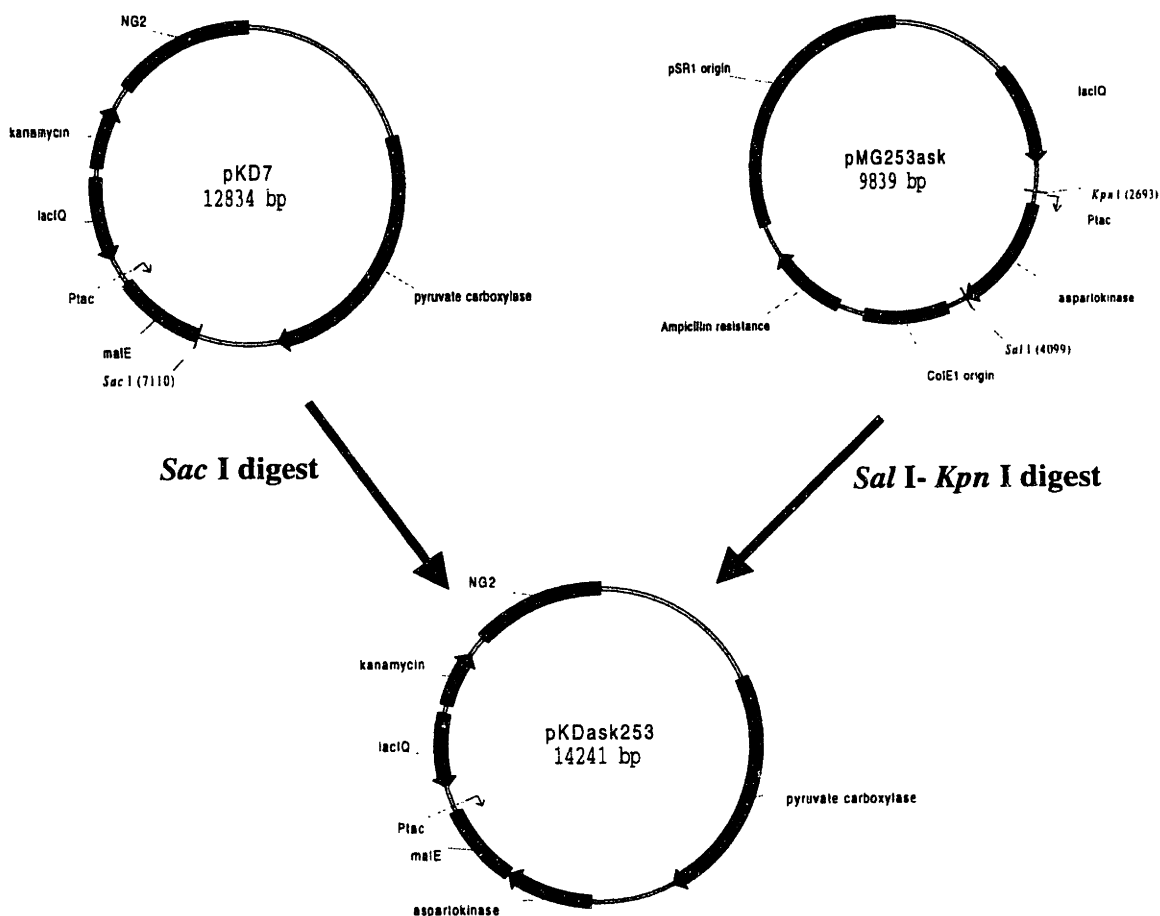
Figure 4.4 Construction of the *E. coli- C. glutamicum* vector pMG253ask that carries the *C. glutamicum* aspartokinase gene next to the *tac* promoter as well as the pyruvate carboxylase gene of the same organism. Drawings are not to scale.
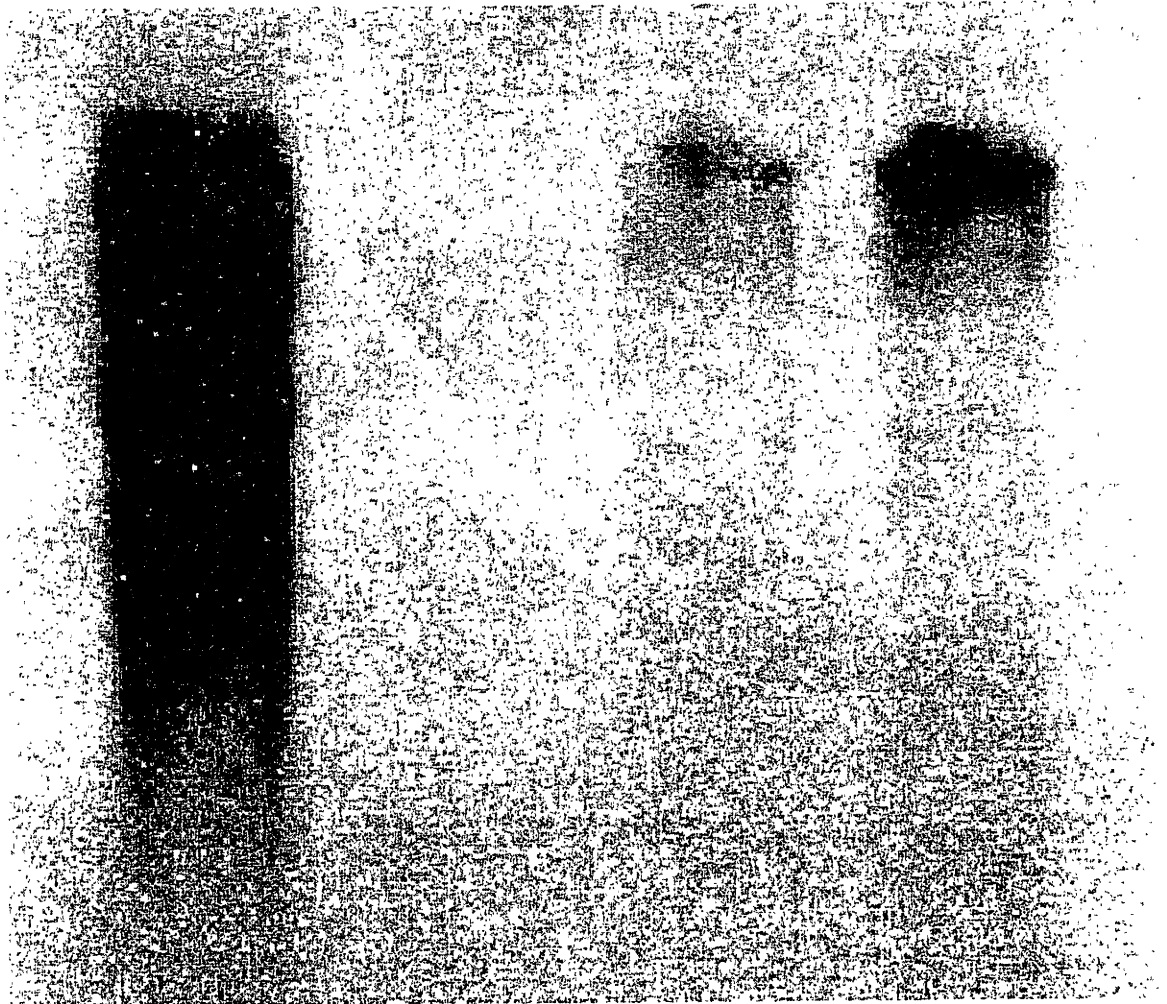
Figure 4.5 Northern blot results on aspartokinase transcription levels. Lane 1: control-PCR fragment containing the 21253 aspartokinase gene. Lane 2: control- total RNA isolated from strain 21253 carrying plasmid pMG108. Lane 3: RNA isolation from strain 21253 carrying plasmid pMG253ask, with no IPTG present in the culture medium. Lane 4: RNA isolation from strain 21253 carrying plasmid pMG253ask, with IPTG present in the culture medium. 20 μg of total RNA/DNA was loaded on each lane.
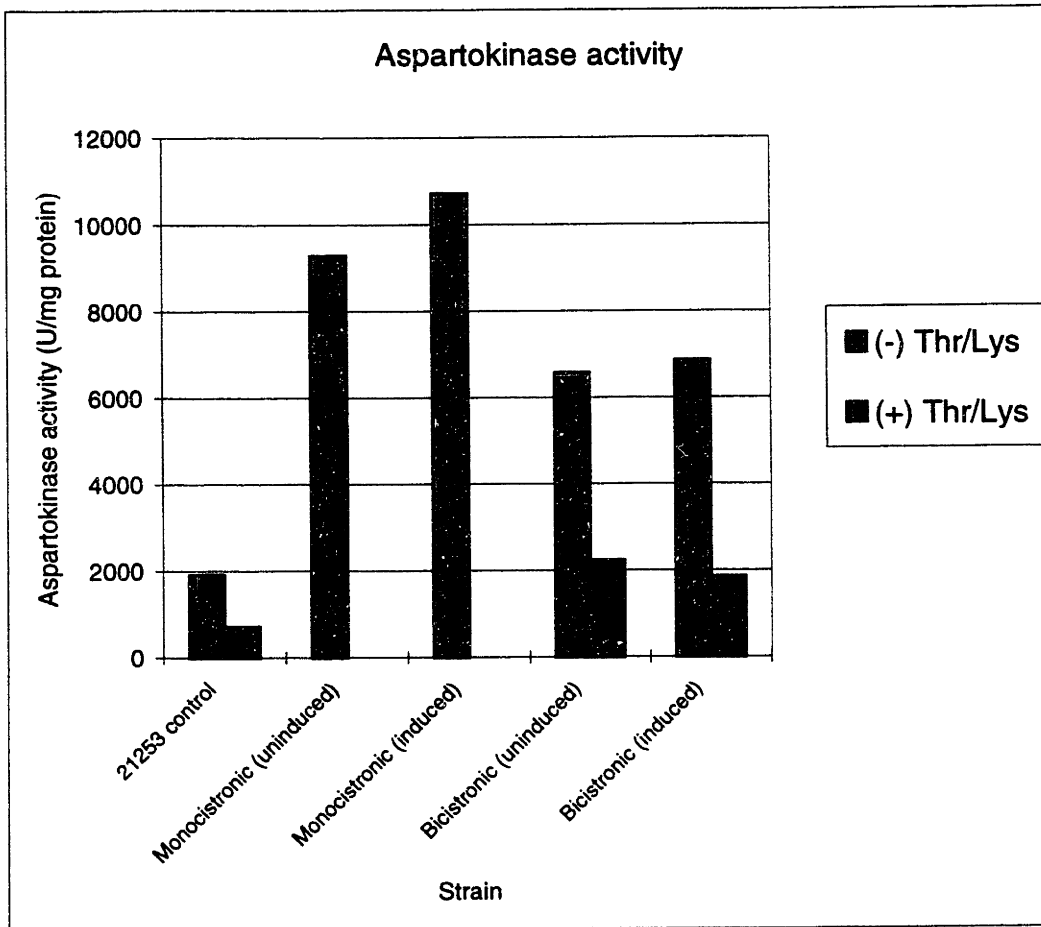
Figure 4.6 Enzymatic activities of various control and recombinant 21253 strains. Enzymatic assays were performed in the absense or presence of L-lysine and threonine (10 mM each). Cells were grown either in the absense or presence of 10 μM of IPTG inducer.
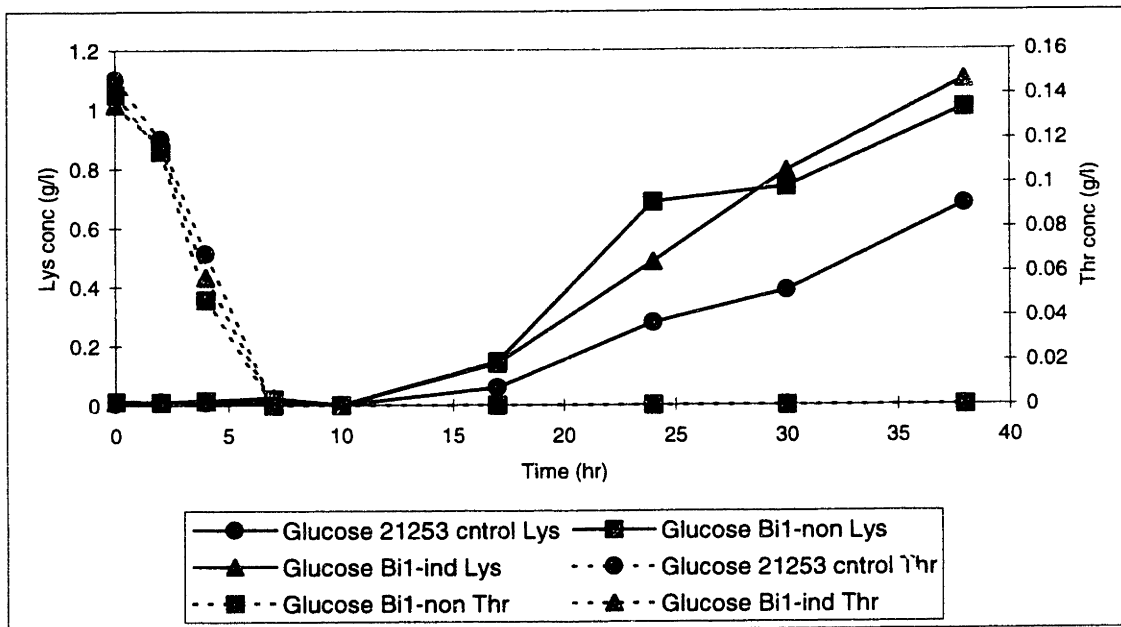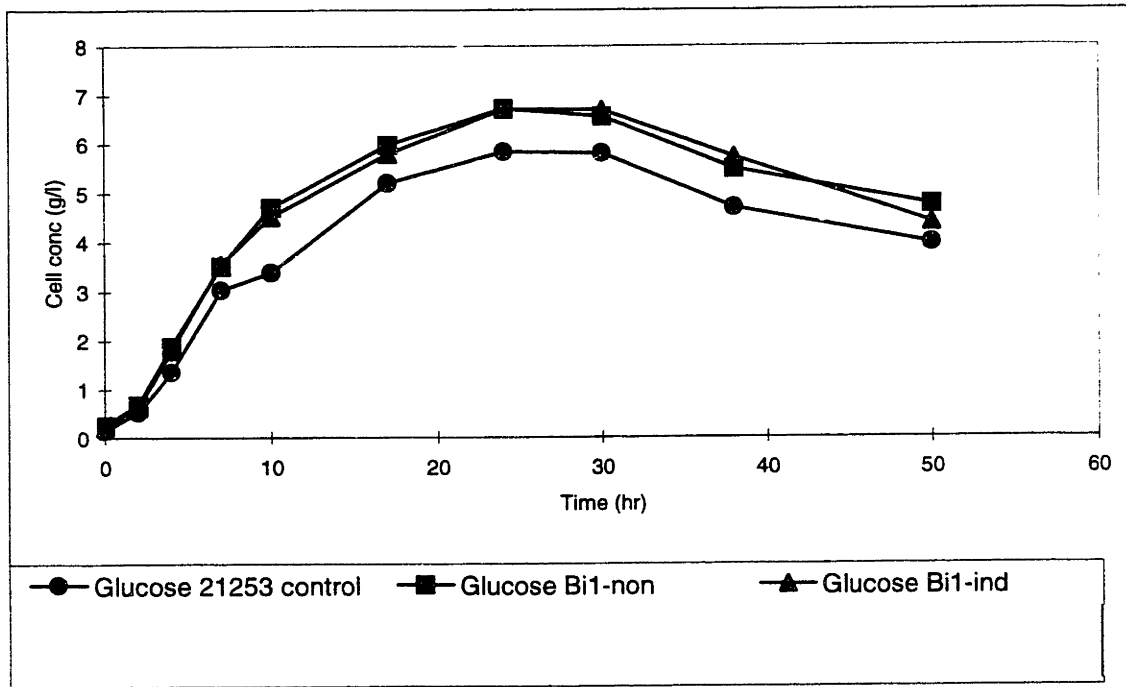
Figure 4.7 Growth, lysine production and threonine consumption of control 21253 strain, 253KDask with no IPTG in the culture medium and 253KDask with 10 μM of IPTG in the culture medium. Glucose was used as the sole carbon source.
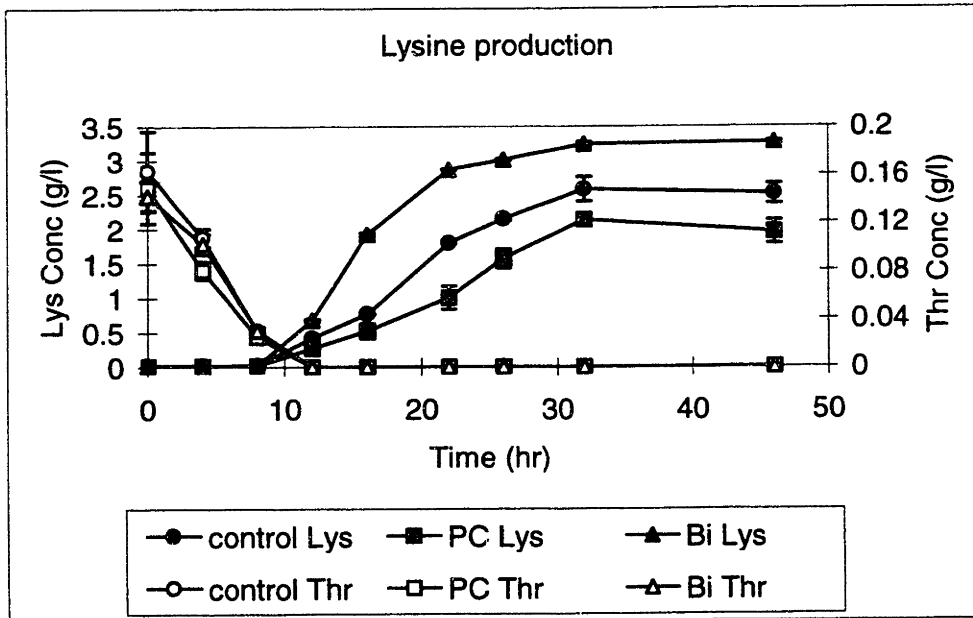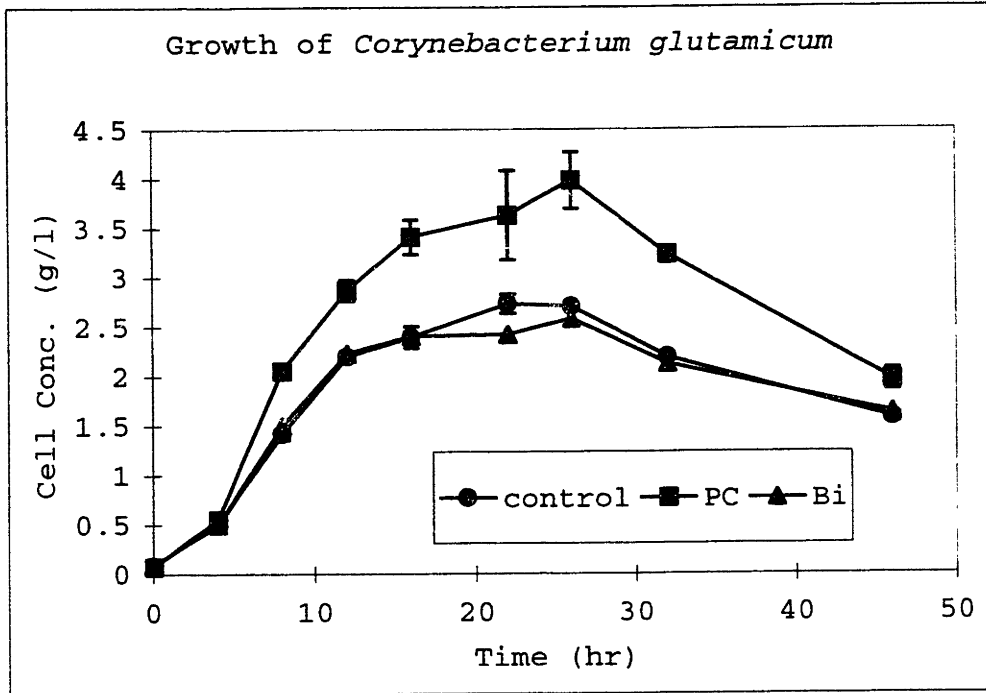
Figure 4.8 Growth, lysine production and threonine consumption of control 21253 strain, 2i253 strain carrying plasmid pKD7 and strain 253KDask with no IPTG in the culture medium. Lactate was used as the sole carbon source.

Strain	21253 control	253KDask-noninduced	253KDask-induced
Q (g lysine/l-hr)	0.03	0.04	0.05
q (mg lysine/g cell-hr)	5.5	6.4	7.9
C_{lysf} (g/l)	0.68	1.0	1.1
μ_m (h^{-1})	0.54	0.5	0.5
X_f (g/l)	5.8	6.6	6.7

Table 4.1 Batch culture parameters measured for 21253 control strain and 253KDask recombinant strain (with and without IPTG induction). Q: volumetric lysine productivity. q: specific lysine productivity. C_{lysf} : final lysine concentration in the shake flask. μ_m: specific growth rate. X_f: final cell concentration.

	21253 control	253KDask-non induced
Q (g lys/l-hr)	0.09	0.24
q (mg lys/g cell-hr)	42.3	106.6
C_{lysf} (g/l)	2.6	3.3
μ_m (hr^{-1})	0.35	0.35
X_f (g/l)	2.6	2.6

Table 4.2. Batch culture parameters measured for 21253 control strain and 253KDask recombinant strain (no IPTG induction). Q: volumetric lysine productivity. Q: specific lysine productivity. C_{lysf} : final lysine concentration in the shake flask. μ_m: specific growth rate. X_f: final cell concentration.
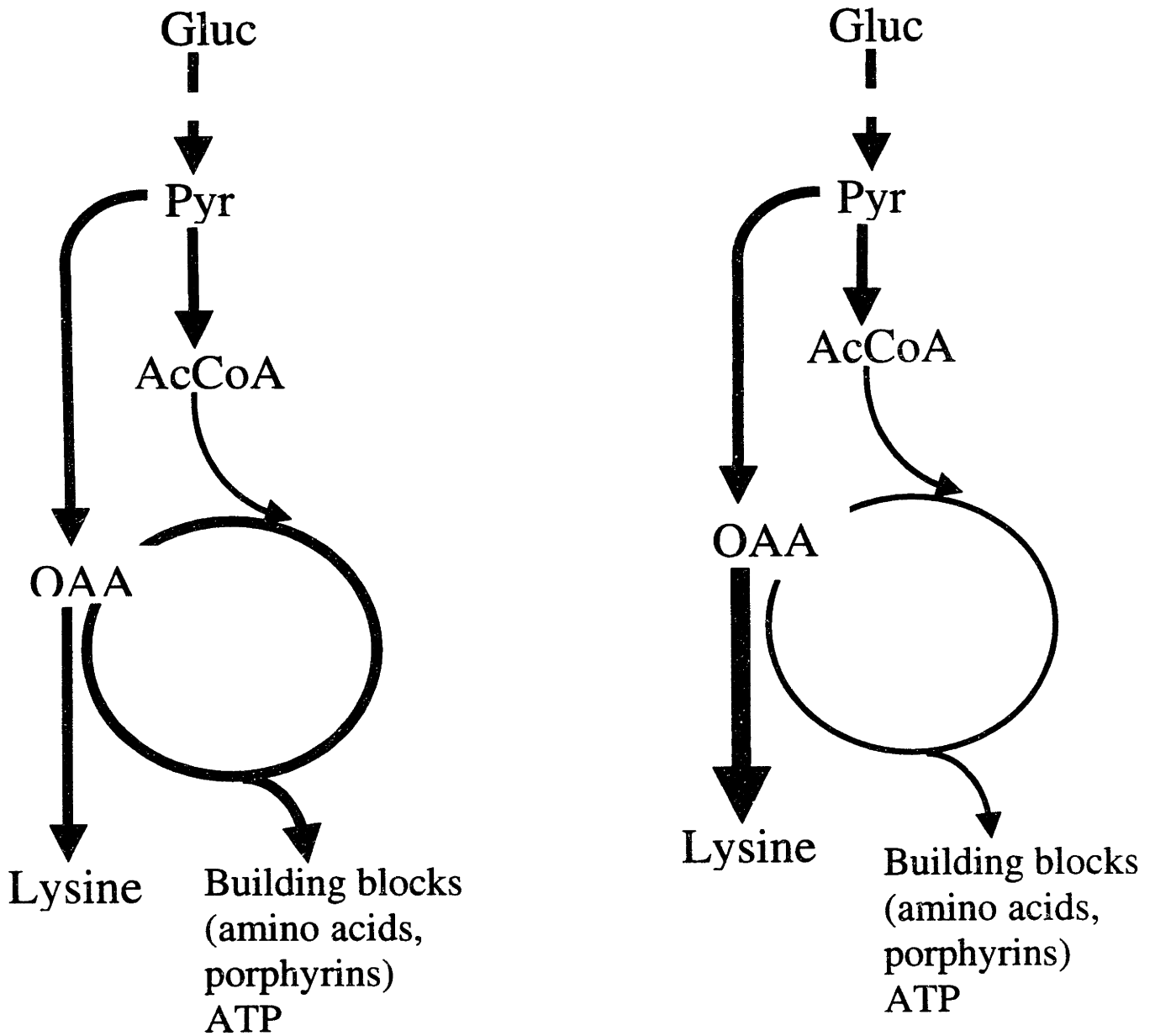
Figure 4.9 A metabolic model:the overexpression of aspartate kinase increases carbon flux towards lysine biosynthesis. Because of that, TCA cycle is depleted of carbon and less ATP and other building blocks (amino acids and porphyrins) are generated. That carbon flux redistribution translates into a better lysine productivity but poorer growth.

CHAPTER 5: INVESTIGATION OF THE ROLE OF BIOTIN IN *C. glutamicum* METABOLISM

5.A Introduction

5.A.1 Chapter Overview

The discovery that at least two enzymes that contain biotin as cofactor exist in *C. glutamicum*, one of which has already been characterized as pyruvate carboxylase, led us to the initiation of a research concentrating on the elucidation of the role of biotin in this strain's metabolism.

Two issues have been the focus of this research: first was the characterization of the second biotinylated enzyme in *C. glutamicum* from a biochemical and genetic point of view. In this chapter the purification of this enzyme is described as well as the N-terminal amino acid sequence. Some genetic aspects regarding this enzyme are also described based on the literature.

The second research topic concentrates on the isolation of the gene of another key enzyme in biotin metabolism, namely the gene of biotin ligase. This is the enzyme that performs the biotinylation i.e the attachment of biotin on a protein in order for it to be functional. Here we describe the isolation of a cosmid from the *C. glutamicum* cosmid library that contains the gene of biotin ligase from this organism. This was achieved by using complementation of an *E. coli* biotin ligase deficient strain. Some thoughts for future directions on this project are also given at the end of the chapter.

5.A.2 Acetyl-CoA carboxylase

The rapid expansion of technologies based on the interactions of biotin with very-high-affinity binding proteins, avidin and streptavidin might obscure the fact that biotin is, in the first place, a vitamin required by all forms of life. This is due to the fact that a number of enzymes in cell metabolism undergo biotinylation as a posttranslational

modification in order to be functional (Figure 5.1). These biotin-dependent enzymes have key roles in gluconeogenesis, lipogenesis, amino acid metabolism and energy transduction. More organisms have fewer than five protein species that are biotinylated. Biotin is covalently attached at the active site of these enzymes that are all carboxylases. They serve the transfer of carbon dioxide from bicarbonate to organic acids to form cellular metabolites. The mechanism of the the biotin-mediated carboxylation is outlined in figure 5.2. The structural differences among these enzymes as well as their catalytic sites are outlined in figure 5.3.

In all cells- whether these be bacteria, yeasts, fungi, plant or animal, acetyl-CoA is carboxylated by the biotin- dependent enzyme complex, acetyl-CoA carboxylase (ACC, EC 6.4.1.2), to give malonyl-CoA (Alberts and Vagelos 1972), which is then converted to fatty acids by fatty acid synthetase (FAS). Like other biotin-dependent carboxylation reactions, the overall reaction can be partitioned into two half reactions: (i) the carboxylation of the biotin prosthetic group by a MgATP- dependent process with bicarbonate and (ii) the transfer of the CO_2 group from the carboxylated prosthetic group to acetyl-CoA to form malonyl-CoA (Alberts and Vagelos 1972), (Polakis, Guchhait et al. 1973). A third subunit, the biotin carboxyl carrier protein (BCCP), carries the biotin prosthetic group covalently attached to a lysine residue. The various catalytic sites can be either in the form of an enzymatic complex, as in the case of the E. coli ACC or can form a single multifunctional polypeptide, as in the case of many eukaryotes and certain bacteria (Al-Feel, Chirala et al. 1992), (Ha, Daniel et al. 1994), (Gornicki, Podkowinski et al. 1994), (Ratledge 1982). Genes encoding acetyl-CoA carboxylase have been isolated from a number of species in the recent years.

Since acetyl-CoA carboxylase catalyzes the initial committed step on the de novo of fatty acid biosynthesis, it has been the subject of very intensive study (Cronan Jr. and Rock 1987), (Wakil, Stoops et al. 1983), (Volpe and Vagelos 1976). It is activated by citrate and is subjected to end product inhibition by malonyl-CoA. The carboxylase is also regulated by covalent modification of the protein by a phosphorylation / dephosphorylation mechanism (Kim 1983). It is now evident that the in vivo phosphorylation of the enzyme is performed by AMP- activated protein kinase, while

protein phosphatase 2A seems to be primarily responsible for dephosphorylation (Cohen and Hardie 1991).

Physiological studies to determine the rate limiting step in *E. coli* phospholipid biosynthesis have led to the suggestion that acetyl-CoA carboxylase is the controlling step. Since 94% of the ATP consumed in phospholipid biosynthesis is expanded in the synthesis of the fatty acid components, it appears logical that acetyl- CoA carboxylase could be a primary site for the regulation of lipid synthesis (Jackowski, Cronan Jr. et al. 1991). I bacteria, such as *E. coli* where fatty acids are primarily incorporated into membrane phospholipids, regulation of fatty acid synthesis is closely coordinated with rates of cell growth. Indeed studies have shown that the rates of transcription of the genes that encode all four subunits of acetyl-CoA carboxylase in *E. coli* are directly related to the rate of cell growth (Li and Cronan 1993). This fact leads to the possibility of controlling the channeling of acetyl-CoA between growth and product formation by genetic manipulation of acetyl- CoA carboxylase. An indirect approach to this was through the *in situ* expression of streptavidin that binds and limits biotin availability (Guan and Wurtele 1996). The possibility of excercising such a control is of pivotal importance in the development of biological processes of amino acid production, where cell permeability is an essential factor in regulating the outcome, as for example in the case of glutamate production (Aida 1973), (Hubbard and Hall 1967). In this case the most important factors regulating the permeability of glutamic acid- producing bacteria are biotin, oleate, surfactants and penicillin. It now appears that the action of biotin, oleate and surfactants causes a change in the fatty acid composition of the cell membrane, particularly in the content of oleate and phospholipid, and as a result the permeability of the cell membrane is altered, allowing glutamate to leak out. It has been shown in the literature that when biotin concentration was increased, the fatty acid content increased and the amino acid yield fell sharply.

The goal of the current research is to identify and chracterize acetyl –CoA carboxylase from *C. glutamicum* following a protein purification approach of the lower biotinylated (~ 80 kDa) biotinylated enzyme present in those cells.

5.A.3 Biotin Ligase

The last step in the effort to investigate the role of biotin in amino acid production involves the identification and characterization of the gene of biotin ligase in *C. glutamicum*.

Biotin is bound to specific proteins via an amide linkage between the biotin carboxyl group and a unique lysine amino group, a reaction catalyzed by biotin ligase (birA) (Samols, Thornton et al. 1988). In *E. coli* biotin ligase catalyzes the transfer of the biotinyl prortion of biotinyl-5'-AMP to the ε-amino group of the specific lysine in the biotin carboxyl carrier protein subunit of acetyl CoA carboxylase, which is the only protein for which biotin is a known cofactor in that strain. Besides this function, biotin ligase is also the repressor protein of biotin biosynthesis in *E. coli*. Complexed with its corepressor, biotinyl-5'-AMP, biotin ligase functions as a classical repressor by binding to an operator site that overlaps the promoters for divergent *bioA* and *bioBCDF* operons (Cronan 1989), (Eisenberg 1984). Two holorepressor monomers bind cooperatively to the two half sites of the operator (Abbott and Beckett 1993) and so repress the initiation of transcription by blocking binding of the RNA polymerase to the two promoters, Pa and Pb (Otsuka and Abelson 1978). It is interesting to note that biotin ligase catalyzes adenylation of biotin, thereby synthesizing its own corepressor. Thus biotin ligase provides feedback control at the biotin synthetic operon by inhibiting transcription in the presence of adequate biotin concentrations. On the basis of the properties of various mutations in biotin ligase, it is thought that DNA binding is mediated by the amino-terminal domain of the protein, whereas the enzymatic activity and corepressor binding site reside in the carboxy-terminal domain (Buoncristiani, Howard et al. 1986). So biotin ligase is a bifunctional protein capable of carrying out the following series of reactions (Eisenberg, Prakash et al. 1982):

$$\text{Biotin} + \text{ATP} + \text{repressor } (R) \rightarrow (\text{biotinyl-5'-AMP} \cdot R) + \text{P-Pi} \qquad (1)$$

$$(\text{Biotinyl-5'-AMP} \cdot R) + \text{apoenzyme} \rightarrow \text{holoenzyme} + \text{AMP} + R \qquad (2)$$

$$(\text{Biotinyl-5'-AMP} \cdot R) + \text{operator} \ (O) \rightarrow (\text{biotinyl-5'-AMP} \cdot R \cdot O) \qquad (3)$$
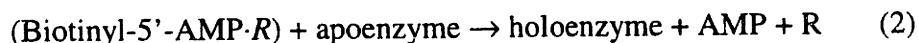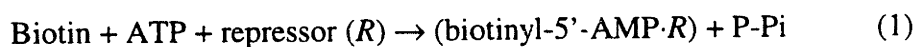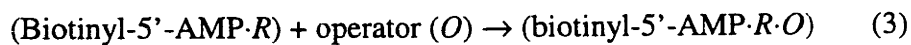
The enzyme biotin ligase is currently emerging as a tool of great importance with the development of novel applications for the enzyme, especially in the field of cell surface ligand- receptor interactions (Altman, Moss et al. 1996), (Braud, Allan et al. 1998), (McMichael and O'Callaghan 1998), (O'Callaghan C, Byford et al. 1999). By engineering constructs containing a protein of interest joined in frame to a peptide recognition sequence for biotin ligase it is possible to produce recombinant protein which can be biotinylated at a defined residue in the recognition sequence. Such highly specific biotinylation permits attachment of the protein to streptavidin derivatives, which can themselves be conjugated to other chemicals, especially fluorochromes, or to surfaces or beads. As streptavidin derivatives are tetrameric and bind four biotin molecules, this provides a straight forward method for multimerizing proteins and increasing the avidity for their receptors. Moreover, the orientation of the biotinylated protein with respect to the streptavidin derivative is precisely defined as a result of the highly specific biotinylation reaction.

Besides *E. coli* the gene of biotin ligase has been sequence from a wide variety of other organisms. From prokaryotic sources the gene of biotin ligase from *Mycobacterium tuberculosis, Paracoccus denitrificans* (Xu, Matsuno-Yagi et al. 1993), *Rhodobacter capsulatus, Methanobacterium thermoautotrophicum* (Mukhopadhyay, Stoddard et al. 1998), *Thermus thermophilus* (Lauer, Rudd et al. 1991) and *Bacillus subtilis* (Bower, Perkins et al. 1995). From eukaryotic and mammalian sources the biotin ligase gene has been sequenced in *Saccharomyces cerevisiae* (Cronan and Wallace 1995) and *Homo sapiens* (Suzuki, Aoki et al. 1994), (Leon-Del-Rio, Leclerc et al. 1995). In almost all the cases a complementation of an *E. coli* mutant strain that is deficient of biotin ligase was employed for the gene isolation. The phenotype of this mutant is complete inability to grow at 42 °C in LB medium and the requirement for high biotin concentration present in M9 minimal medium for the strain to grow at 30 °C.

In the following paragraphs a presentation of various attempts to isolate the gene of this enzyme from *C. glutamicum* are presented and the results of these experiments are further discussed.

5.B Materials- Experimental Methods

5.B.1 Strains-Media

ATCC strain 21253 was utilized for producing protein crude extract and as a donor of *C. glutamicum* genomic DNA. All protein and genetic manipulations were preformed as described in previous chapters. All media used for *C. glutamicum* cultivation have been described previously.

E. coli mutant strain BM4062 [*birA* (TS)] deficient in biotin ligase (Barker and Campbell 1980), (Campbell, Chang et al. 1980) was a kind gift from Professor Alan Campbell (Stanford University) and was utilized for complementation experiments. LB was utilized as the rich medium for *E. coli* growth, while M9 minimal medium was used as a glucose minimal medium, prepared as described in the literature (Miller 1972).

5.C Results

5.C.1 Protein purification of Acetyl-CoA-Carboxylase

After cell lysis, the protein crude extract was resuspended in homogenization buffer (triethanolamine 20mM, NaCl 20 mM, dtithiothreitol 1 mM, pH 7.6) at a final protein concentration of 15-20 mg/ml. Initial attempts to further concentrate and achieve an initial purification of the protein by ammonium sulphate and PEG precipitation could not provide an optimal concentration for any of these agents that could lead to a substantial change in the protein purity. The cell lysate had a final protein content of 600-820 mg and a final volume of about 40 ml. Because of the presence of significant turpidity, due to the presence of nucleic acids and lipids in the cell homogenate, an extra

centrifugation was performed at 15,000 rpm for an hour at 4 °C followed by a treatment with DNase and Rnase enzymes for 15 min at 37 °C.

In the next step the 40 ml protein crude extract was loaded on a DEAE- Sephacel anion exchange column. The dimensions of this column were 2.5×25 cm and was previously equilibrated by homogenization buffer (about 2 lt) at a flow rate of 1 ml/min and at 4 °C. The protein crude extract (40 ml) was loaded on the column at a flow rate of 0.5 ml/min. After the completion of the loading the column was further washed with about 500 ml of 10% buffer B (20 mM triethanolamine, 1M NaCl) and 90% of buffer A (20 mM triethanolamine, 20 mM NaCl), so that the NaCl concentration in the column during the wash was about 118 mM.

Elution of the protein was achieved with a 500 ml linear NaCl gradient from 118 mM initial concentration (10% B, 90% A) to 1 M final (100%B). Thirtysix fractions were collected of 10 ml each and the presence of the 80 kDa biotinylated enzyme was analyzed using the same western blot technique described earlier. Most of the enzyme appeared to be present in fractions 10-23. We found that the early fractions contained exclusively the 80 kDa biotinylated protein with no contamination by pyruvate carboxylase while the later ones containing some pyruvate carboxylase protein.

Fractions 10- 23 were combined in a total volume of 130 ml and a concentration of this protein mixture followed by hollow fiber ultrafiltration and a 10,000 MW cutoff. The final volume of the concentrate was approximately 38 ml.

In the next step, the protein concentrate was loaded on a 10 ml agarose tetrameric avidin column. This column had previously been equilibrated with buffer PBS (phosphate 20 mM, NaCl 20 mM, pH 7.5) at a flow rate of 1 ml/min according to the instructions of the manufacturer (Pierce Chemical Company, Rockford Illinois). The protein solution was loaded at a flow rate of 0.33 ml/min. After the loading a washing step followed using 100 ml of PBS and a flow rate of 1 ml/min. The elution of the protein was performed using 70 ml of guanidine·HCl 8M, pH 1.5 and a flow rate of 1 ml/min as suggested by the manufacturer. The 70 ml elution solution was mixed with 24.5 ml Tris buffer 1M and pH 7.5 in order to be neutralized and in order to dilute the concentrated guanidine·HCl. The eluate was again concentrated using hollow fiber ultrafiltration with a 10,000 MW cutoff.

In order to remove guanidine·HCl, the concentrated eluate from the tetrameric avidin column was then passed through a PD10 gel filtration column that was previously equilibrated with buffer Bis-Tris 20 mM, NaCl 20 mM and dithiothreitol 1 mM (pH 6.8). After loading this column, elution was performed with the same Bis-Tris buffer and the first protein peak containing the protein was again concentrated by ultrafiltration (centrifugation) to a final volume of 0.7 ml (MW cutoff 10,000).

In the final step of the purification procedure, the protein mixture was passed through a monoQ R/R anion exchange column prepacked by the manufacturer (Pharmacia) to a final volume of 1 ml. This column had previously been equilibrated with the same Bis-Tris buffer mentioned above and the loading of the protein mixture was also done at a flow rate of 1 ml/min. After washing the column with 20 ml of Bis-Tris buffer elution of the protein was performed with a linear gradient of NaCl from 20 mM to 1M (20 ml total volume) and at a flow rate of 1 ml/min. Twenty 1 ml fractions were collected from the column. The 80 kDa biotinylated enzyme appeared in fractions 13-17 which were mixed together and concentrated by centrifugal ultrafiltration to a final volume of 0.5 ml. The final amount of protein obtained by this procedure was about 250 µg. The overall process established for the purification of the 80 kDa biotinylated enzyme is outlined in figure 5.4.

By performing SDS-Page electrophoresis and staining with silver staining the purity of the 80 kDa protein was verified (Figure 5.5). The purity was also assessed by tranfering the protein from a polyacrylamide gel to a nitrocellulose membrane and then staining the membrane using colloidal gold. Only one major protein band was detected by this method with a slight contamination by another protein with a molecular weight of about 120 kDa (most likely pyruvate carboxylase).

5.C.2 N-terminal amino acid sequence of acetyl-CoA- carboxylase

About 100 µg of the purified 80 kDa biotinylated enzyme were subjected to N-terminal sequence using the classical Edeman degradation reaction. The N- terminal sequence was performed at MIT's biopolymers lab and the sequence of the first 25 amino

acids of this protein were obtained (Figure 5.6). No blocking of the initial methionine was detected. This N- terminal sequence obtained matched exactly the sequence of the acetyl-CoA carboxylase enzyme which at that time appeared in the literature (Jager, Peters-Wendisch et al. 1996). Amino acid sequences of internal peptides obtained by chemotrypsin digestion of the protein were also obtained.

5.C.3 PCR-based approach for obtaining the biotin ligase gene

To isolate the *C. glutamicum* gene of biotin ligase, two different approaches were followed.

The first one was similar to the approach followed previously for obtaining the gene of pyruvate carboxylase. The protein sequences of biotin ligases isolated from eleven different organisms were examined and two strongly homologous regions were identified. The biotin ligase sequences examined were derived from the following organisms: *M. tuberculosis* (acc. no. Z92771), *E. coli* (acc. no. AE000471), *B. subtilis* (acc. no. U20445), *P. denitrificans* (acc. no. L02354), *R. capsulatus* (acc. no. AF029365), *M. thermoautotrophicum* (acc. no. AE000942), *A. thaliana* (acc. no. U41369), *B. pertussis* (acc. no. AF016461), *Schizosaccharomyces pombe* (acc. no. Z97992), *S. cerevisiae* (acc. no. NC_001136) and *Homo sapiens* (acc. no. NM_000411). The first conserved region was corresponding to an ATP binding site with a motif QaaGRGRaG(R/N) while the second, a few amino acids down from the ATP binding site, was corresponding to a biotin binding site with a motif (L/V/I)KWPND(L/V/I). Three types of PCR DNA primers were designed for each one of those two regions. The first set of primers was based on the *M. tuberculosis* DNA sequence, the second was a set of degenerate primers based on the bacterial DNA sequences and the third was again a set of degenerate primers based on all sequences mentioned above.

All nine possible PCR combinations were tried using *C. glutamicum* genomic DNA as template. 56 °C was used as the annealing temperature, *Taq* was utilized as the DNA polymerase and 35 PCR cycles were performed in total. Two sets of primer combinations produced DNA bands with molecular weights of about 210 bp as expected

based on the biotin ligase genes published in the literature. The first combination of primers was corresponding to the *M. tuberculosis* based PCR primers (birAbmyc-birAemyc) while in the second set the beginning primer was a degenerate primer based on a consensus homology among bacterial biotin ligases while the end primer was again corresponding to the *M. tuberculosis* DNA sequence (birAbbact-birAemyc).

The two PCR products generated were ligated into vector pCR-Script, in a procedure similar to the one described previously and then subjected to sequencing using T3 and T7 primers. None of the two PCR fragments gave an open reading frame with significant homology to any known biotin ligases. Due to the fact that biotin ligases have a very low homology on their DNA level in general, the PCR- based approach for isolating the gene of this enzyme was suspended.

5.C.4 Complementation tests using mutant BM4062

As mentioned previously, in most of the cases that have been reported in the literature, the gene of biotin ligase from various sources has been obtained by complementation tests utilizing *E. coli* birA deficient strains. Since a cosmid library of *C. glutamicum* was available to us, such an approach was followed during our experimental efforts to obtain the gene of biotin ligase from this organism.

The 291 cosmids that contain the complete genome of *C. glutamicum* were isolated from 5 ml cultures and were mixed in groups of 10, except for the last (29 th) group that was containing 11 different cosmids. Each cosmid was present in these groups at an equal amount except for about twenty of them, due to low yields achieved during their isolation. The cosmids were stored at −80 °C until used for transformation purposes.

Each one of the 29 cosmid mixtures was then transformed into the *E. coli* mutant strain BM4062. Initial selection was done based on growth on LB/agar plates at 42 °C. This resulted in the isolation of sixteen colonies derived from seven cosmid mixtures. This colonies were able to regrow when striked on LB plates and left at 42 °C overnight, even though at apparently different levels. In order to assess the growth level of the colonies they were next inoculated into 5 ml LB liquid medium and incubated at 42 °C

overnight. Only five of the sixteen colonies were able to reach levels of growth similar to that of the wildtype *E. coli* strain used as a control for that experiment while the rest showed an intermediary to minimal level of growth.

In order to verify whether those colonies were indeed positives, a complementation test of the second phenotype of strain BM4062 was performed. For that purpose all five colonies isolated previously as well as the wildtype were inoculated into M9 glucose minimal medium with (control) and without the presence of 4 μM of biotin in the culture medium. The tubes were left to grow overnight at 30 °C. While the wildtype was able to grow both with and without biotin, only one colony, named 15birA, was able to grow without biotin at levels comparable to that of the wildtype. The other four reached a high final biomass (almost identical to that of the wildtype) in the presence of biotin and showed an intermediary growth profile in the absence of this vitamin. These growth profiles were verified by repeating this complementation test a second time.

5.D Discussion

5.D.1 Acetyl-CoA Carboxylase

Metabolic Engineering has been defined as the area of the rational redirection of the flux of metabolites towards production optimization and cell physiology improvement. Even though a great portion of research has been directed towards the first goal, there are few examples in the literature that have focused mainly on the second. The investigation of the role of biotin in amino acid production is one of the very few examples of a single metabolite been able to affect both product formation and cell growth simultaneously, due to its role as a rare cofactor. As such, the elucidation of the enzymes that utilize biotin becomes the very first step in an effort to create such tunable systems that can allow someone to redirect this vitamin either towards production or towards growth.

The approach that was followed on identifying acetyl-CoA-carboxylase i.e. obtaining a pure protein and identifying its amino acid sequence has relied on the

presence of the biotin molecule which allows someone to use immobilized avidin as an affinity column. The first step however in the purification procedure, and before we reach the affinity chromatography step has to involve the separation of the two biotinylated enzymes that exist in *C. glutamicum* and the isolation of fractions with the 80 kDa protein as the sole biotinylated enzyme in them. As such, an anion exchange chromatography column (DEAE Sephacel) was used. Its selection was based on previous reports in literature that have utilized this type of chromatography for pyruvate carboxylase purification (Mayer, Wallace et al. 1980), (Rylatt, Keech et al. 1977). Indeed the use of this column provided fractions in which acetyl-CoA-carboxylase was entirely or almost completely separated from pyruvate carboxylase as that was evident by western blots. The elution of acetyl-CoA-carboxylase started occurring at a NaCl concentration estimated to be about 350 mM, reached a pick at around 550 mM and continued until almost the end of the protein elution. On the other hand, the elution of pyruvate carboxylase started only after NaCl concentration in the column had reached approximately 600 mM. It is interesting to note that this type of chromatography did not produce any protein fractions with pyruvate carboxylase isolated from acetyl-CoA-carboxylase.

After the fractions of interest were combined and concentrated, they were loaded on a tetrameric avidin column at a very low flow rate (0.33 ml/min). The slow flow rate was chosen because of reports that have appeared in the literature suggesting that avidin is a slow binding inhibitor of pyruvate carboxylase. It has actually been calculated that the rate constant of formation of the enzyme- avidin complex is 1.42×10^5 $M^{-1}s^{-1}$ for the enzyme isolated from chicken liver. The elution from the column was performed as suggested by the manufacturer using a very high concentration of a chaotropic agent, such as guanidine·HCl and a very low pH of 1.5. These conditions are necessary because of the very strong binding of tetrameric avidin with biotin. In fact the avidin-biotin interaction is the strongest non covalent bond that exists in nature with a dissociation constant of 10^{-14}.

Even though the expectation after the elution from the avidin column was a completely pure acetyl-CoA-carboxylase, that was not the case indeed. There were impurities present, as they were monitored after the removal of guanidine·HCl, especially in the lower molecular weight level. That did not come as a surprise. Tetrameric avidin is

a protein with a very high pKa, of about 10. This is exactly the reason why the column eluent included so many impurities: since the protein binding occurres at pH 7.5, besides the avidin-biotin interaction that leads to the capture of the biotinylated protein, electrostatic interactions also take place, leading to a non specific binding of proteins that are strongly negative (low pKa). This is exactly the reason why a final purification step is necessary for the removal of the impurities, and as such a monoQ R/R anion exchange column was selected.

Two protein peaks were detected on the chromatograph. Even though acetyl-CoA-carboxylase was present in all of them, most of the protein appears to be eluted in the second protein peak which corresponds to 480 mM-600 mM NaCl concentration. The detection for this last step was performed with colloidal gold staining instead of western blot. As mentioned in the discussion section, only a minor contamination by pyruvate carboxylase was evident by this very sensitive protein detection method. This contamination did nopt appear at all when SDS-Page followed by silver staining was performed. We therefore concluded that the 80 kDa biotinylated enzyme was purified to complete purity.

In terms of yield, from 600-800 mg of total protein loaded on the DEAE-Sephacel column, 250 µg of pure protein were obtained, so that corresponds to a 0.04%-0.03% yield. Obviously this is a very low yield and it is likely that most of the protein was lost during the first step of anion exchange. As mentioned previously. only the fractions that appear to contain acetyl-CoA-carboxylase completely separated from pyruvate carboxylase were further used during the purification procedure. A significant amount of the 80 kDa enzyme was eluted at higher salt concentrations, together with pyruvate carboxylase, and this is most likely the step of major loss of the protein of interest since those fractions were disregarded.

The final point of interest is the fact that after the avidin column step, the 80kDa band was the major protein appearing on silver stained acrylamide gels, while all the other contaminants were appearing in much smaller amounts. It is therefore speculated that during the initial steps of cell lysis and anion exchange chromatography, a second subunit most likely corresponding to the acetyl-CoA binding site of the holoenzyme was

detached. The fact that no enzymatic assays were available at the time of the purification does not let us conclude in complete certainty the presence of this second subunit and the denaturation of the native enzyme. However based on the literature, this subunit is almost always present is acetyl-CoA-carboxylases purified from various sources and we would expect that this is the case in *C. glutamicum* also. Further work, focusing on the enzymatic properties of acetyl-CoA-carboxylase in the crude and purified form is needed in order to elucidate this point.

5.D.2 Biotin Ligase

Two were the main reasons behind our efforts to sequence and clone the gene of biotin ligase from *C. glutamicum*. As mentioned above, the first one has to do with the fact that biotin ligase is the very last enzyme involved in the utilization of biotin that has not been identified in this strain. That is especially true after the last reports in the literature that provide the sequences of the genes involved in the biotin synthesis pathway i.e. 7,8-diaminopelargonic acid (DAPA) aminotransferase (*bioA*), dethiobiotin (DTB) synthetase (*bioD*), and biotin synthase (*bioB*) in both *C. glutamicum* and *C. lactofermentum* (Serebriiskii, Vassin et al. 1996), (Hatakeyama, Kohama et al. 1993a), (Hatakeyama, Kohama et al. 1993b). The second one has to do with the different levels of expression shown by the three different constructs that contain the gene of pyruvate carboxylase. As mentioned previously all these constructs contained the gene of pyruvate carboxylase but different DNA regions upstream and downstream of the gene. Since there is precedence in the literature of the presence of biotin ligase immediately after the gene of pyruvate carboxylase in *M. thermoautotrophicum* (Mukhopadhyay, Stoddard et al. 1998) the question that we would like to address is whether this is the case in *C. glutamicum* also and how the simultaneous presence of biotin ligase and pyruvate carboxylase in a multicopy plasmid would influence the expression level of the second enzyme.

The first approach for the gene isolation of biotin ligase was similar to that previously followed for pyruvate carboxylase. Unlike the *pc* case though the results that

were obtained for the biotin ligase case were not encouraging. None of the candidate PCR products produced by this approach gave any significant homology to other biotin ligases. This is not as surprising as it may look at first. Indeed the level of homology among biotin ligases reported in the literature is not as strong as that observed among pyruvate carboxylases. For example the level of homology between the biotin ligases of *B. subtilis* and *E. coli* is 27% (Bower, Perkins et al. 1995), while between those of *P. denitrificans* and *E. coli* is 31% (Xu, Matsuno-Yagi et al. 1993). And even though there is a highly conserved biotin binding site in all these proteins (conserved even on the DNA level) the degree of homology at the ATP binding site is much lower. It is also interesting to note that the homology approach has not been followed in any reports of the cloning of biotin ligase from any organism so far.

On the other hand the complementation approach was the one that finally gave a few candidate cosmids. Out of the five cosmids initially isolated by screening for temperature tolerance, only one grew at levels similar to those of the wildtype *E. coli*. Most likely the other four cosmids contain only parts of the gene of biotin ligase or they are missing DNA regions important for the full expression of the gene. The cosmid that gave the complementation belonged to group number 15 that contains none of the cosmids that contain the gene of pyruvate carboxylase (namely IIE9, IIIB7, IIIG7 and IIIF10). So the initial hypothesis that biotin ligase may be to a chromosomal region close to the gene of pyruvate carboxylase is proven not to be the case in *C. glutamicum*. It is important to note that the plasmids that contain the *pc* gene were also introduced into the biotin ligase deficient *E. coli* srain BM4062 without giving any complementation either of its heat sensitive phenotype or the requirement for the presence of elevetated biotin concentration when M9 minimal medium is used to grow the strain.

After isolating the cosmid giving the complementation, further experiments were performed in order to isolate and clone a small DNA fragment that would contain the complete or part of the gene of biotin ligase. However these experiments were fruitfull for some time due to an unusually high smearing appearing on the agarose gel after digesting with various restriction enzymes. Recently a 3 kb DNA fragment was isolated from this

cosmid that gave complementation of the *E. coli* mutant strain. Further work is needed in order to obtain the sequence of the biotin ligase gene.
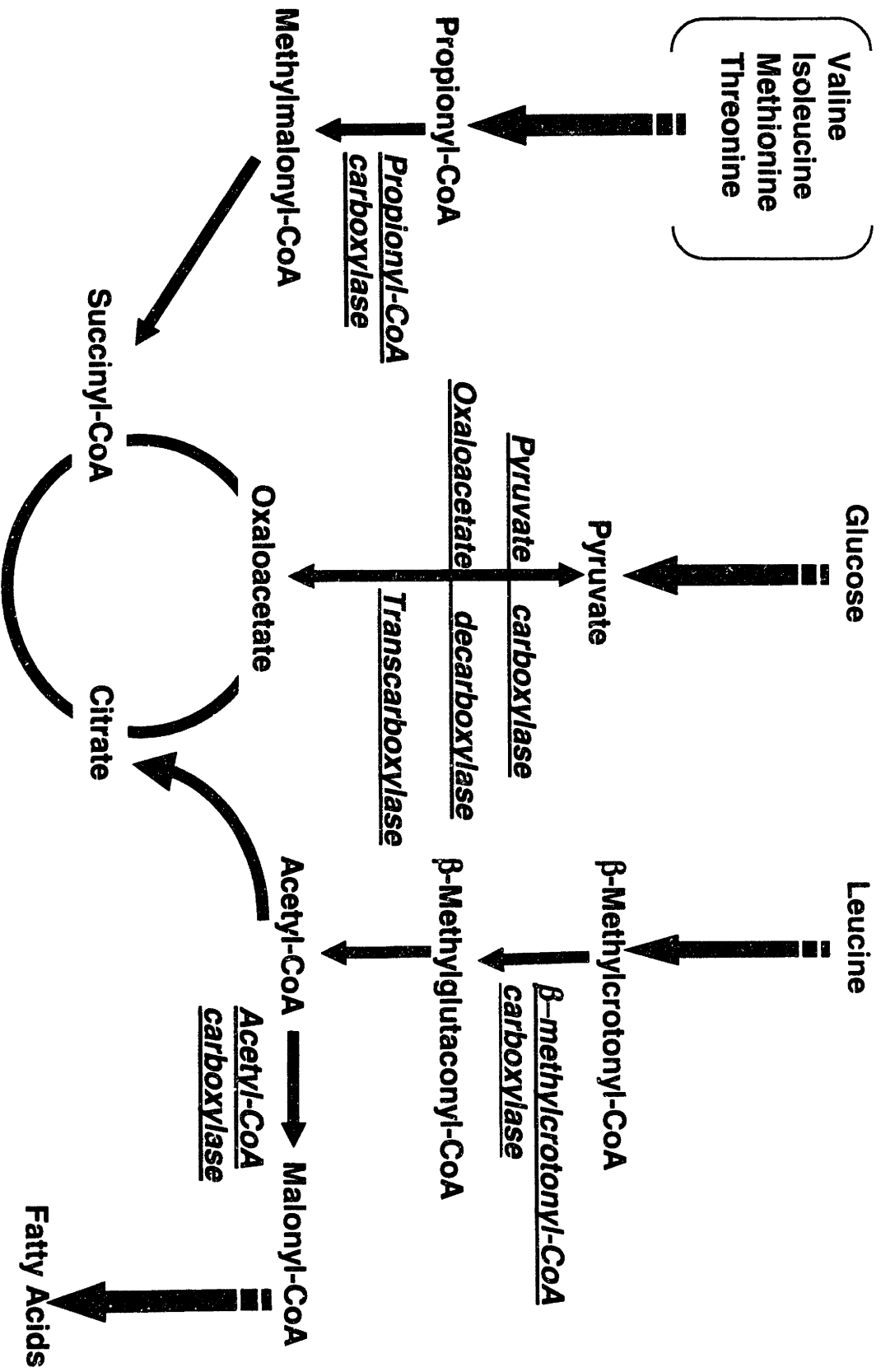
Figure 5.1 Biotinylated enzymes within the cell metabolism

164
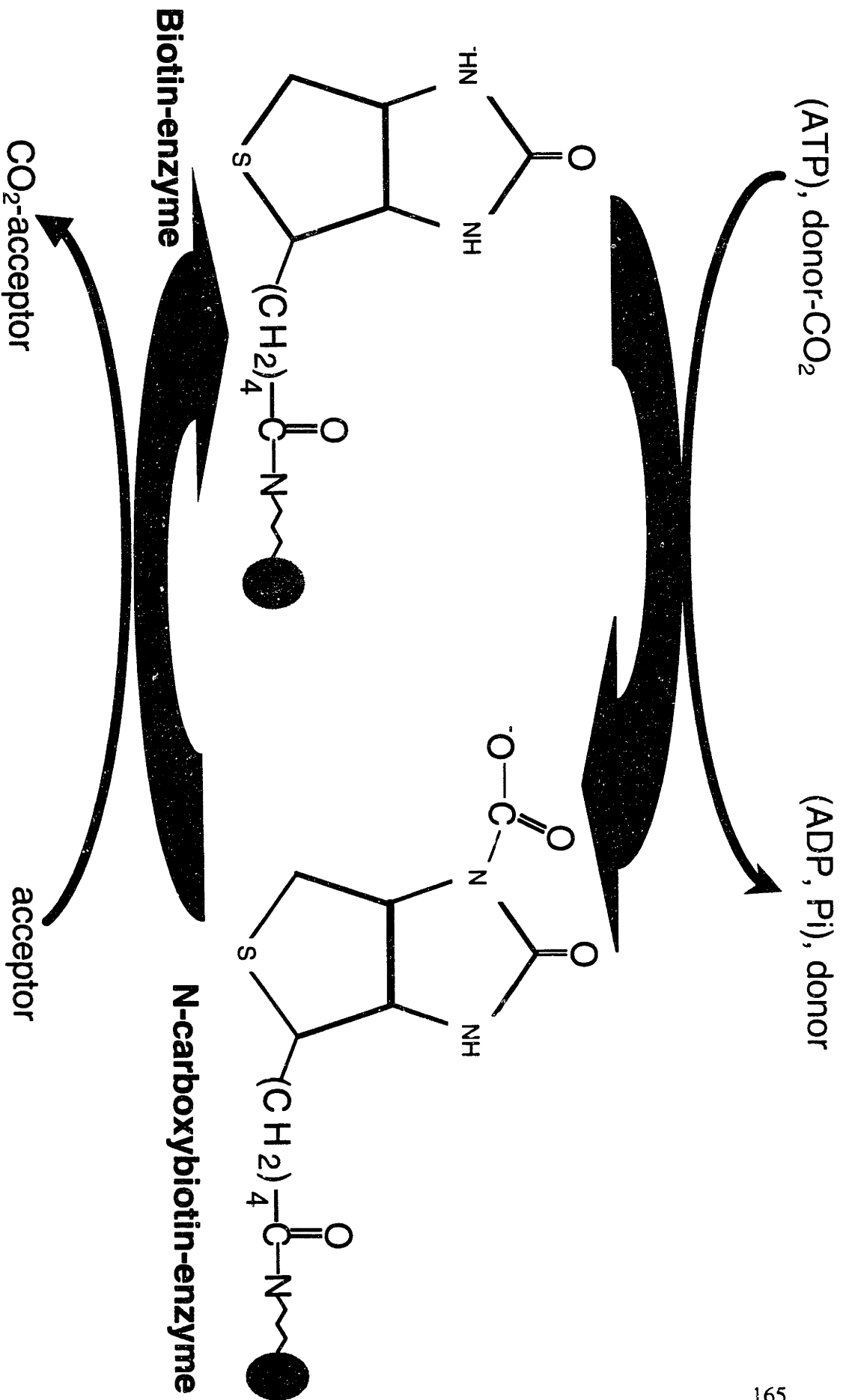
Figure 5.2 Enzymatic mechanism of biotin mediated carboxylation.

165
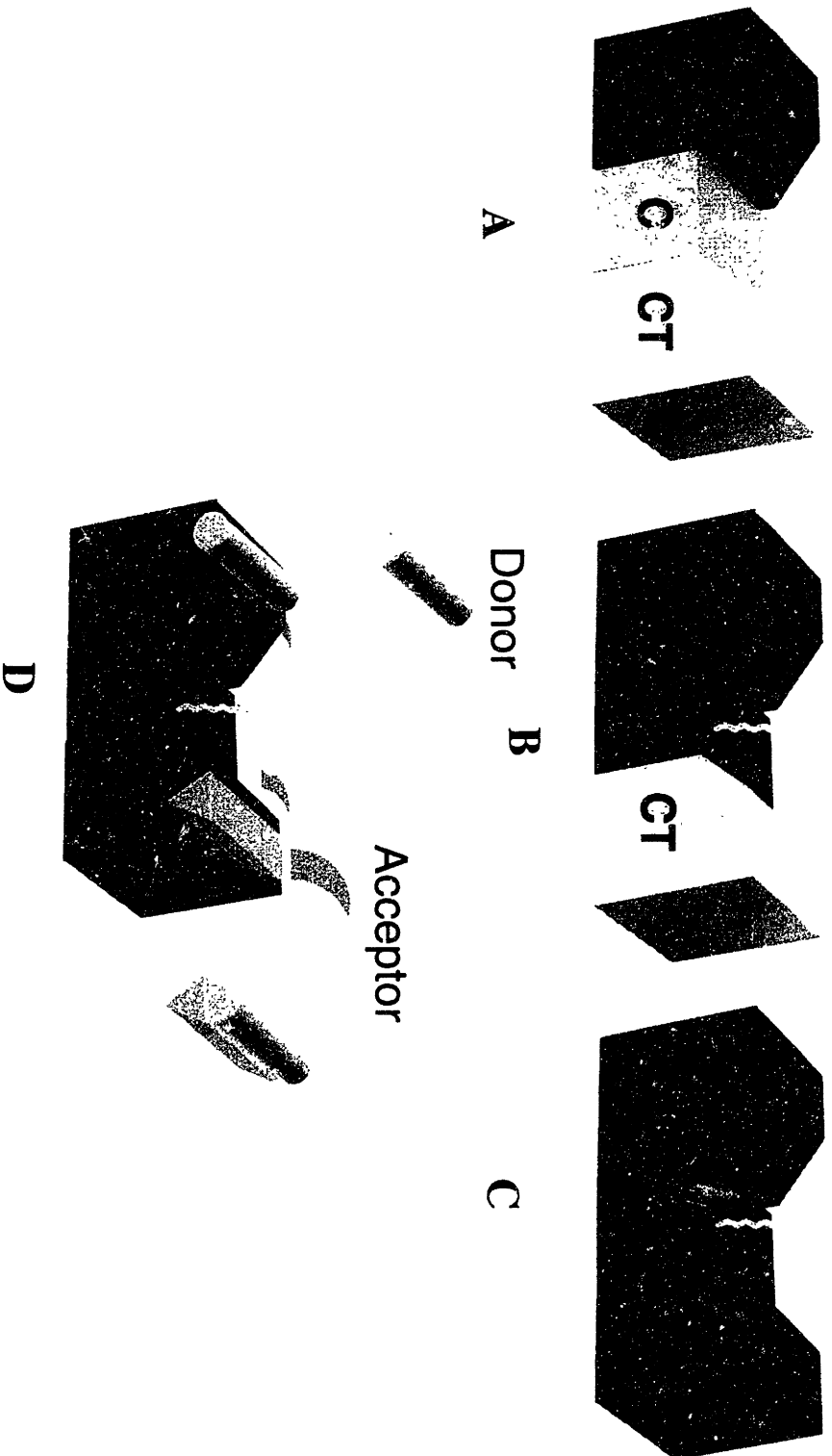
Figure 5.3 Feodor Lynen's world of biotinylated enzymes. BC: biotin carboxylation site. C: biotinyl carrying site. CT: carboxyl-transfer site. (A) The three catalytic sites can be distinct subunits (AcCoA-carboxylase of *E. coli*). (B) Two catalytic sites are on the same subunit (*Achromobacter* β-methylcrotonyl CoA carboxylase). (C) All catalytic sites are on the same subunits (yeast pyruvate carboxylase). (D) A schematic represantation of biotin mediated carboxylation.

166
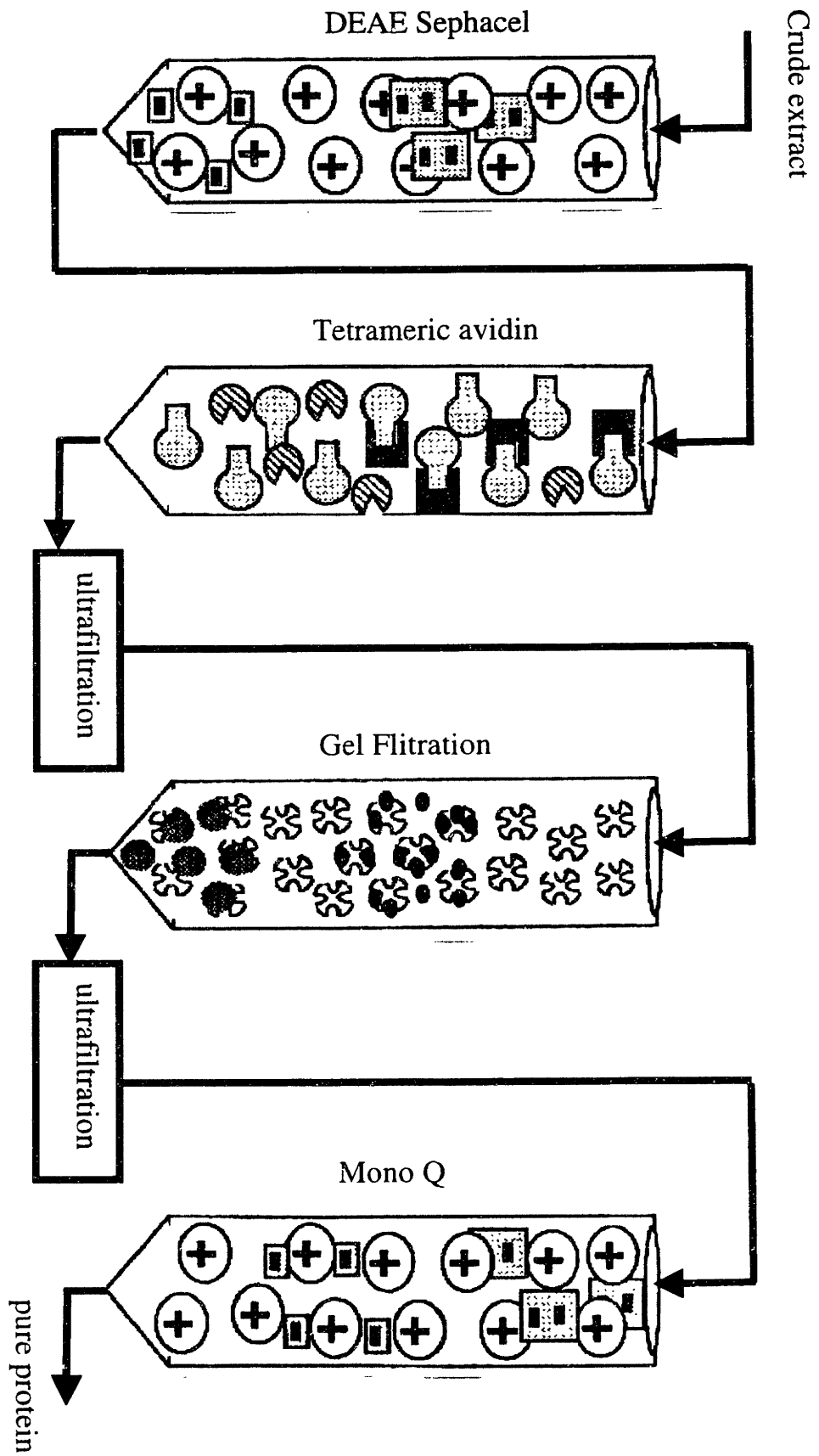
Figure 5.4 Purification scheme of acetyl-CoA-carboxylase of *C. glutamicum*

DEAE Sephacel

Crude extract

Tetrameric avidin

ultrafiltration

Gel Flitration

ultrafiltration

Mono Q

pure protein
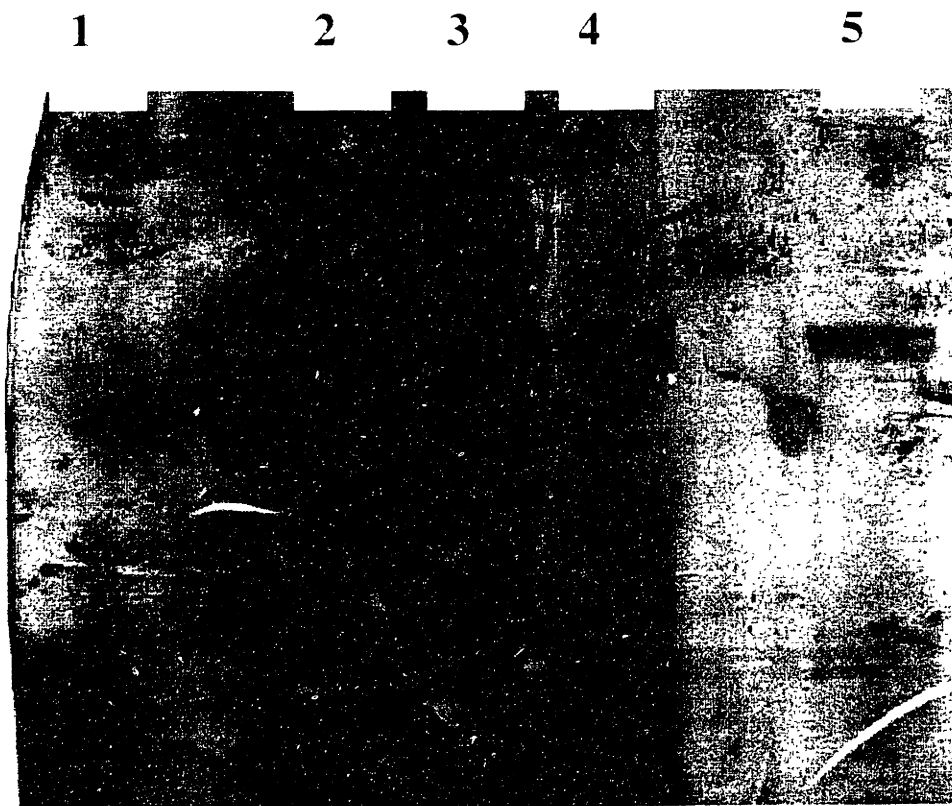
Figure 5.5 Purification of acetyl-CoA-carboxylase as seen on SDS-PAGE gel (silver staining)
Lane 1 protein standards
Lane 2 crude extract
Lane 3 DEAE-Sephacel flow through
Lane 4 PD-10 flow through
Lane 5 mono Q flow through

I.D.: 87132

Investigator: KOPPAS
Laboratory: SHOCKEY
Molec. Weight: 80 kDa
Sample solution: 25mM Bis-Tris, ~100mM NaCl, 1mM DTT

Figure 5.6 N-terminal sequence of the purified 80 kDa biotinylated enzyme

169

```
S. cerevisiae           MNVLVYNGPGTTPGSVKHAVESLRDFLEPYYAVSTVNVKVLQTEPWMSKT
S. pombe                MNVLIYNGNGASKICLLRTFQSLLPFVVPLYAMRFVDASTLEKEPWPAST
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  --------------------------------------------------
H. sapiens              MEDRLHMDNGLVPQKIVSVHLQDSTLKEVKDQVSNKQAQILEPKPEPSLE
B. subtilis             --------------------------------------------------
B. pertussis            --------------------------------------------------
A. thaliana             --------------------------------------------------
M. tuberculosis         --------------------------------------------------


S. cerevisiae           SAVVFPGGADLPYVQACQPIISRLKHFVSKQGGVFIGFCAGGYFGTSRVE
S. pombe                ALLVMPGGRDMGYCSSFNETIYRKITDFVKRGGAYLGLCAGGYFGSAVVD
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  --------------------------------------------------
H. sapiens              IKPEQDGMEHVGRDDPKALGEEPKQRRGSASGSEPAGDSDRGGGPVEHYH
B. subtilis             --------------------------------------------------
B. pertussis            --------------------------------------------------
A. thaliana             -----------------------------------MEAVRSTTTLSN
M. tuberculosis         --------------------------------------------------


S. cerevisiae           FAQG----DPTMEVSGSRDLRFFPGTSRGPAYNGFQYNSEAGARAVKLNL
S. pombe                FRMP----DSDLNVVGKRKLQFFPGTCAGPTFPGFTYDSEDGARRASIIV
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  --------------------------------------------------
H. sapiens              LHLSSCHECLELENSTIESVKFASAENIPDLPYDYSSSLESVADETSPEP
B. subtilis             --M--------RSTLRKDLIELFS------QAGSEFISGQKIS------
B. pertussis            --------------------------------------------------
A. thaliana             PHL--------LNILVLRSLKPLH-------RLSFSFSASAMES------
M. tuberculosis         --------------------------------------------------


S. cerevisiae           PD--GS-----QFSTYFNGGAVFVDADKFDNVEILATYAEHPDVPSSDSG
S. pombe                DGMQSSPV---HTHIYFNGGGSFLETENYSNVKVVARYQET---------
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  --------------------------------------------------
H. sapiens              EGRRVNLTGKAPNILLYVGSDSQEALGRFHEVRSVLADCVDIDSYILYHL
B. subtilis             --------------------------------------------------
B. pertussis            --------------------------------------------------
A. thaliana             --------------------------------------------------
M. tuberculosis         --------------------------------------------------


S. cerevisiae           KGQSEN----PAAVVLCTVGRGKVLLTGPHPEFNVRFMRKSTDK------
S. pombe                --DFEK----SAAIIYVKVGKGNVVLTGIHPEFSAEGSPILDKR------
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  ---------------------------MVEDYDPCEV------------
H. sapiens              LEDSALRDPWTDNCLLLVIATRESIPEDLYQKFMAYLSQGGKVLGLSSSF
B. subtilis             ----------DALGCSRTAVWKHIEELRKEGYEVEAVRRKGYR------
B. pertussis            --------------MGAITIAPMSAEVRSLSLPEPAALAR---------
A. thaliana             -----------DASCSLVLCGKSSVETEVAKGLKNKNSLKLPDN------
M. tuberculosis         --------------------------MTDFDRLRPPLDERSLRDQ---------


S. cerevisiae           -----HFLETVVENLKAQEIMRLKFMRTVLTKTGLNCNNDFNYVRAPNLT
S. pombe                -----DEKTRLELLSYILKLLGLKVPKDTSKRGQPTLTDQY---------
R. capsulatus           --------------------------------------------------
P. denitrificans        --------------------------------------------------
M. thermoautotrophicum  --------------------------------------------------
H. sapiens              TFGGFQVTSKGALHKTVQNLVFSKADQSEVKLSVLSSGCRYREGPVRLSP
B. subtilis             -----LI-------------------------------------------
B. pertussis            --------------------------------------------------
A. thaliana             -----TK-------------------------------------------
M. tuberculosis         --------------------------------------------------
```

```
                          PLFMASAPNKRNYLQEMENNLAHHGMHANNVELCSEL---------------
S. cerevisiae             PLFMASAPNKRNYLQEMENNLAHHGMHANNVELCSEL---------------
S. pombe                  -LFPNNVETKRFIEKALTNKVKNQ------------D-------------
R. capsulatus             --------------------------------------------------
P. denitrificans          --------------------------------------------------
M thermoautotrophicum     --------------------------------------------------
H. sapiens                GRLQGHLENEDKDRMIVHVPFGTRGGEAVLCQVHLELPPSSNIVQTPEDF
B subtilis                -----------------KKPGKLSESEIR------------F--------
B pertussis               -------------------------A-----------I------------
A. thaliana               ------------VSLILESEAKNL------------V-------------
M. tuberculosis           -------------------------------------L------------
```

```
S cerevisiae              -NAETDSFQFYRGYRASYDAA------------------------------
S. pombe                  -EDTLYTFQFSDISSEIPEHQ------------------------------
R. capsulatus             --------MARHVLGQTDT--------------------------------
P. denitrificans          -------------MLARTDS-------------------------------
M. thermoautotrophicum    -EVDTEYIGHEIKCFDEVDS------------------------------
H sapiens                 NLLKSSNFRRYEVLREILTTLGLSCDMKQVPALTPLYLLSAAEEIRDPLM
B subtilis                -GLKTEVMGQHLIYQDVISS------------------------------
B pertussis               -AARLPAFRDISWVGSTGST------------------------------
A thaliana                -KDDDNSFNLSLFMNSIITHR------------------------------
M tuberculosis            -IGAGSGWRQLDVVAQTGST------------------------------
```

```
S cerevisiae              --SSSLLHKEPDEVPKTVIFPGVDEDIPPFQYTPNFDMKEY---------
S. pombe                  --LANLDISADLSDSDNEIVKIWYGDEEKICKKAKPSFDLE---------
R capsulatus              ---------------TMAEAARLAPH-------------------------
P denitrificans           ---------------TNAEALKLAPG-------------------------
M thermoautotrophicum     ---------------TNNVAKRMAEDG------------------------
H sapiens                 QWLGKHVDSEGEIKSGQLSLRFVSSYVSEVEITPSCIPVVTNMEAFSSEH
B subtilis                ---------------TQKTAHELANNN-------------------------
B pertussis               --------------NADLLARARAGG--------------------------
A thaliana                --FGRPLIWSPRLSSTHDVVSHNFSE-------------------------
M tuberculosis            --------------NADLLARAASGA-------------------------
```

```
S cerevisiae              ---FKYLNVQNTIGSLLLYGEVVTSTSTILNNNKSLLSSIPESTLLHVGT
S pombe                   ---LYAKLINGCRFGLPIIVAPVIRSTQTLLDKNYRFLDSTNTGFTVLGN
R capsulatus              ------------------------------------------LAGPAWVFAH
P denitrificans           ------------------------------------------LSGSAWVLAR
M thermoautotrophicum     ------------------------------------------ATEGTVVIAK
H sapiens                 FNLEIYRQNLQTKQLGKVILFAEVTPTTMRLLDGLMFQTPQEMGLIVIAA
B subtilis                ------------------------------------------APEGTLVVAD
B pertussis               ------------------------------------------GARPCLLGAH
A thaliana                ------------------------------------------LPVGSVCVTD
M. tuberculosis           ------------------------------------------DIDGVVLIAE
```

```
S cerevisiae              I VS R GGNT INPKGVCASTAVVTMPLQSPVTNRNISVVFVQYLSML
S pombe                   Y TA R GQNM VSPYG-TLAFSFIINVDAK--NFSTTPIALFQYLMAL
R capsulatus              R TA R RGRD LMPSG-NFAATLVFRP-----AVTPATAALYSFVAAL
P denitrificans           E FA R RGRE VMPAG-NFAGTLVLRP-----QGGALAAAQLSFVAAL
M thermoautotrophicum     T SR R RGKP ISPQG-GIWMSIILRP-----EVHPSRAPLLTLVAGV
H sapiens                 R TE K GGNV LSPVGCALSTLLISIPLR---SQLGQRIPFVQHLMSV
B subtilis                K TA R MSRV HSQEGNGIWMSLILRP-----DIPLQKTPQLTLLAAV
B. pertussis              L ET R AGRP QNRVGAALMFSCAFDV-----DLPPSQLPALSPLAGM
A thaliana                I FK R TKNV ESPKG-CLMYSFTLEM------EDGRVVPLIQYVVSL
M tuberculosis            H TA R HGRG AATARAQIILSVGVRVV----DVPVQAWGWLSLAAGL
```

```
S cerevisiae              YCKAILSYAPG--FSDIPVRI        LYRLS?TYYKRKNLKLVNTGFEH
S pombe                   VVRGIREYAPG--YENIPAFI        IYVPVDKGG----------------
R capsulatus              LEAALTEVMGP---G-PVTAI        VLLNG--------------------
P denitrificans           LYDALGLACGP---A-ARLAI        VLLNG--------------------
M thermoautotrophicum     VARTLRDECG------LDVGI        ILIGD--------------------
H. sapiens                VVEAVRSIPEY---QDINLRV        IYYSD--------------------
B subtilis                VVQGIEAAAG------IQTDI        ILING--------------------
B. pertussis              ACEALRQLVGQ---P-AGLLV        VQWHD--------------------
A thaliana                VTEAVKDVCDKKGLPYIDVKI        LYVNG--------------------
M tuberculosis            VLDSVAPLIAV---PPAETGL        VLARG--------------------
```

```
S cerevisiae              TKLPLGDIEPAYL IS LLVNTH--FINNKYCLLLGC INLTS-DGPTTS
S pombe                   ----INFQGKQYM LS IIVTSN--YRKNVLHLVVGC INVSN-LGPTVS
R capsulatus              -----------G IA ILLESV-GTGGQISHLAVGI VNLAK-APPAAA
P denitrificans           -----------G VA ILLESS-GSGPGVQAVAVGI VNLAG-APDAGA
M thermoautotrophicum     -----------K VC ILTEAH-ARFNTLEYVVVGV IDTN--VDISHF
H sapiens                 ----------LM IQ VLVNST--LMGETFYILIGC FNVTN-SNPTIC
B subtilis                -----------K TV ILTEMQ-AEEDRVRSVIIGI INVN--QQSDDF
B pertussis               -----------A LA VLVESVRNPAGTGHAVVIGM LNLHGGDALSQA
A thaliana                -----------L VG ILCTST----YRSKKFNVSV VGLN--VDNGQP
M tuberculosis            -----------G LA ILAEVA-----QPFVVLGV LNVT--QAP-EE
```

```
S cerevisiae        LQTWIDILNEERQQLHLDLLPAIKAEKLQALYMNNLEVILKQFINYGAAE
S pombe             LNTLVDEWNKNSDNPRLEKF---SFEKLLASVLNQFDRYHRLLLEEGFSL
R capsulatus        LEPGAVAPVSLLGETGLV-I---DPLDPLAPLARGFAHYAATFERYGFAP
P denitrificans     VEPGATPPVSVQGETGHA-V---DPEEFLDLLAPAFARWQAQLDTYGFAP
M thermoautotrophicum PDDLREGATSIKNELKRD-I---KSSELIARFLRNFEETYNJFKDGKMDE
H sapiens           INDLITEYNKQHKAELKP----LRADYLIARVVTVLEKLIKEFQDKGPNS
B. subtilis         PDELKDIATSLSQASGEK-I---DRAGVIQHILLCPEKRYRDYMTHGFTP
B pertussis         LQRSVADWNDVAGPGEKGPD---AAADLVAAVARAWQDAAHTLQQEGFGA
A thaliana          TTCLNAVLKGMAPESNLL-----KREEILGAFFHKFEKFFDLFMDQGFKS
M tuberculosis      VDPDATSLLDLGVAAPDR------NRIASRLLRELEARIIQWRN-ANPQ


S cerevisiae        ILPSYYELWLHSNQIVTLPDHGNTQAMITGITEDYG-LLIAKE█VSGSST
S pombe             ILPEYYQYWLHSNQTVNLASGGK--AIIQGITSDFG--FLLAQ█LNENNE
R capsulatus        IRAAWLARAARIGQRITARTGT---ETLTGTFETID---DTGA█ILATGS
P denitrificans     IRNAWLARAARLGEPIIARTGT---AESHGIFEGID---DSGA█ILRGPA
M thermoautotrophicum ILTEWRKLSKTIGRRVEIRKQLG--EIVHGDAVGIN---SEGA█ILELDD
H sapiens           VLPLYYRYWVHSCQQVHLGSAE---GPKV-SIVGLD---DSGP█QVHQEG
B subtilis          IKLLWESYALGIGTNMRARTLN---GTFYGKALGID---DEGV█LLETQE
B pertussis         FGARFARVDALAGRPVNVIDRGE--ILYTGAASGVD---DHGR█LVQAAD
A thaliana          LEELYYRTWLHSEQRVIVEDKVE--DQVVQNVVTIQGLTSSGY█LAVGDD
M tuberculosis      LAADYRARSLTIGSRVRVELPGG--QDVVGIARDID---DQGR█CLDVGG


S cerevisiae        QPTGNVYNLQPDGNTFDIFKSLIAKKVQS
S pombe             PTT-KVVHLQPDGNSFDLMRNLITRKT--
R capsulatus        E-----RRAISAADIFF------------
P denitrificans     G-----RQVIPAAEVFFGG----------
M thermoautotrophicum GT---LRKVISGECIHR------------
H sapiens           GE---VVTVHPDGNSFDMLRNLILPYRR-
B subtilis          G----IKKIYSADIELG-----------
B pertussis         G----MRPVSVGEISIRPQA---------
A thaliana          NQ---MYELHPDGNSFDFFKGLVRRKI--
M tuberculosis      R-----TVVVSAGDVVHLR----------
```

CHAPTER 6: CONCLUSIONS AND RECOMMENDATIONS

6.A Summary of Thesis Results

The central focus of this thesis has been the elucidation and understanding of the metabolic network that includes central carbon metabolism and lysine production pathway in *C. glutamicum*. In the area of metabolic engineering, it is of critical importance to have a complete view of metabolic pathways before the optimization of the metabolism can take place. This was the motivation behind this thesis. A combination of molecular biology, bacterial batch cultures and enzymatic properties produced data that allowed us to develop a metabolic model in order to understand how carbon flows through metabolic pathways. This model was then verified by applying the same experimental tools and led to a lysine production improvement.

More specifically, this thesis was initiated by attempting to obtain a clear enzymatic evidence for the presence of a pyruvate carboxylating enzyme in *C. glutamicum*. Even though there was phenotypic and physiological evidence for the presence of such an enzyme in this strain, previous experimental attempts had failed to provide unequivocal *in vitro* enzymatic activity. In this study, such an activity is reported in *C. glutamicum* protein crude extracts. Contrary to our expectations though it is present at very low levels and more significantly it is uncoupled from the presence or absence of ATP. It is very likely that gluconeogenic enzymes such as oxaloacetate decarboxylase and malic enzyme that do not require ATP to function contribute to the anaplerotic needs of *C. glutamicum*. Still, the need for much higher enzymatic activities in order to explain lysine productivities observed led the experimental efforts towards the identification of other glycolytic pyruvate carboxylases, such as pyruvate carboxylase.

An alternative route in order to establish the presence of this enzyme was the identification of its gene in the *C. glutamicum* chromosome. The fact that pyruvate carboxylase, a very common anaplerotic enzyme, had been sequenced from a variety of organisms, the most significant being *M. tuberculosis*, enabled the identification of

173

various homologous regions within the amino acid sequence. Those regions were utilized to design PCR primers that eventually provided a piece of the gene of pyruvate carboxylase. The complete sequence was later obtained by identifying cosmids from a *C. glutamicum* cosmid library that contain the complete sequence and by directly sequencing those cosmids. The identification of the gene of pyruvate carboxylase provides for the first time unequivocal evidence that such carboxylation indeed takes place in *C. glutamicum* in a glycolytic direction. More important, the completion of the metabolic map around pyruvate proves the complexity of the regulation of this node and makes its genetic optimization the next experimental challenge.

The overexpression of the gene of pyruvate carboxylase proved to be a challenge more difficult than what it was anticipated. It was achieved only with a piece of DNA of specific length that includes the gene of this enzyme as well as certain upstream and downstream regions. What was even more significant was the physiological effect of such an overexpression: in terms of growth the recombinant strain grew better and in terms of production, contrary to the expectations. lysine produced was either equal or less than that of the control strain. It was therefore concluded that indeed pyruvate carboxylase overexpression increases the carbon flux into the TCA cycle. However. due to the presence of a metabolic bottleneck in the lysine biosynthesis pathway, this carbon flux increase cannot translate into increase of product formation. Instead, carbon stays in the TCA cycle or exits and converts into other products necessary for growth and that translates into better growth.

In order to prove this metabolic model, a potential metabolic bottleneck that was targeted was aspartate kinase. Its overexpression yielded a recombinant strain with a higher lysine productivity but lower growth rate. Obviously the increase of aspartate kinase activity translates into higher flux towards lysine production. However this increase deprives carbon from the TCA cycle and that results in the decrease in growth rate observed. The combination of overexpression of pyruvate carboxylase and aspartate kinase thus became the next obvious step. The results where indeed the ones predicted: a strain with similar growth rates as the control but higher lysine productivity. This is one

of the very few succesfull employments of metabolic engineering that results in an increase of the production of a primary metabolite such as lysine.

In a separate study, two other enzymes were investigated namely acetyl-CoA-carboxylase and biotin ligase. The purpose of this study was to complete the investigation of the role of biotin in amino acid production. The biotinylated subunit of acetyl-CoA-carboxylase was purified to completion and its N-terminal and internal amino acid sequences were obtained. Using complementation as a screening method of a biotin ligase deficient *E. coli* strain, a cosmid containing the gene of biotin ligase was also identified.

The contribution of this thesis lies in the integration of molecular biology, enzymology and physiological data obtained from batch cultures in order to address metabolic engineering challenges. The results of these three areas have been successfully integrated in order to predict metabolic targets and in order to finally achieve strain optimization, the final goal of every metabolic engineering question. The approach presented in this thesis provides a useful framework for a holistic study of metabolism and most important presents one of the very few examples of succesfull metabolic pathway manipulation.

6.B Conclusions

The following major conclusions can be drawn from this work:

1. Pyruvate carboxylase exists in *C. glutamicum*. The enzyme is very unstable and does not show significant *in vitro* activity in protein crude extract.

2. Pyruvate carboxylase overexpression in *C. glutamicum* cannot be achieved by simply cloning the gene of this enzyme in a *Corynebacterium* vector. The presence of upstream and/or downstream regions is necessary in order for this to be achieved.

3. Contrary to expectations the overexpression of pyruvate carboxylase results in an increase in growth but has no effect or even decreases the production of lysine.

4. The overexpression of aspartate kinase results in an increase in lysine production with a simultaneous decrease in growth.

5. Simultaneous overexpression of pyruvate carboxylase and aspartate kinase results has no effect on growth but results in a higher lysine specific productivity.

6. The 80 kDa biotinylated enzyme that exists in *C. glutamicum* corresponds to the biotinylated subunit of acetyl-CoA-carboxylase.

7. A cosmid has been isolated that most likely contains the gene of biotin ligase from *C. glutamicum*. This is based on complementation evidence of a biotin ligase deficient *E. coli* strain.

6.C Recommendations for Future Work

6.C.1 Investigation of other rate-limiting steps in the lysine pathway

One of the most significant findings of this thesis is the improvement of lysine productivity by performing a simultaneous overexpression of pyruvate carboxylase and aspartate kinase. This is one of the very few reports where an improvement of lysine production has been accomplished by applying molecular biology techniques in a lysine overproducing strain (Eggeling 1994) (Eggeling, De Graaf et al. 1996). This relates to the unique split pathway by which lysine is synthesized (Schrumpf, Schwarzer et al. 1991) (Sonntag, Eggeling et al. 1993) and the identification of three flux-controlling steps. Control is exerted at the entry of aspartate into the assembling pathway (aspartate kinase) and also at the branching point where aspartate semialdehyde is converted to either lysine or threonine (Cremer, Eggeling et al. 1991). Most exciting is the discovery that additional control is exerted by the export of L-lysine from the cell (Schrumpf, Eggeling et al. 1992) (Broer, Eggeling et al. 1993), and that a new type of carrier protein is present in *C*.

glutamicum which serves as a valve to regulate the intracellular lysine concentration (Vrljic, Sahm et al. 1996).

As already mentioned extensively previously, overexpression of aspartate kinase has an effect on lysine productivity. The second enzyme that also has been studied extensively in the literature and proved to have a significant effect on lysine yield was dihydrodipicolinate synthase (dapA). Overexpression of this enzyme and detailed flux studies revealed that, instead of simply tailoring the pathway by removing a bottleneck, the increased enzyme activity has far-reaching consequences with favourable global effects on the carbon flux within the aspartate family of amino acids and even within the central metabolism.

An oversynthesis of dihydrodipicolinate synthase results in an increase of lysine accumulation from 220 mM to 270 mM. The synthase, encoded by *dapA*, is located at the branch point of metabolite distribution to either lysine or threonine and competes with homoserine dehydrogenase for the common substrate aspartate semialdehyde. When graded *dapA* expression was used, as well as quantification of enzyme activities, intracellular metabolite concentrations and flux rates, it became apparent that the global response of the metabolic system to the synthase overactivity was double: an increased flux towards lysine production was accompanied by a decreased flux towards threonine. This resulted in a decreased growth rate, but more surprising was an increase of intracellular levels of pyruvate-derived valine and alanine. Thus the synthase serves as a barrier within lysine synthesis, and the flux through this low-affinity enzyme is proportional to its total amount. However, the improvement does not only consist in the removal of a bottleneck, but is due to a subtle flux redistribution at the dehydrogenase/synthase branch point. Combined with this redistribution is an introduced growth limitation, which results in increased availabilities of metabolites within the central metabolism (Eggeling, Oberle et al. 1998).

Based on all these data, a natural question that arises is what will happen to the flux distribution if a pyruvate consuming enzyme, such as pyruvate carboxylase, is simultaneously overexpressed with dihydrodipicolinate synthase. Is it possible that this double oversynthesis of the two enzymes can lead to an even higher increase in carbon

flux towards lysine production? Based on the data obtained from the overexpression of the other potential bottleneck, aspartate kinase, most likely this double overexpression will lead to the removal of growth limitation introduced with the overexpression of the synthase. It is obvious that the removal of the rate limiting step at the aspartate semialdehyde branch point leads to the introduction of a new one at the pyruvate branch point. This new rate limiting step translates into higher alanine and valine intracellular levels and it can only be due to low pyruvate carboxylase activity. Furthermore, the increase of carbon flux towards lysine again depletes the TCA cycle from carbon, as again it was the case with apsartate kinase, and that is why a lower growth rate is observed for the recombinant strain. It is therefore reasonable to expect that a bicistronic recombinant strain with both pyruvate carboxylase and dihydrodipicolinate synthase overexpressing, will again show similar growth profiles and higher lysine productivity comparing to the control. Of course the level of improvement of the parental strain comparing to that of the bicistronic pyruvate carboxylase- apartate kinase strain cannot be assessed at this point. It would be however useful as a last step to explore an even further possibility for strain improvement. That would consist of a simultaneous overexpression of all three previous targets, i.e. pyruvate carboxylase, aspartate kinase and dihydrodipicolinate synthase.

Overall, it appears that pyruvate carboxylase overexpression offers a new alternative to intracellularly introduced growth limitations as a way to improve product formation. It appears that this enzyme can compansate for the carbon loss from the TCA cycle that the removal of rate limiting steps within the lysine biosynthesis pathway causes. Its overexpression is therefore proposed to be used as an attractive means to create recombinant strains with growth properties similar to those of the parental strains, whenever growth appears to be a necessary step towards strain optimization.

6.C.2 Identification of the gene of biotin ligase

As mentioned in chapter 5, despite extensive experimental effort, in this work we have only been able to report the isolation of a cosmid that most likely contains the gene

of biotin ligase from *C. glutamicum*. The identification of this cosmid was based on complementation tests run with the use of an *E. coli* mutant strain that is biotin ligase deficient. This cosmid was able to complement both phenotypes of this mutant strain i.e. its inability to grow at 42 °C as well as its inability to grow in M9 glucose minimal medium without the presence of significant amounts of biotin. However, due to difficulty in obtaining clear DNA fragments following restriction enzyme digests, it has not been possible to target the presence of this gene in a fraction of this cosmid. The continuation of restriction digests is therefore proposed with a variety of different restriction enzymes complemented by efforts to measure a biotin ligase enzymatic activity in the mutant *E. coli* strain that harbors this cosmid. Such an enzymatic assay has been described extensively previously in the literature (Eisenberg, Prakash et al. 1982).

On a different level, even if the native gene of biotin ligase from *C. glutamicum* is not isolated an attempt can be made to express various heterologous biotin ligase genes in this organism from other related sources, as for example *M. tuberculosis*, *E. coli*, *R. capsulatus* etc. There are two major issues that such an overexpression can address. The first has to do with the activity levels of the biotinylated enzymes in *C. glutamicum*, i.e. pyruvate carboxylase and acetyl-CoA-carboxylase. Is it possible that biotinylation is a limiting step in obtaining a fully functional pyruvate carboxylase? If that would be a case, the physiological effects of pyruvate carboxylase (better growth and decrease in lysine production) should be seen enhanced in a recombinant strain that overexpresses pyruvate carboxylase and biotin ligase. Also the enzymatic activity of pyruvate carboxylase should be measured higher in such a strain. The second reason has to do with the ability to exercise control over the biotinylation process overall. Ideally, we would like to be able to direct the flow of extracellular biotin to either pyruvate carboxylase or acetyl-CoA-carboxylase, depending whether the focus is on growth or product formation. The cloning of the gene of biotin ligase under a tunable promoter, such as *tac*, would be able to provide us with such a system. That of course would require that the gene of the enzyme whose activity we desire to enhance (pyruvate carboxylase or acetyl-CoA-carboxylase) is also under the same promoter inside the chromosome of *C. glutamicum*.

In the area of acetyl-CoA-carboxylase, it needs to be pointed out that even though the gene of the biotinylated subunit of this enzyme has been sequenced, other possible subunits that need to be present have not been identified. Even for the biotinylated subunit itself that has been sequenced, there is no enzymatic proof that this protein derives from acteyl-CoA-carboxylase of *C. glutamicum*. The only evidence that points to that direction come from homology allignements and database search. In order to further elucidate this topic, two research directions are proposed.

The first one deals with the identification of the gene(s) of other subunits that consitute, together with the biotinylated subunit that has been purified in this work, the holoenzyme of acetyl-CoA-carboxylase. The identification of this subunit can be done either enzymatically- protein purification as performed in this work, with the requirement of measuring enzymatic activities between the intermediary steps- or by molecular biology techniques.

The second one relates with the enzymatic activity of acetyl-CoA-carboxylase at various stages during cell growth and especially during amino acid production. One of the major contributions of such a study would be a correlation between amout of glutamate or lysine excreted with the activity of this enzyme. That would provide the necessary experimental evidence for the indirect correlation proposed by various researchers between the activity of acetyl-CoA-carboxylase, the phospholipid composition of cell membrane and the amount of amino acids diffused from the cell into the extracellular medium.

As a final experimental step, an attempt to disrupt and then reintroduce the gene(s) that compose acetyl-CoA-carboxylase would further facilitate the studies on the physiological role of this enzyme in amino acid production. Such attempts have been made in the past. without any conclusive evidence for their success.

6.C.3 Biotin biosynthetic pathway

Unlike other actinomyces, such as *M. tuberculosis*. *C. glutamicum* is a biotin auxotroph. The most amazing aspect in this subjcet is the fact that *Corynebacteria* appear

to have mos of the enzymes required for biotin biosynthesis from pimelic acid, as it has been reported in the literature. In fact perhaps the only gene still required for converting *C. glutamicum* into a biotin prototroph strain is *bioF*. It would de advantageous to allow *Corynebacterium* to be biotin self sufficient since large amounts of D-biotin are required during the fermentation production of amino acids with this bacterium. Furthemore this strain with its high glucose utilization capability and ability to excrete many types of metabolites may be a suitable organism for the fermentative production of D-biotin.

The introduction of the necessary gene(s) for converting of *C. glutamicum* into a biotin prototroph should be done directly into the chromosome of the organism. As donors of such genes various sources can be used, such as *M. tuberculosis*, *E. coli* and *B. sphaericus*. These genes have a history of good expression levels when cloned into other organisms and in the case of *E. coli* and *B. sphaericus* their biotin biosynthesis operons have been studied extensively.

The potential of using *Corynebacterium* as a biotin producer has been demonstrated by Ogino *et al.* [1974a, 1974b] (Ogino, Fujimoto et al. 1974a) (Ogino, Fujimoto et al. 1974b) who reported a biotin conversion using n-paraffin utilizing *Corynebacterium* species from DL-*cis*-tetrahydro-2-oxo-4-*n*-pentyl-thieno-(3,4-D)-imidazoline, which is a biotin analog having a methyl group instead of a carboxyl group in the biotin molecule. It is possible therefore for *C. glutamicum* to become a biotin producers, along with other strains that already have been studied such as *E. coli* (Brown, Speck et al. 1991), *B. sphaericus* (Ohsawa, Kisou et al. 1992) and *Serratia marcescens* (Sakurai, Imai et al. 1993).

6.C.4 Mathematical Modeling of Central Carbon Metabolism

With the discovery of pyruvate carboxylase and the elucidation of the kinetics of this enzyme, the mathematical model originally developed by Vallino to describe the carbon flow around the PEP branchpoint can be reexamined. The original model regarded phosphoenaolpyruvate carboxylase as the major anaplerotic enzyme and did not account for the flux distribution by the presence of pyruvate carboxylase, malic enzyme

and oxaloacetate decarboxylase. The development of mutant and recombinant *C. glutamicum* strains that lack or overexpress pyruvate carboxylase respectively would enable to uncover the regulation mechanisms surrounding the partitioning of fluxes into oxaloacetate. These types of information are hard to obtain experimentally and would demonstrate the utility and usefulness of the mathematical modeling approach.

CHAPTER 7:   REFERENCES

Abbott, J. and D. Beckett (1993). "Cooperative binding of the *Escherichia coli* repressor of biotin biosynthesis to the biotin operator sequence." Biochemistry **32**: 9649-9656.

Abe, S. K., K. Takayama, et al. (1967). "Taxonomical studies on glutamic-acid producing bacteria." Journal of General and Applied Microbiology **13**: 279-301.

Aida, K. (1973). "Control of Microbial Metabolism." The Microbial Production of Amino Acids **K. Yamada, S. Kinoshita, T. Tsunoba and K. Aida, Ed.**(John Wiley & Sons: Tokyo): 91-121.

Alberts, A. and P. R. Vagelos (1972). "Acyl CoA carboxylase." The Enzymes **P. D. Boyer Ed.**(Academic Press: New York 3rd ed.): 37-82.

Alberts, A. W., A. M. Nervi, et al. (1969). "Acetyl CoA carboxylase II. Demonstration of biotin-protein and biotin-carboxylase subunits." Proceedings of the National Academy of Sciences USA **63**: 1319-1332.

Al-Feel, W., S. S. Chirala, et al. (1992). "Cloning of the yeast FAS3 gene and primary structure of yeast acetyl-CoA carboxylase." Proc Natl Acad Sci U S A **89**(10): 4534-8.

Al-ssum, R. M. and P. J. White (1977). "Activities of anaplerotic enzymes and acetyl coenzyme A carboxylase in biotin-deficient *Bacillus megaterium*." Journal of General Microbiology **100**: 203-209.

Altman, J. D., P. A. H. Moss, et al. (1996). "Phenotypic analysis of antigen-specific T lymphocytes [published erratum appears in Science 1998 Jun 19;280(5371):1821]." Science **274**(5284): 94-6.

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol 215(3): 403-10.

Archer, J. A. and A. J. Sinskey (1993). "The DNA sequence and minimal replicon of the *Corynebacterium glutamicum* plasmid pSR1: evidence of a common ancestry with plasmids from C. diphtheriae." J Gen Microbiol 139(Pt 8): 1753-9.

Archer, J. A. C., M. T. Follettie, et al. (1989). "Biology of *Corynebacterium glutamicum*: a molecular approach." C.L. Hershberger, S.W. Queener and G. Hegeman (ed.) Genetcis and molecular biology of industrial microorganisms **American Society for Microbiology, Washington DC**: 27-33.

Ashworth, J. M., H. L. Kornberg, et al. (1965). "The role of phosphoenolpyruvate carboxylase in *Escherichia coli*." Biochemistry Journal 94: 28P.

Atkinson, D. E. (1977). Cellular Energy Metabolism and its regulation. New York.

Attwood, P. V. (1995). "The structure and the mechanism of action of pyruvate carboxylase." Int J Biochem Cell Biol 27(3): 231-49.

Ballard, F. J., R. W. Hanson, et al. (1970). "Immunochemical studies with soluble and mitochondrial pyruvate carboxylase activities from rat tissues." Biochem J 119(4): 735-42.

Baiker, D. F. and A. M. Campbell (1980). "Use of bio-lac fusion strains to study regulation of biotin biosynthesis in *Escherichia coli*." J Bacteriol 143(2): 789-800.

Barksdale, L. (1970). "*Corynebacterium diphtheriae* and its relatives." Bacteriology Review 34: 378-422.

Bathe, B., J. Kalinowski, et al. (1996). "A physical and genetic map of the *Corynebacterium glutamicum* ATCC 13032 chromosome." Mol Gen Genet **252**(3): 255-65.

Bondaryk, R. P. and H. Paulus (1985). "Expression of the gene for *Bacillus subtilis* aspartokinase II in *Escherichia coli*." J Biol Chem **260**(1): 592-7.

Bouzier, A. K., P. Voisin, et al. (1998). "Glucose and lactate metabolism in C6 glioma cells: evidence for the preferential utilization of lactate for cell oxidative metabolism." Dev Neurosci **20**(4-5): 331-8.

Bower, S., J. Perkins, et al. (1995). "Cloning and characterization of the *Bacillus subtilis* birA gene encoding a repressor of the biotin operon." J Bacteriol **177**(9): 2572-5.

Brabetz, W., W. Liebl, et al. (1991). "Studies on the utilization of lactose by *Corynebacterium glutamicum*, bearing the lactose operon of *Escherichia coli*." Arch Microbiol **155**(6): 607-12.

Braud, V. M., D. S. Allan, et al. (1998). "HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C [see comments]." Nature **391**(6669): 795-9.

Brewster, N. K., D. L. Val, et al. (1994). "Regulation of pyruvate carboxylase isozyme (PYC1, PYC2) gene expression in *Saccharomyces cerevisiae* during fermentative and nonfermentative growth." Arch Biochem Biophys **311**(1): 62-71.

Broer, S., L. Eggeling, et al. (1993). "Strains of *Corynebacterium glutamicum* with different lysine productivities may have different lysine excretion systems." Applied and Environmental Microbiology **59**: 316-321.

Brown, S. W., D. Speck, et al. (1991). "The production of biotin by recombinant strains of *Escherichia coli*." Journal of Chemical Technology and Biotechnology **50**: 115-121.

Buckley, J. J., S. Libor, et al. (1969). "Biotin subunits of acetyl CoA carboxylase and pyruvate carboxylase from a thermophilic *Bacillus*." Archives of Biochemistry and Biophysics **192**: 396-400.

Buoncristiani, M. R., P. K. Howard, et al. (1986). "DNA-binding and enzymatic domains of the bifunctional biotin operon repressor (BirA) of *Escherichia coli*." Gene **44**(2-3): 255-61.

Campbell, A., R. Chang, et al. (1980). "Biotin regulatory (bir) mutations of *Escherichia coli*." J Bacteriol **142**(3): 1025-8.

Cazzulo, J. J., T. K. Sundaram, et al. (1971). "Synthesis of pyruvate carboxylase from its apoenzyme and (+)-biotin in *Bacillus stearothermophilus*. Purification and properties of the apoenzyme and the holoenzyme synthetase." Biochem J **122**(5): 653-61.

Cazzulo, J. J., T. K. Sundaram, et al. (1969). "Regulation of pyruvate carboxylase formation from the apo-enzyme and biotin in a thermophilic bacillus." Nature **223**(211): 1137-8.

Cazzulo, J. J., T. K. Sundaram, et al. (1970). "Properties and regulation of pyruvate carboxylase from *Bacillus stearothermophilus*." Proc R Soc Lond B Biol Sci **176**(42): 1-19.

Chao, Y. P. and J. C. Liao (1993). "Alteration of growth yield by overexpression of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxykinase in *Escherichia coli*." Appl Environ Microbiol **59**(12): 4261-5.

Charles, A. M. and D. W. Willer (1984). "Pyruvate carboxylase from *Thiobacillus novellus*: properties and possible functions." Canadian Journal of Microbiology **30**: 532-539.

Cohen, G. N. and I. Saint- Geront (1987). Biosynthesis of threonine, lysine and methionine. Washington, D.C., American Society for Microbiology.

Cohen, P. and D. G. Hardie (1991). "The actions of cyclic AMP on biosynthetic processes are mediated indirectly by cyclic AMP-dependent protein kinase." Biochim Biophys Acta **1094**(3): 292-9.

Collins, M. D. and C. S. Cummins (1986). "Genus *Corynebacterium* Lehmann and Neumann 1896." Bergey's manual of systematic bacteriology **II**(P.H.A. Sneath, N.S. Mair, M.E. Sharpe and J.G. Holt (ed.)): 1266-1276.

Colon, G. E., M. T. Follettie, et al. (1993). "Redirections of carbon flux at a *Corynebacterium glutamicum* threonine metabolic branch point by controlled enzyme overexpression." Annual Meeting of the American Society of Microbiology **Atlanta, GA**,(Washington DC: American Society of Microbiology): 320.

Colon, G. E., M. S. Jetten, et al. (1995). "Effect of inducible thrB expression on amino acid production in *Corynebacterium lactofermentum* ATCC 21799." Appl Environ Microbiol **61**(1): 74-8.

Colon, G. E., T. T. Nguyen, et al. (1995). "Production of isoleucine by overexpression of ilvA in a *Corynebacterium lactofermentum* threonine producer." Appl Microbiol Biotechnol **43**(3): 482-8.

Cremer, J., L. Eggeling, et al. (1991). "Control of the lysine biosynthetic sequence in *Corynebacterium glutamicum* as analyzed by overexpression of the individual corresponding genes." Applied and Environmental Microbiology **57**: 1746-1752.

Cronan, J. E., Jr. (1989). "The E. coli bio operon: transcriptional repression by an essential protein modification enzyme." Cell **58**(3): 427-9.

Cronan, J. E., Jr. and J. C. Wallace (1995). "The gene encoding the biotin-apoprotein ligase of *Saccharomyces cerevisiae*." FEMS Microbiol Lett **130**(2-3): 221-9.

Cronan Jr., J. E. and C. O. Rock (1987). "Biosynthesis of Membrane Lipids." *Escherichia coli and Salmonella typhimurium* **F.C. Neidhardt Ed.**(American Society for Microbiology: Washington, DC.): 474-497.

Davies, J. (1990). "Interspecific gene transfer: where next?" Trends Biotechnol **8**(8): 198-203.

Deb, J. K. and N. Nath (1999). "Plasmids of corynebacteria." FEMS Microbiology Letters **175**: 11-20.

Dekleva, M. L. and W. R. Strohl (1988). "Activity of phosphoenolpyruvate carboxylase of an anthracycline-producing streptomycete." Can J Microbiol **34**(11): 1241-6.

Diesterhaft, M. D. and E. Freese (1973). "Role of pyruvate carboxylase, phosphoenolpyruvate carboxykinase, and malic enzyme during growth and sporulation of *Bacillus subtilis*." J Biol Chem **248**(17): 6062-70.

Dunican, L. K. and E. Shivnan (1989). "High frequency transformation of whole cells of amino acid producing coryneform bacteria using high voltage electroporation." Bio/Technology **7**: 1067-1070.

Dunn, M. F., S. Encarnacion, et al. (1996). "Pyruvate carboxylase from *Rhizobium etli*: mutant characterization, nucleotide sequence, and physiological role." J Bacteriol **178**(20): 5960-70.

Eggeling, L. (1994). "Biology of L-lysine overproduction by *Corynebacterium glutamicum*." Amino Acids **6**: 261-272.

Eggeling, L., A. A. De Graaf, et al. (1996). "Quantifying and directing metabolite flux: application to amino acid overproduction." Advances in Biochemical Engineering **54**: 2-30.

Eggeling, L., S. Oberle, et al. (1998). "Improved L-lysine yield with *Corynebacterium glutamicum*: use of dapA resulting in increased flux combined with growth limitation." Appl Microbiol Biotechnol **49**(1): 24-30.

Eggeling, L. and H. Sahm (1999). "L-Glutamate and L-Lysine: traditional products with impetous developments." Applied Microbiology and Biotechnology **52**: 146-153.

Eikmanns, B. J., L. Eggeling, et al. (1993). "Molecular aspects of lysine, threonine, and isoleucine biosynthesis in *Corynebacterium glutamicum*." Antonie Van Leeuwenhoek **64**(2): 145-63.

Eisenberg, M. A. (1984). "Regulation of the biotin operon." Annals of the New York Academy of Sciences **447**: 335-349.

Eisenberg, M. A., O. Prakash, et al. (1982). "Purification and properties of the biotin repressor. A bifunctional protein." J Biol Chem **257**(24): 15167-73.

Fall, R. R., A. W. Alberts, et al. (1975). "Analysis of bacterial biotin-proteins." Biochim Biophys Acta 379(2): 496-503.

Fazel, A., Y. Guillou, et al. (1983). "A hybrid proteolytic fragment of *Escherichia coli* aspartokinase I-homoserine dehydrogenase I. Structure, inhibition pattern, dissociation properties, and generation of two homodimers." J Biol Chem 258(22): 13570-4.

Fipula, D., A. H. Ally, et al. (1986). "Complete nucleotide sequence of a native plasmid of *Brevibacterium lactofermentum*." Nucleic Acids Research 14: 514-518.

Follettie, M. T., O. P. Peoples, et al. (1993). "Gene structure and expression of the *Corynebacterium flavum* N13 ask-asd operon." J Bacteriol 175(13): 4096-103.

Follettie, M. T., O. P. Peoples, et al. (1993). "Metabolic Engineering of Corynebacteria." Molecular Microbiology 5: 315-325.

Follettie, M. T., H. K. Shin, et al. (1988). "Organization and regulation of the *Corynebacterium glutamicum* hom-thrB and thrC loci." Mol Microbiol 2(1): 53-62.

Fry, D. C., S. A. Kuby, et al. (1986). "ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F1-ATPase, and other nucleotide-binding proteins." Proc Natl Acad Sci U S A 83(4): 907-11.

Gerwin, B. I., B. E. Jacobsen, et al. (1969). "Transcarboxylase VIII. Isolation and properties of a biotin-carboxyl carrier protein." Proceedings of the National Academy of Sciences USA 64: 1315-1318.

Gornicki, P., J. Podkowinski, et al. (1994). "Wheat acetyl-coenzyme A carboxylase: cDNA and protein structure." Proc Natl Acad Sci U S A 91(15): 6860-4.

190

Guan, X. and E. S. Wurtele (1996). "Reduction of growth and acetyl-CoA carboxylase activity by expression of a chimeric streptavidin gene in *Escherichia coli*." Appl Microbiol Biotechnol **44**(6): 753-8.

Gubler, M., M. Jetten, et al. (1994). "Cloning of the pyruvate kinase gene (pyk) of *Corynebacterium glutamicum* and site-specific inactivation of pyk in a lysine-producing *Corynebacterium lactofermentum* strain." Appl Environ Microbiol **60**(7): 2494-500.

Gubler, M., S. M. Park, et al. (1994). "Effects of phosphoenolpyruvate carboxylase deficiency on metabolism and lysine production in *Corynebacterium glutamicum*." Applied Microbiology and Biotechnology **40**: 857-863.

Gubler, M. and A. Sinskey (1993). Influence of transcription rate on protein secretion in *Corynebacterium glutamicum*. Washington, D.C., American Society for Mocrobiology.

Ha, J., S. Daniel, et al. (1994). "Cloning of human acetyl-CoA carboxylase cDNA." Eur J Biochem **219**(1-2): 297-306.

Hanahan, D. (1983). "Studies on transformation of *Escherichia coli* with plasmids." J Mol Biol **166**(4): 557-80.

Hartman, R. E. (1970). "Carbon dioxide fixation by extracts of *Streptococcus faecalis* var. *liquefaciens*." Journal of Bacteriology **102**: 341-347.

Hatakeyama, K., K. Kohama, et al. (1993). "Analysis of the biotin biosynthesis pathway in coryneform bacteria: cloning and sequencing of the bioB gene from *Brevibacterium flavum*." DNA Seq **4**(2): 87-93.

Hatakeyama, K., K. Kohama, et al. (1993b). "Genomic organization of the biotin biosynthetic genes of coryneform bacteria: cloning and sequencing of the *bioA-bioD* genes from *Brevibacterium flavum*." DNA Sequence **4**: 177-184.

Haynes, J. A. and M. L. Britz (1989). "Electrotransformation of *Brevibacterium lactofermentum* and *Corynebacterium glutamicum* : growth in Tween 20 increases transformation frequencies." FEMS Microbiology Letters **61**: 329-334.

Higa, A. I., S. R. Milrad de Forchetti, et al. (1976). "CO2-fixing enzymes in *Pseudomonas fluorescens*." J Gen Microbiol **93**(1): 69-74.

Hiller, A. J. and G. R. Jago (1978). "The metabolism of [^{14}C]bicarbonate by *Streptococcus lactis*: the fixation of [^{14}C]bicarbonate by pyruvate carboxylase." Journal of Diary Research **45**: 445-451.

Hodgson, A. L., J. Krywult, et al. (1990). "Nucleotide sequence of the erythromycin resistance gene from the *Corynebacterium* plasmid pNG2." Nucleic Acids Res **18**(7): 1891.

Hubbard, J. S. and A. N. Hall (1967). "Effects of biotin on glutamate production and fatty acid composition in *Bacillus cereus* 14B22." Canadian Journal of Microbiology **14**: 1039-1048.

Ingram, L. O., T. Conway, et al. (1987). "Genetic engineering of ethanol production in *Escherichia coli*." Appl Environ Microbiol **53**(10): 2420-5.

Ishino, S., K. Yamaguchi, et al. (1984). "Involvement of mesodiaminopimelate D-dehydrogenase in lysine biosynthesis in *Corynebacterium glutamicum*." Agricultural and Biological Chemistry **48**: 2557-2560.

Jackowski, S., J. E. Cronan Jr., et al. (1991). "Lipid Metabolism in Prokaryotes." Biochemistry of Lipids, Lipoproteins and Membranes **D. E. Vance and J. Vance Ed.**(Elsevier Science Publishing Co. : Amsterdam): 43-85.

Jacobsen, J. R. and C. Khosla (1998). "New directions in metabolic engineering." Current Opinion in Chemical Biology **21**: 133-137.

Jager, W., P. G. Peters-Wendisch, et al. (1996). "A *Corynebacterium glutamicum* gene encoding a two-domain protein similar to biotin carboxylases and biotin-carboxyl-carrier proteins." Arch Microbiol **166**(2): 76-82.

Jetten, M. S., M. T. Follettie, et al. (1995). "Effect of different levels of aspartokinase on the lysine production by *Corynebacterium lactofermentum.*" Appl Microbiol Biotechnol **43**(1): 76-82.

Jetten, M. S. and A. J. Sinskey (1995). "Recent advances in the physiology and genetics of amino acid-producing bacteria." Crit Rev Biotechnol **15**(1): 73-103.

Jetten, M. S. M., G. A. Pitoc, et al. (1994). "Regulation of phosphoenol-pyruvate- and oxaloacetate-converting enzymes in *Corynebacterium glutamicum.*" Applied Microbiology and Biotechnology **41**: 47-52.

Jitrapakdee, S., G. W. Booker, et al. (1996). "Cloning, sequencing and expression of rat liver pyruvate carboxylase." Biochem J **316**(Pt 2): 631-7.

Kalinowski, J., B. Bachmann, et al. (1990). "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspartate beta-semialdehyde dehydrogenase gene asd in *Corynebacterium glutamicum.*" Mol Gen Genet **224**(3): 317-24.

Kalinowski, J., J. Cremer, et al. (1991). "Genetic and biochemical analysis of the aspartokinase from *Corynebacterium glutamicum*." Mol Microbiol 5(5): 1197-204.

Karasawa, M., O. Tosaka, et al. (1986). "Application of protoplast fusion to the development of L-threonine and L-lysine producers." Agricultural and Biological Chemistry 50: 339-346.

Katsumata, R., A. Ozaki, et al. (1984). "Protoplast transformation of glutamate-producing bacteria with plasmid DNA." J Bacteriol 159(1): 306-11.

Kawahara, Y., Y. Yoshihara, et al. (1990). "Stimulatory effect of glycine betaine on L-lysine feremntation." Applied Microbiology and Biotechnology 34: 87-90.

Keech, D. B. and J. C. Wallace (1985). "Pyruvate Carboxylase." CRC Press Series in Enzyme Biology **Boca Raton, Florida**.

Keilhauer, C., L. Eggeling, et al. (1993). "Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the ilvB-ilvN-ilvC operon." J Bacteriol 175(17): 5595-603.

Kim, K.-H. (1983). "Regulation of Acetyl-CoA Carboxylase." Current Topics of Cell Regulation 20: 139-184.

Kinoshita, S., K. Nakayama, et al. (1958). "L- Lysine production using microbial auxotroph." Journal of General and Applied Microbiology 4: 128-129.

Kiss, R. D. and G. Stephanopoulos (1991). "Metabolic activity control of the L-lysine fermentation by restrained growth fed-batch strategies." Biotechnology Progress 7: 501-509.

Koffas, M., C. Roberge, et al. (1999). "Metabolic Engineering." Annual Reviews of Biomedical Engineering 1: 535-557.

Kondo, H., Y. Kazuta, et al. (1997). "Cloning and nucleotide sequence of *Bacillus stearothermophilus* pyruvate carboxylase." Gene 191(1): 47-50.

Kornberg, H. L. (1966). "Anaplerotic sequences and their role in metabolism." Essays in Biochemistry II: 1-31.

Kumar, V., S. Ramakrishnan, et al. (1992). "*Saccharomyces cerevisiae* cells secreting an *Aspergillus niger* beta-galactosidase grow on whey permeate." Biotechnology (N Y) 10(1): 82-5.

Kumer, G. K., F. C. Haase, et al. (1988). "Involvement and identification of a tryptophanyl residue at the pyruvate- binding site of transcarboxylase." Biochemistry 27: 5978-5983.

Lachica, V. F. and P. A. Hartman (1969). "Inhiibition of CO_2 fixation in group *D. streptococci*." Canadian Journal of Microbiology 15: 57-62.

Lauer, G., E. A. Rudd, et al. (1991). "Cloning, nucleotide sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia coli* DNA ligase." J Bacteriol 173(16): 5047-53.

Leon Del Rio, A. and R. A. Gravels (1994). "Sequence requirements for the biotinylation of carboxyl-terminal fragments of human propionyl-CoA carboxylase α subuntil expressed in *Escherichia coli*." Journal of Biological Chemistry 269: 22964-22968.

Leon-Del-Rio, A., D. Leclerc, et al. (1995). "Isolation of a cDNA encoding human holocarboxylase synthetase by functional complementation of a biotin auxotroph of *Escherichia coli.*" Proc Natl Acad Sci U S A **92**(10): 4626-30.

Li, S. J. and J. E. Cronan, Jr. (1993). "Growth rate regulation of *Escherichia coli* acetyl coenzyme A carboxylase, which catalyzes the first committed step of lipid biosynthesis." J Bacteriol **175**(2): 332-40.

Liao, C. L. and D. E. Atkinson (1971). "Regulation at the phosphoenolpyruvate branchpoint in *Azotobacter vinelandii*: pyruvate kinase." J Bacteriol **106**(1): 37-44.

Libor, S. M., T. K. Sundaram, et al. (1978). "Pyruvate carboxylase from thermophilic *Bacillus*. Studies on the speicificity of activation by acyl derivatives of coenzyme A and on the properties of catalysis in the absence of activator." Biochemistry Journal **169**: 543-550.

Liebl, W. (1992). "The Genus *Corynebacterium*- Nonmedical." The Prokaryotes, Second Edition **II**(Editors: Balows, A.; Trueper, H.G.; Dworkin, M.; Harder, W.; Schleifer, K.H.): 1157-1171.

Liebl, W., A. Bayerl, et al. (1989). "High efficiency electroporation of intact *Corynebacterium glutamicum* cells." FEMS Microbiol Lett **53**(3): 299-303.

Lim, F., C. P. Morris, et al. (1988). "Sequence and domain structure of yeast pyruvate carboxylase." J Biol Chem **263**(23): 11493-7.

Lynen, F. (1974). "Structures of biotin enzymes." Lipman Symposium: Energy, Regulation and Biosynthesis in Molecular Biology **Richter, D., Ed., de Grutyer, Berlin**: 671.

MacQuitty, J. J. (1988). "Impact of biotechnology on the chemical industry." ACS Symposium Series **362**: 221-233.

Martin, J. F., R. Santamaria, et al. (1987). "Cloning systems in amino acid producing corynebacteria." Bio/Technology **5**: 137-146.

Mayer, F., J. C. Wallace, et al. (1980). "Further electron microscope studies on pyruvate carboxylase." Eur J Biochem **112**(2): 265-72.

Mazodier, P. and J. Davies (1991). "Gene transfer between distantly related bacteria." Annu Rev Genet **25**: 147-71.

Mazodier, P., R. Petter, et al. (1989). "Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species." J Bacteriol **171**(6): 3583-5.

McMichael, A. J. and C. A. O'Callaghan (1998). "A new look at T cells." J Exp Med **187**(9): 1367-71.

Menendez, J. and C. Gancedo (1998). "Regulatory regions in the promoters of the *Saccharomyces cerevisiae* PYC1 and PYC2 genes encoding isoenzymes of pyruvate carboxylase." FEMS Microbiol Lett **164**(2): 345-52.

Messerotti, L. J., A. J. Radford, et al. (1990). "Nucleotide sequence of the replication region from the *Mycobacterium-Escherichia coli* shuttle vector pEP2." Gene **96**(1): 147-8.

Miller, J. H. (1972). Experiments in molecular genetics. Cold Spring Harbor New York, Cold Spring Harbor Laboratory.

Milrad de Forchetti, S. R. and J. J. Cazzulo (1976). "Some properties of the pyruvate carboxylase from *Pseudomonas fluorescens*." <u>J Gen Microbiol</u> **93**(1): 75-81.

Miwa, K., K. Matsui, et al. (1985). "Construction of novel shuttle vectors and a cosmid vector for the glutamic acid-producing bacteria *Brevibacterium lactofermentum* and *Corynebacterium glutamicum*." <u>Gene</u> **39**(2-3): 281-6.

Modak, H. V. and D. J. Kelly (1995). "Acetyl-CoA-dependent pyruvate carboxylase from the photosynthetic bacterium *Rhodobacter capsulatus*: rapid and efficient purification using dye-ligand affinity chromatography." <u>Microbiology</u> **141**(Pt 10): 2619-28.

Morinaga, Y., M. Tsuchiya, et al. (1987). "Expression of *Escherichia coli* promoters in *Brevibacterium lactofermentum* using the shuttly vector pEB003." <u>Journal of Biotechnology</u> **5**: 305-312.

Mukhopadhyay, B., S. F. Stoddard, et al. (1998). "Purification, regulation, and molecular and biochemical characterization of pyruvate carboxylase from *Methanobacterium thermoautotrophicum* strain deltaH." <u>J Biol Chem</u> **273**(9): 5155-66.

Nakayama, K., H. Tanaka, et al. (1966). "Studies on lysine formation, part V. Concerted feedback inhibition of aspartokinase and the absence of lysine inhibition on aspartic semialdehyde-pyruvate condensation in *Micrococcus glutamicus*." <u>Agricultural and Biological Chemistry</u> **30**: 611-616.

Nerem, R. M. (1991). "Cellular Engineering." <u>Annual Biomedical Engineering</u> **19**: 529-545.

Obermayer, M. and F. Lynen (1976). "Structure of biotin enzymes." <u>Trends in Biochem. Sciences</u> **1**: 169.

O'Brien, R., D. T. Chuang, et al. (1977). "Novel enzymic machinery for the metabolism of oxalacetate, phosphoenolpyruvate, and pyruvate in *Pseudomonas citronellolis*." J Biol Chem **252**(4): 1257-63.

O'Callaghan C, A., M. F. Byford, et al. (1999). "BirA enzyme: production and application in the study of membrane receptor-ligand interactions by site-specific biotinylation." Anal Biochem **266**(1): 9-15.

Ogino, S., S. Fujimoto, et al. (1974a). "Cooxidation of *dl-cis*-tetrahydro-2-oxo-4-n-pentyl-thieno-(3,4-d)-imidazoline (dl-TOPTI) by soil isolates of the genus *Corynebacterium*." Agricultural and biological Chemistry **38**: 275-278.

Ogino, S., S. Fujimoto, et al. (1974b). "Production of biotin by microbial transformation of *dl-cis*-tetrahydro-2-oxo-4-n-pentyl-thieno-(3,4-d)-imidazoline (dl-TOPTI)." Agricultural and Biological Chemistry **38**: 707-712.

Ohsawa, I., T. Kisou, et al. (1992). "Bioconversion of pimelic acid into biotin by *Bacillus sphaericus bioB* transformants." Journal of Fermentation and Bioengineering **73**: 121-124.

Ohta, K., D. S. Beall, et al. (1991). "Genetic improvement of *Escherichia coli* for ethanol production: chromosomal integration of *Zymomonas mobilis* genes encoding pyruvate decarboxylase and alcohol dehydrogenase II." Appl Environ Microbiol **57**(4): 893-900.

Ohta, K., D. S. Beall, et al. (1991). "Metabolic engineering of *Klebsiella oxytoca* M5A1 for ethanol production from xylose and glucose." Appl Environ Microbiol **57**(10): 2810-5.

Oki, T. and K. Ogata (1968). "Size and morphology of *Brevibacterium* phages." Agricultural and Biological Chemistry **32**: 241-248.

Otsuka, A. and J. Abelson (1978). "The regulatory region of the biotin operon in *Escherichia coli*." Nature **276**(5689): 689-94.

Owttrim, G. W. and B. Coleman (1986). "Purification and characterization of phosphoenolpyruvate carboxylase from a cyanobacterium." Journal of Bacteriology **168**: 207-212.

Park, S. M. (1996). Investigation of Carbon Fluxes in Central Metabolic Pathways of *Corynebacterium glutamicum*. Chemical Engineering. Cambridge, Massachusetts Institute of Technology.

Park, S. M., C. Shaw-Reid, et al. (1997b). "Elucidation of anaplerotic pathways in *Corynebacterium glutamicum* via ^{13}C-NMR spectroscopy and GC-MS." Applied Microbiology and Biotechnology **47**: 430-440.

Park, S. M., A. Sinskey, et al. (1997a). "Metabolic and physiological studies of *Corynebacterium glutamicum* mutants." Biotechnology and Bioengineering **55**: 864-879.

Patek, M., J. Ludvik, et al. (1985). "New bacteriophage-like particles in *Corynebacterium glutamicum*." Virology **140**(2): 360-3.

Peters-Wendisch, P. G., B. J. Eikmanns, et al. (1993). "Phosphoenolpyruvate carboxylase in *Corynebacterium glutamicum* is dispensable for growth and lysine production." FEMS Microbiology Letters **112**: 269-274.

Peters-Wendisch, P. G., C. Kreutzer, et al. (1998). "Pyruvate carboxylase from *Corynebacterium glutamicum*: characterization, expression and inactivation of the pyc gene." Microbiology **144**(Pt 4): 915-27.

Peters-Wendisch, P. G., V. F. Wendisch, et al. (1996). "C3-Carboxylation as an anaplerotic reaction in phosphoenolpyruvate carboxylase- deficient *Corynebacterium glutamicum*." <u>Archives of Microbiology</u> **165**: 387-396.

Peters-Wendisch, P. G., V. F. Wendisch, et al. (1997). "Pyruvate carboxylase as an anaplerotic enzyme in *Corynebacterium glutamicum*." <u>Microbiology</u> **143**: 1095-1103.

Phibbs, P. V., T. W. Feary, et al. (1974). "Pyruate carboxylase deficiency in pleiotropic carboxhydrate-negative mutant strains of *Pseudomonas aeruginosa*." <u>Journal of Bacteriology</u> **118**: 999-1004.

Polakis, S. E., R. B. Guchhait, et al. (1973). "Stringent control of fatty acid synthesis in *Escherichia coli*. Possible regulation of acetyl coenzyme A carboxylase by ppGpp." <u>J Biol Chem</u> **248**(22): 7957-66.

Post, L. E., D. J. Post, et al. (1990). "Dissection of the functional domains of *Escherichia coli* carbamoyl phosphate synthetase by site-directed mutagenesis." <u>J Biol Chem</u> **265**(14): 7742-7.

Radford, A. J. and A. L. Hodgson (1991). "Construction and characterization of a *Mycobacterium-Escherichia coli* shuttle vector." <u>Plasmid</u> **25**(2): 149-53.

Ratledge, C. (1982). "The Biology of the *Mycobacteria*." <u>Academic Press: London</u>.

Renner, E. D. and R. W. Bernlohr (1972). "Characterization and regulation of pyruvate carboxylase of *Bacillus lichenformis*." <u>Journal of Bacteriology</u> **109**: 764-768.

Riddles, P. W., R. L. Blakeley, et al. (1978). "Ellman's Reagent: 5,5'-Dithiobis (2-nitrobenzoic Acid)- a Reexamination." <u>Analytical Biochemistry</u> **94**: 75-81.

Ruiz-Amil, M., G. De Torrontegui, et al. (1965). "Properties and function of yeast pyruvate carboxylase." J Biol Chem **240**(9): 3485-92.

Rylatt, D. B., D. B. Keech, et al. (1977). "Pyruvate carboxylase: the amino acid sequence at the biotin-attachment site of the enzymes isolated from chicken, turkey and sheep liver [proceedings]." Biochem Soc Trans **5**(5): 1544-6.

Sakurai, N., Y. Imai, et al. (1993). "Molecular breeding of a biotin-hyperproducing *Serratia marcescens* strain." Applied and Environmental Microbiology **59**: 3225-3232.

Sambrook, J., E. F. Fritsch, et al. (1989). "Molecular cloning: a laboratory manual." Cold Spring Harbor Laboratory **Cold Spring Harbor, NY**.

Samols, D., C. G. Thornton, et al. (1988). "Evolutionary conservation among biotin enzymes." Journal of Biological Chemistry **263**: 6461-6464.

Sano, K. and I. Shiio (1970). "Microbial production of L-lysine III. Production by mutants resistant to S-(2-aminoethyl)-L-cysteine." Journal of General and Applied Microbiology **16**: 373-391.

Santamaria, R., J. A. Gil, et al. (1984). "Characterization of an endogenous plasmid and development of cloning vectors and a transformation system in *Brevibacterium lactofermentum*." Journal of General Microbiology **130**: 2237-2246.

Santamaria, R. I., J. A. Gil, et al. (1985). "High-frequency transformation of *Brevibacterium lactofermentum* protoplasts by plasmid DNA." J Bacteriol **162**(1): 463-7.

Sanwal, B. D. (1970). "Allosteric controls of amphilbolic pathways in bacteria." Bacteriol Rev **34**(1): 20-39.

Schafer, A., J. Kalinowski, et al. (1990). "High-frequency conjugal plasmid transfer from gram-negative *Escherichia coli* to various gram-positive coryneform bacteria." J Bacteriol **172**(3): 1663-6.

Schafer, A., A. Schwarzer, et al. (1994). "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from *Corynebacterium glutamicum* ATCC 13032 and analysis of its role in intergeneric conjugation with *Escherichia coli*." J Bacteriol **176**(23): 7309-19.

Schiller, J., N. Growman, et al. (1980). "Plasmids in *Corynebacterium diphtheriae* mediating erythromycin resistance." Antimicrobial Agents and Chemotherapy **18**: 814-821.

Schrumpf, B., L. Eggeling, et al. (1992). "Isolation and prominent characteristics of an L-lysine hyperproducing strain of *Corynebacterium glutamicum*." Applied Microbiology and Biotechnology **37**: 566-571.

Schrumpf, B., A. Schwarzer, et al. (1991). "A functionally split pathway for lysine synthesis in *Corynebacterium glutamicium*." J Bacteriol **173**(14): 4510-6.

Schulz, G. E. and R. H. Schrimer (1979). "Principles of protein structure." Cantor, C.R., Ed **Springer Verlag, New York**: 205.

Scrutton, M.C. and M. R. Young (1972). "Pyruvate Carboxylase." The Enzymes **6**: 1-35.

Scrutton, M. C. (1978). "Fine control of the conversion of pyruvate (phosphoenolypyruvate) to oxaloacetate in various species." FEBS Lett **89**(1): 1-9.

Scrutton, M. C. and B. L. Taylor (1974). "Isolation and characterization of pyruvate carboxylase from *Azotobacter vinelandii* OP." Arch Biochem Biophys **164**(2): 641-54.

Serebriiskii, I. G., V. M. Vassin, et al. (1996). "Two new members of the bio B superfamily: cloning, sequencing and expression of bio B genes of *Methylobacillus flagellatum* and *Corynebacterium glutamicum*." Gene **175**(1-2): 15-22.

Serwold-Davis, T. M. and N. B. Growman (1986). "Mapping and cloning of *Corynebacterium dephtheriae* plasmid pNG2 and characterization of its relatedness to plasmids from skin coryneforms." Antimicrobial Agents and Chemotherapy **123**: 69-72.

Shiio, = I., S. I. Otsuka, et al. (1959). "Glutamic acid formation from glucose by bacteria. I. Enzymes of the Embden-Meyerhof-Parnas pathway, the Krebs-cycle, and the glyoxylate bypass in cell extracts of *Brevibacterium flavum* No. 2247." Journal of Biochemistry **46**: 1303-1311.

Shiio, I. and R. Miyajima (1969). "Concerted inhibition and its reversal by end products of aspartate kinase in *Brevibacterium flavum*." Journal of Biological Chemistry **65**: 849-859.

Shiio, I., H. Yoshino, et al. (1990). "Isolation and properties of lysine- producing mutants with feedback- resistant aspartokinase derived from *Brevibacterium flavum* strain with citrate synthase- and pyruvate kinase- defects and feed back resistant phosphoenol pyruvate carboxylase." Agricultural and Biological Chemistry **54**: 3275-3282.

Simon, R., M. O'Connell, et al. (1986). "Plasmid vectors for the genetic analysis and manipulation of rhizobia and other gram-negative bacteria." Methods Enzymol **118**: 640-59.

Sonnen, H., J. Schneider, et al. (1990). "Characterization of phi GA1, an inducible phage particle from *Brevibacterium flavum*." J Gen Microbiol **136**(Pt 3): 567-71.

Sonnen, H., J. Schneider, et al. (1990). "Corynephage Cog, a virulent bacteriophage of *Corynebacterium glutamicum*, and its relationship to phi GA1, an inducible phage particle from *Brevibacterium flavum*." J Gen Virol **71**(Pt 8): 1629-33.

Sonntag, K., L. Eggeling, et al. (1993). "Flux partitioning in the split pathway of lysine synthesis in *Corynebacterium glutamicum*. Quantification by 13C- and 1H-NMR spectroscopy." Eur J Biochem **213**(3): 1325-31.

Stephanopoulos, G. and J. J. Vallino (1991). "Network rigidity and metabolic engineering in metabolite overproduction." Science **252**(5013): 1675-81.

Stryer, L. (1988). "Biochemistry." 3rd edition **Freeman, NY**.

Stucka, R., S. Dequin, et al. (1991). "DNA sequences in chromosomes II and VII code for pyruvate carboxylase isoenzymes in *Saccharomyces cerevisiae*: analysis of pyruvate carboxylase-deficient strains." Mol Gen Genet **229**(2): 307-15.

Sumper, M. and C. Riepertinger (1972). "Structural relationship of biotin-containing enzymes." European Journal of Biochemistry **29**: 237-246.

Sundaram, T. K. (1973). "Physiological role of pyruvate carboxylase in a thermophilic bacillus." J Bacteriol **113**(2): 549-57.

Sundaram, T. K., J. J. Cazzulo, et al. (1971). "Synthesis of pyruvate carboxylase from its apoenzyme and (+)-biotin in *Bacillus stearothermophilus*. Mechanism and control of the reaction." Biochem J **122**(5): 663-9.

Suzuki, Y., Y. Aoki, et al. (1994). "Isolation and characterization of mutations in the human holocarboxylase synthetase cDNA." Nat Genet **8**(2): 122-8.

Taylor, B. L., S. Routman, et al. (1975). "The control of the synthesis of pyruvate carboxylase in *Pseudomonas citronellolis*. Evience from double labeling studies." J Biol Chem **250**(6): 2376-82.

Theodore, T. S. and E. Englesberg (1964). "Mutant of *Salmonella typhimurium* deficient in the carbon dioxide-fixing enzyme phosphoenolpyruvate carboxylase." Journal of Bacteriology **88**: 946-950.

Thierbach, G., J. Kalinowski, et al. (1990). "Cloning of a DNA fragment from *Corynebacterium glutamicum* conferring aminoethyl cysteine resistance and feedback resistance to aspartokinase." Appl Microbiol Biotechnol **32**(4): 443-8.

Toh, H., T. Kawamura, et al. (1994). "Molecular evolution of phosphoenolpyruvate carboxylase." Plant Cell Environ. **17**: 31-43.

Toh, H., H. Kondo, et al. (1993). "Molecular evolution of biotin-dependent carboxylases." Eur J Biochem **215**(3): 687-96.

Tolan, J. S. and R. K. Finn (1987). "Fermentation of D-xylose to ethanol by genetically modified *Klebsiella planticola*." Applied and Environmental Microbiology **53**: 2039-2044.

Tomioka, N., K. Shinozaki, et al. (1981). "Molecular cloning and characterization of ribosomal RNA genes from a blue green alga, *Anacystis nidulans*." Molecular and General Genetics **184**: 359-363.

Tosaka, O., H. Morioka, et al. (1979). "The role of biotin-dependent pyruvate carboxylase in L-Lysine production." Agricultural and Biological Chemistry **43**: 1513-1519.

206

Tosaka, O., K. Takinami, et al. (1978). "L-Lysine production by S-(2-amioethyl)L-cysteine and α-amino-β-hydroxyvaleric acid resistant mutants of *Brevibacterium lactofermentum*." Agricultural and Biological Chemistry **42**: 745-752.

Towbin, H., T. Staehelin, et al. (1979). "Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications." Proceedings of National Acedemy of Sciences, USA **76**: 4350-4354.

Trautwetter, A., C. Blanco, et al. (1987). "Characterization of the corynebacteriophage CG33." J Gen Microbiol **133**(Pt 10): 2945-52.

Trautwetter, A., C. Blanco, et al. (1987). "Structural characteristics of the *Corynebacterium lilium* bacteriophage CL31." J Virol **61**(5): 1540-5.

Treadway, S. L., K. S. Yanagimachi, et al. (1999). "Isolation and characterization of indene bioconversion genes from *Rhodococcus* strain I24." Appl Microbiol Biotechnol **51**(6): 786-93.

Trieu-Cuot, P., C. Carlier, et al. (1987). "Plasmid transfer by conjugation from *Escerichia coli* to gram positive bacteria." FEMS Micorbiology Letters **48**: 289-294.

Uy, D., S. Delaunay, et al. (1999). "A method for the determination of pyruvate carboxylase activity during the glutamic acid fermentation with *Corynebacterium glutamicum*." J Microbiol Methods **39**(1): 91-6.

Vallino, J. J. (1991). Identification of branch-point restrictions in microbial metabolism through metabolic flux analysis and local network perturbations. Chemical Engineering. Cambridge, Massachusetts Institute of Technology.

Vallino, J. J. and G. Stephanopoulos (1993). "Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction." Biotechnology and Bioengineering **41**: 633-646.

Vallino, J. J. and G. Stephanopoulos (1994). "Carbon flux distributions at the pyruvate branch point in *Corynebacterium glutamicum* during lysine overproduction." Biotechnology Progress **10**: 320-326.

Vallino, J. J. and G. Stephanopoulos (2000). "Metabolic flux distributions in *Corynebacterium glutamicum* during growth and lysine overproduction. Reprinted from Biotechnology and Bioengineering, Vol. 41, Pp 633-646 (1993) [classical article]." Biotechnol Bioeng **67**(6): 872-85.

van der Oost, J., M. Ciaramella, et al. (1998). "Molecular biology of hyperthermophilic Archaea." Adv Biochem Eng Biotechnol **61**: 87-115.

van der Rest, M. E., C. Lange, et al. (1999). "A heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogeneic plasmid DNA." Appl Microbiol Biotechnol **52**(4): 541-5.

Volpe, J. J. and P. R. Vagelos (1976). "Mechanisms and regulation of biosynthesis of saturated fatty acids." Physiol Rev **56**(2): 339-417.

Vrljic, M., H. Sahm, et al. (1996). "A new type of transporter with a new type of cellular function: L-lysine export from *Corynebacterium glutamicum*." Mol Microbiol **22**(5): 815-26.

Wakil, S. J., J. K. Stoops, et al. (1983). "Fatty acid synthesis and its regulation." Annu Rev Biochem **52**: 537-79.

Wendisch, V. F., A. A. de Graaf, et al. (2000). "Quantitative determination of metabolic fluxes during coutilization of two carbon sources: comparative analyses with *Corynebacterium glutamicum* during growth on acetate and/or glucose." J Bacteriol **182**(11): 3088-96.

Wexler, I. D., Y. Du, et al. (1994). "Primary amino acid sequence and structure of human pyruvate carboxylase." Biochim Biophys Acta **1227**(1-2): 46-52.

Xu, X., A. Matsuno-Yagi, et al. (1993). "DNA sequencing of the seven remaining structural genes of the gene cluster encoding the energy-transducing NADH-quinone oxidoreductase of *Paracoccus denitrificans*." Biochemistry **32**(3): 968-81.

Yakunin, A. F. and P. C. Hallenbeck (1997). "Regulation of synthesis of pyruvate carboxylase in the photosynthetic bacterium *Rhodobacter capsulatus*." J Bacteriol **179**(5): 1460-8.

Yamaguichi, R., M. Terabe, et al. (1986). "Determination of the complete nucleotide sequence of *Brevibacterium lactofermentum* plasmid pAM330 and analysis of its genetic information." Agricultural and Biological Chemistry **50**: 2771-2778.

Yarmush, M. L. and F. Berthiaume (1997). "Metabolic engineering and human disease." Nat Biotechnol **15**(6): 525-8.

Yokota, A. and I. Shiio (1988). "Effects of reduced citrate synthase activity and feedback resistant phosphoenol pyruvate carboxylase on lysine productivities of *Brevibacterium flavum* mutants." Agricultural and Biological Chemistry **52**: 455-463.

Yoshihama, M., K. Higashiro, et al. (1985). "Cloning vector system for *Corynebacterium glutamicum*." J Bacteriol **162**(2): 591-7.

Zhang, J., W. L. Xia, et al. (1993). "Adipose pyruvate carboxylase: amino acid sequence and domain structure deduced from cDNA sequencing." Proc Natl Acad Sci U S A 90(5): 1766-70.

Zhang, M., C. Eddy, et al. (1995). "Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*." Science 267: 240-243.

Zhang, Y., J. Praszkier, et al. (1994). "Molecular analysis and characterization of a broad-host-range plasmid, pEP2." J Bacteriol 176(18): 5718-28.

THESIS PROCESSING SLIP

FIXED FIELD: ill._____ name _____

index_____ biblio _____

▶ **COPIES:** Archives   Aero   Dewey   Barker   Hum

Lindgren   Music   Rotch   Science   Sche-Plough

TITLE VARIES: ▶ ☐ _____

_____

_____

NAME VARIES: ▶ ☐ _____

_____

IMPRINT: (COPYRIGHT)_____

_____

▶ **COLLATION:** _____

_____

▶ **ADD: DEGREE:** _____ ▶ **DEPT.:** _____

▶ **ADD: DEGREE:** _____ ▶ **DEPT.:** _____

SUPERVISORS: _____

_____

NOTES:

	cat'r:	date:
		page:
▶ **DEPT:** _____		▶

▶ **YEAR:** _____ ▶ **DEGREE:** _____

▶ **NAME:** _____

_____