Order Number 9402399

Cloning and characterization of the *Streptococcus thermophilus* galactokinase gene

Mustapha, Azlin, Ph.D.

The University of Nebraska - Lincoln, 1993

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

CLONING AND CHARACTERIZATION OF THE *Streptococcus thermophilus* GALACTOKINASE GENE

by

Azlin Mustapha

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Food Science and Technology

Under the Supervision of Professor Robert W. Hutkins

Lincoln, Nebraska

August, 1993

DISSERTATION TITLE

CLONING AND CHARACTERIZATION OF THE

Streptococcus thermophilus GALACTOKINASE GENE

BY

AZLIN MUSTAPHA

SUPERVISORY COMMITTEE:

APPROVED	DATE
Robert Hutkins	7-30-93
Signature	
ROBERT W. HUTKINS	
Typed Name	
Lloyd B. Bullerman	7/30/93
Signature	
LLOYD B. BULLERMAN	
Typed Name	
Randy L. Wehling	7/30/93
Signature	
RANDY L. WEHLING	
Typed Name	
Kenneth W. Nickerson	7/30/93
Signature	
KENNETH W. NICKERSON	
Typed Name	
Michael M. Meyer	7/30/93
Signature	
Typed Name	
Signature	
Typed Name	

GRADUATE COLLEGE
UNIVERSITY OF NEBRASKA

CLONING AND CHARACTERIZATION OF THE *Streptococcus thermophilus* GALACTOKINASE GENE

Azlin Mustapha, Ph.D.

University of Nebraska, 1993

Advisor: Robert W. Hutkins

The physiology and genetics of carbohydrate metabolism in *Streptococcus thermophilus* have only recently been studied. *S. thermophilus* utilizes a lactose permease and the enzyme β-galactosidase to transport and hydrolyze lactose inside the cell. Additionally, most wild-type strains of this organism are also unable to ferment lactose completely and release the galactose portion back into the medium. Galactose-fermenting strains of *S. thermophilus* have been shown to metabolize galactose via the enzymes of the Leloir pathway, two of which (galactose permease and galactokinase) are rate-limiting. However, even these few Gal^+ strains release this sugar into the medium when grown on lactose.

It is economically and scientifically feasible to increase the metabolic diversity of this important thermophilic starter culture. The main intent of this work, therefore, was to achieve a greater understanding of galactose metabolism in *S. thermophilus* with regard to its basic physiological and genetic characteristics. The specific objectives of this work were to first isolate a strain of *S. thermophilus* which can ferment galactose without releasing it into the medium, and to isolate, identify, and purify the galactokinase gene (*galK*)

from this organism. Next, the purified *galK* gene was cloned in *Escherichia coli*, and the gene was further analyzed by determining its size, restriction map, and nucleotide, as well as amino acid sequence. This *galK* gene was also compared with that of other organisms, and finally, flanking regions were analyzed in order to identify and locate any other adjacent genes.

The *galK* gene from *S. thermophilus* F410 was successfully cloned in *E. coli* in this study. The putative size of the gene was determined to be about 3.9 kb which translated to a protein monomer of about 49 kDa. The galactokinase from *S. thermophilus* F410 was found to exhibit significant homology to that of *Lactobacillus helveticus, E. coli, Haemophilus influenzae,* and *Kluyveromyces lactis*. Two additional genes, *lacS* and *lacZ*, have been proposed to be located downstream of this *galK* based on significant homology found between this region and the *lacSZ* genes and gene products from *S. thermophilus* A147.

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge the National Dairy Board and the University of Nebraska Center for Biotechnology for the provision of funds for this research project. Most importantly, deepest gratitude is extended to my advisor, Dr. Bob Hutkins, for giving me the opportunity to work under his supervision. If not for his guidance, patience, and support, this work would have never been completed. Thanks are also extended to Drs. Nickerson, Wehling, Bullerman, and Meagher for serving on my commitee and for their helpful suggestions.

To my labmates and friends, Suzette, Theresa, Melody, Doug, Beatriz, and Jayne, I wish to express my thanks for their wonderful friendship. Special thanks go to Dave, who has been a constant wall of support for me, especially during the preparation of this thesis.

I am grateful to Suzette for her technical assistance in conducting *galK* assays, and to Robin and Becky for their abundant help throughout this research effort. I also wish to thank Shane and Susan for sharing their computer expertise.

Finally, I thank my parents and grandparents for their unrelenting love and support. It is to them that I dedicate this thesis.

TABLE OF CONTENTS

PAGE

PAGE

PAGE

LIST OF FIGURES

PAGE

LIST OF TABLES

LITERATURE REVIEW

METABOLISM OF CARBOHYDRATES BY *Streptococcus thermophilus*: PHYSIOLOGY AND GENETICS

I. INTRODUCTION

Lactic acid bacteria have long been used in the manufacture of various fermented food products. The dairy industry represents one of the major users of lactic acid bacteria for the production of cheese, yogurt, and other fermented milk products. The importance of this group of organisms has thus led to extensive research efforts aimed at gaining a better understanding of their physiological and genetic properties. Knowledge about the different properties of mesophilic starter cultures, especially the lactococci, has rapidly accumulated, and now many of the details regarding their carbohydrate metabolism and regulation, as well as genetic aspects involved in various traits are well established (55).

Until recently, *Streptococcus thermophilus*, an equally important dairy starter culture, has surprisingly stimulated much less research efforts. It is now known that carbohydrate fermentation in this organism differs markedly from that of the lactococci and other mesophilic lactic acid bacteria. In fact, metabolism of sugars by *S. thermophilus* occurs more like that of *Escherichia coli* than other, more closely related lactic acid bacteria. Moreover, as will be described later in this review, recent studies indicate that many of the genes responsible for sugar catabolism in *S. thermophilus* have higher homology with *E. coli* genes than those of other lactic acid bacteria.

This review will first cover the numerous taxonomic rearrangements of the streptococci and the independent status of *S. thermophilus*. An overview

of the basic biochemical properties of this organism will follow. Various aspects of carbohydrate utilization in *S. thermophilus* will then be discussed, with the main focus being on lactose and galactose transport, metabolism, and regulation, and their differences from that of *Lactococcus*. The genetic aspects of lactose and galactose metabolism will also be examined. Finally, future applications for genetically modifying specific traits in order to obtain improved starter strains will be presented.

II. TAXONOMY OF Streptococcus thermophilus

Since it was first reported by Rosenbach in 1884 (23), the genus *Streptococcus* has undergone numerous taxonomic rearrangements. This is due mainly to the extensive heterogeneity of this group of organisms. In 1933, Lancefield (46) divided the streptococci into different groups based on the presence of group-specific antigens. This serological testing proved to be a very important tool in subsequent classification studies of the streptococci, and was used widely. Based on the Lancefield serological classification, *Streptococcus lactis, Streptococcus cremoris,* and *Streptococcus lactis* subsp. *lactis* all possess the group N antigen (hence, the classification, Group N streptococci), whereas *S. thermophilus* (first described by Orla-Jensen in 1919) did not possess any group-specific antigen.

The earliest systematic classification scheme of the group streptococci was proposed by Sherman in 1937 (95), who based his characterization of the

different species on simple physiological, biochemical, and serological studies. Sherman divided this genus into four broad groups- enterococci, pyogenic, lactic, and viridans. *S. thermophilus*, was placed by Sherman in the viridans group based on its somewhat close relatedness to the species in that group. It was acknowledged however, that the physiological and serological characteristics of *S. thermophilus* were so distinct from the other streptococci, that it clearly did not belong in any of the four groups.

In 1978, Jones (35) reorganized the streptococci into seven different groups- pyogenic, pneumococci, oral, fecal, lactic, anaerobic, and other streptococci. His classification emerged as a result of more extensive and sophisticated methods of differentiation, including DNA-DNA homology, guanine and cytosine (G + C) %mol ratios, protein homology, additional serological and physiological testings, and numerical taxonomic studies. This classification subsequently replaced the earlier Sherman scheme in the eighth edition of Bergey's Manual of Systematic Bacteriology (23), though it was mainly for convenience and not due to any taxonomic validity.

Yet another classification scheme was proposed by Sharpe in 1979 (94), who divided the streptococci into four major groups- pyogenic, fecal, oral, and lactic. This author placed *S. thermophilus* in the lactic group, with the acknowledgement that it was for practical purposes only, and not due to any degree of relatedness of this organism to the species of that group. Her classification was based on combined physiological, biochemical, and

serological analysis, as opposed to strict emphasis on serological groupings of earlier reports (35, 95, 111).

More recent studies on this highly diverse group of organisms have resulted in a major reclassification of some of its members at both the genus and species level. Results from various laboratories (33, 40, 48, 88, 89) have led to the reclassification of *Streptococcus faecalis* and *Streptococcus faecium*, as *Enterococcus faecalis* and *Enterococcus faecium*, respectively, in 1984; and the transfer of the group N lactic streptococci (*S. lactis* subsp. *lactis, S. lactis* subsp. *diacetylactis*, and *S. lactis* subsp. *cremoris*), *Lactobacillus hordniae, Lactobacillus xylosus, Streptococcus garvieae, Streptococcus raffinolactis,* and *Streptococcus plantarum* to a new genus, *Lactococcus*, in 1985. This reclassification scheme has been accepted and is now widely used.

Although most of the taxonomy work has been conducted on the mesophilic or Group N streptococci, it was clear that a problem also lies with the grouping of *S. thermophilus*. This organism has always been closely associated with the group N streptococci (*Lactococcus*) because of its use as a starter culture in the dairy industry, even though it is physiologically, biochemically, serologically, and genetically distinct from this group. Furthermore, Kandler (36) showed that the structure of the *S. thermophilus* cell wall peptidoglycan (L-lysine-L-alanine-L-alanine type) was identical to that of *S. faecalis*.

In 1982, Kilpper-Balz *et al.* (40) demonstrated close DNA homology

(75%-97%) between *S. thermophilus* and the oral *Streptococcus salivarius*. Farrow and Collins, in 1984 (18), confirmed this by showing high DNA homology between the two groups (in the range of 61-100% under optimum conditions), as well as similar fatty acid profiles. These findings led these authors to propose the reclassification of *S. thermophilus* as *Streptococcus salivarius* subsp. *thermophilus* (18). Although the current volume of Bergey's (23) retained the former name, many investigators began using the new name for this organism, as recommended by Farrow and Collins (18). Further DNA hybridization studies performed by Schleifer and Kilpper-Balz (90), however, demonstrated DNA homology values of 60% only under optimal, but only 30% under stringent hybridization conditions. These authors then proposed the revival of the two separate species: *Streptococcus thermophilus* and *Streptococcus salivarius*. Subsequent DNA hybridization studies later performed by Schleifer *et al.* (91) in 1991, further confirmed these reports, and based on these results as well as the supportive findings using numerical taxonomic studies (6), and simply due to the distinct environments from which these two organisms are found (milk and mouth respectively), it was officially proposed by Schleifer *et al.* in 1991 (91) to reclassify *S. salivarius* subsp. *thermophilus* as *S. thermophilus*.

It is clearly shown by all the studies mentioned here that *S. thermophilus* does not belong in any one particular grouping with the other streptococci, but rather is distinct from them. As mentioned above, the species *S. thermophilus*

has been revived, and for this thesis, this organism will be referred to as such.

III. PROPERTIES OF *Streptococcus thermophilus*

The identification and characterization of *S. thermophilus* has been based on a multitude of studies, ranging from basic morphological, physiological, and biochemical work (69, 95, 94) to serological (46) and more recently, genetic studies (33, 40). The latter have encompassed more sophisticated techniques such as DNA-DNA and DNA-RNA hybridization techniques, and numerical taxonomic investigations (6).

The genus *Streptococcus* is described in the latest edition of Bergey's (23) as Gram positive organisms that are mainly facultatively anaerobic, metabolically fermentative, producing mainly lactic acid but no gas, catalase-negative, non-sporeforming spherical or ovoid cells occurring in pairs or chains in liquid media. *S. thermophilus* is additionally described to be 0.7-0.9 μm in diameter and able to grow at 45°C. The streptococci are heterotrophic and are generally fastidious, requiring simple carbohydrates for an energy source, and preformed amino acids as a nitrogen source (23).

One of the best descriptions of *S. thermophilus* was provided by Sherman (95), who stated that this organism is notable "more by the things it is unable to do, rather than by those which it can do" (see Table 1). *S. thermophilus* does not possess a group-specific antigen, which immediately distinguishes it from the other streptococci (18, 69, 95). According to

Table 1. Main properties of *S. thermophilus*

Genotype	
DNA base content (GC %)	38%-40%
Nucleic acid homology	different from allied streptococci
Relationship of isofunctional enzymes (aldolases) with those of:	*L. lactis* subsp. *lactis*: Yes *L. lactis* subsp. cremoris: Yes *L. lactis* subsp. *diacetylactis*: Yes
Phenotype	
Morphology	- Spherical or ovoid cells in pairs or long chains - Strong polymorphism in old cells
Ecology	Milk, dairy products only
Growth at 10°C	-
Growth at 45°C	+
Heat resistance at 65°C, 30 min	+
Growth in 2% NaCl	-
Serological group	Absence of group antigen
Arginine hydrolysis	-
Fermented sugars:	
- fructose, glucose	+
- lactose, sucrose	+
- galactose	(-)
- maltose, xylose, arabinose	+
- raffinose	(-)
- trehalose, inulin, glycerol, mannitol, sorbitol	- -
Action on litmus milk	Rapid acidification; A Coagulation; C Very slow and often incomplete reduction of litmus; r

+ = positive reaction for 90% or more strains
(-) = negative reaction for 90% or more strains
- = reaction always negative

(Adapted from Ref. 1)