Development of improved α-amylases

by

Nivetha Ramachandran

Thesis presented in partial fulfilment of the requirements for the doctoral degree of Sciences at Stellenbosch University.

April 2005

Supervisor:
Dr Ricardo R. Cordero Otero

Co-supervisor:
Prof. Isak S. Pretorius

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

_____ _____

Nivetha Ramachandran **Date**

SUMMARY

The technological advancement of modern human civilisation has, until recently, depended on extensive exploitation of fossil fuels, such as oil, coal and gas, as sources of energy. Over the last few decades, greater efforts have been made to economise on the use of these non-renewable energy resources, and to reduce the environmental pollution caused by their consumption. In a quest for new sources of energy that will be compatible with a more sustainable world economy, increased emphasis has been place on researching and developing alternative sources of energy that are renewable and safer for the environment.

Fuel ethanol, which has a higher octane rating than gasoline, makes up approximately two-thirds of the world's total annual ethanol production. Uncertainty surrounding the long-term sustainability of fuel ethanol as an energy source has prompted consideration for the use of bioethanol (ethanol from biomass) as an energy source. Factors compromising the continued availability of fuel ethanol as an energy source include the inevitable exhaustion of the world's fossil oil resources, a possible interruption in oil supply caused by political interference, the superior net performance of biofuel ethanol in comparison to gasoline, and a significant reduction in pollution levels. It is to be expected that the demand for inexpensive, renewable substrates and cost-effective ethanol production processes will become increasingly urgent.

Plant biomass (including so-called 'energy crops', agricultural surplus products, and waste material) is the only foreseeable sustainable source of fuel ethanol because it is relatively low in cost and in plentiful supply. The principal impediment to more widespread utilisation of this important resource is the general absence of low cost technology for overcoming the difficulties of degrading the recalcitrant polysaccharides in plant biomass to fermentable sugars from ethanol can be produced. A promising strategy for dealing with this obstacle involves the genetic modification of *Saccharomyces cerevisiae* yeast strains for use in an integrated process, known as direct microbial conversion (DMC) or consolidated bioprocessing (CBP). This integrated process differs from the earlier strategies of SHF (separate hydrolysis and fermentation) and SSF (simultaneous saccharification and fermentation, in which enzymes from external sources are used) in that the production of polysaccharide-degrading enzymes, the hydrolysis of biomass and the fermentation of the resulting sugars to ethanol all take place in a single process by means of a polysaccharide-fermenting yeast strain.

The CBP strategy offers a substantial reduction in cost if *S. cerevisiae* strains can be developed that possess the required combination of substrate utilisation and product formation properties. *S. cerevisiae* strains with the ability to efficiently utilise polysaccharides such as starch for the production of high ethanol yields have not been described to date. However, significant progress towards the development of such amylolytic strains has been made over the past decade.

With the aim of developing an efficient starch-degrading, high ethanol-yielding yeast strain, our laboratory has expressed a wide variety of heterologous amylase-encoding genes in *S. cerevisiae*. This study forms part of a large research programme aimed at improving these amylolytic 'prototype' strains of *S. cerevisiae*. More specifically, this study investigated

the *LKA1*- and *LKA2*-encoded α-amylases (Lka1p and Lka2p) from the yeast *Lipomyces kononenkoae*. These α-amylases belong to the family of glycosyl hydrolases (EC 3.2.1.1) and are considered to be two of the most efficient raw-starch-degrading enzymes. Lka1p functions primarily on the α-1,4 linkages of starch, but is also active on the α-1,6 linkages. In addition, it is capable of degrading pullulan. Lka2p acts on the α-1,4 linkages.

The purpose of this study was two-fold. The first goal was to characterise the molecular structure of Lka1p and Lka2p in order to better understand the structure-function relationships and role of specific amino acids in protein function with the aim of improving their substrate specificity in raw starch hydrolysis. The second aim was to determine the effect of yeast cell flocculence on the efficiency of starch fermentation, the possible development of high-flocculating, *LKA1*-expressing *S. cerevisiae* strains as 'whole-cell biocatalysts', and the production of high yields of ethanol from raw starch.

In order to understand the structure-function relationships in Lka1p and Lka2p, standard computational and bioinformatics techniques were used to analyse the primary structure. On the basis of the primary structure and the prediction of the secondary structure, an N-terminal region (1-132 amino acids) was identified in Lka1p, the truncation of which led to the loss of raw starch adsorption and also rendered the protein less thermostable. Lka1p and Lka2p share a similar catalytic TIM barrel, consisting of four highly conserved regions previously observed in other α-amylase members. Furthermore, the unique Q^{414} of Lka1p located in the catalytic domain in place of the invariant H^{296} (TAKA amylase), which offers transition state stabilisation in α-amylases, was found to be involved in the substrate specificity of Lka1p. Mutational analysis of Q^{414} performed in the current study provides a basis for understanding the various properties of Lka1p in relation to the structural differences observed in this molecule. Knowing which molecular features of Lka1p contribute to its biochemical properties provides us with the potential to expand the substrate specificity properties of this α-amylase towards more effective processing of its starch and related substrates.

In attempting to develop 'whole-cell biocatalysts', the yeast's capacity for flocculation was used to improve raw starch hydrolysis by *S. cerevisiae* expressing *LKA1*. It was evident that the flocculent cells exhibited physicochemical properties that led to a better interaction with the starch matrix. This, in turn, led to a decrease in the time interval for interaction between the enzyme and the substrate, thus facilitating faster substrate degradation in flocculent cells. The use of flocculation serves as a promising strategy to best exploit the expression of *LKA1* in *S. cerevisiae* for raw starch hydrolysis.

This thesis describes the approaches taken to investigate the molecular features involved in the function of the *L. kononenkoae* α-amylases, and to improve their properties for the efficient hydrolysis of raw starch. This study contributes to the development of amylolytic *S. cerevisiae* strains for their potential use in single-step, cost-effective production of fuel ethanol from inexpensive starch-rich materials.

OPSOMMING

Die tegnologiese vooruitgang van die moderne samelewing was tot onlangs toe van die ontginning van fossielenergiebronne soos steenkool en gas afhanklik. Gedurende die afgelope paar jare is aansienlike pogings aangewend om die gebruik van hierdie nie-hernieubare energiebronne te besuinig en om omgewingsbesoedeling agv die verbruik daarvan te verminder. In ‘n soeke na nuwe energiebronne wat verenigbaar is met ‘n volhoubare wêreldekonomie, word toenemend klem op die navorsing en ontwikkeling van alternatiewe hernieubare- en omgewingsvriendelike energiebronne gelê.

Brandstofetanol, wat van ‘n hoër oktaangehalte as petrol is, maak ongeveer twee derdes van die wêreld se jaarlikse etanolproduksie uit. Die onsekerhede rondom die langtermyn volhoubaarheid van brandstofetanol as ‘n energiebron het tot die oorweging van bioetanol (etanol vanaf biomassa) as energiebron aanleiding gegee. Faktore wat die volgehoue beskikbaarheid van brandstofetanol as energiebron in gedrang bring, is die onvermydelike uitputting van die wêreld se fossieloliebronne, die moontlike onderbreking van olievoorsiening as gevolg van politieke inmenging, die uitstekende werkverrigitng van bioetanol in vergelyking met petrol en die aansienlike laer besoedelingsvlakke. Dit kan dus verwag word dat die behoefte aan goedkoop, hernieubare substrate en koste-effektiewe etanolproduksieprosesse meer dringend sal word.

Plantbiomassa (insluitende die sogenaamde “energiegewasse” en landbou surplusprodukte en -afvalmateriaal) is die enigste volhoubare brandstofetanolbron agv die relatief lae koste en beskikbaarheid. Die sentrale tegnologiese hindernis vir die algemene gebruik van hierdie belangrike hulpbron, is die afwesigheid van bekostigbare tegnologie om die weerspannigheid van die plantbiomassa te oorkom. Die metaboliese manipulering van *Saccharomyces cerevisiae* gisrasse vir gebruik in ‘n geïntegreerde proses wat as “direkte mikrobiese omskakeling” of “gekonsolideerde bioprosessering” (KBP) bekendstaan, lewer ‘n belowende strategie om hierdie hindernis te oorkom. Hierdie geïntegreerde proses verskil van die vroeër-ontwikkelde SHF- en SSF-strategieë deurdat die produksie van polisakkariedafbraakensieme, die hidrolise van die biomassa en die fermentering van die suikers na etanol tydens ‘n enkele proses deur middel van ‘n polisakkaried-fermenterende gisras plaasvind.

Die KBP-strategie kan ‘n aansienlike kosteverlaging meebring indien *Saccharomyces cerevisiae*-rasse wat oor die vereisde kombinasie van substraatverbruik- en produkvormingseienskappe beskik, ontwikkel kan word. Geen geskikte polisakkariedafbrekende gisrasse is tot dusver beskryf nie. Aansienlike vordering in die ontwikkeling van sulke amilolitiese rasse is egter deur die loop van die afgelope dekade gemaak.

‘n Wye verskeidenheid mikrobiese amilase-gene is in ons laboratorium in *S. cerevisiae* uitgedruk met die doel om ‘n effektiewe styselafbrekende, hoë etanolproduserende gisras te ontwikkel. Hierdie studie vorm ‘n deel van ‘n groot navorsingsprogram wat op die verbetering van hierdie amilolitiese “prototipe”-gisrasse van *S. cerevisiae* gefokus is en het meer spesifiek die ondersoek van die struktuur-funksie ooreenkoms van die *LKA1*- en *LKA2*-gekodeerde α-amilases (Lka1p en Lka2p) van die gis *Lipomyces kononenkoae* behels. Hierdie α-amilases behoort aan die glikosielhidrolase-familie (EC 3.2.1.1) en word as twee van die mees effektiewe rou styselafbrekende ensieme beskou. Lka1p funksioneer primêr op die α-1,4 verbindings van stysel, maar is ook op die α-1,6 verbindings aktief. Hierdie ensiem beskik verder ook oor die vermoë om pullulaan af te breek. Lka2p funksioneer op die α-1,4 verbindings van stysel.

Die doelwit van hierdie studie was tweërlei. Die eerste doelwit was die molekulêre karakterisering van Lka1p en Lka2p vir die verdere ondersoek van die struktuur-funksie verhouding en die rol van spesifieke aminosure in proteienfunksie, asook die moontlike verbetering van hul substraatspesifisiteit in rou styselhidrolise. Die tweede doelwit was die bepaling van die effek van gisselflokkulering op die effektiwiteit van styselfermentering en die moontlike ontwikkeling van hoë flokkulerende *S. cerevisiae* gisrasse wat *LKA1* uitdruk as “heelsel-biokataliste” en hoë konsentrasies etanol vanaf rou stysel kan produseer.

Om die struktuur-funksie verhouding in Lka1p en Lka2p te verstaan, is standaard rekenaar- en bioinformatikategnieke ingespan en die primêre struktuur is geëvalueer. Op grond van die primêre struktuur en die voorspelling van die sekondêre struktuur, is ‘n N-terminale gebied (1-132 aminosure) in Lka1p geïdentifiseer. Die verkorting van hierdie gebied het die verlies van rou styseladsorpsie tot gevolg gehad en die molekuul was ook minder hittestabiel. Lka1p en Lka2p deel ‘n soortgelyke katalitiese TIM-silinder, wat uit vier hoogsgekonserveerde dele wat ook in ander α-amilases voorkom, bestaan. Die unieke Q^{414} van Lka1p wat in die katalitiese domein in die plek van die konstante H^{296} (TAKA amilase) geleë is, verleen oorgangstoestandstabilisering aan α-amilases en is by die substraatspesifisiteit van Lka1p betrokke. Mutasie-analises van Q^{414} wat in die betrokke studie uitgevoer is, verskaf ‘n basis vir die verklaring van die verskeie eienskappe van Lka1p wat met die strukturele variasies wat in hierdie molekuul waargeneem is, verband hou.

In ‘n poging om “heelsel-biokataliste” te ontwikkel, is die gis se flokkuleringsvermoë gebruik om rou styselhidrolise deur *S. cerevisiae* wat *LKA1* uitdruk, te verbeter. Dit was duidelik dat die flokkulerende selle oor fisies-chemiese eienskappe beskik het wat tot verbeterde interaksie met die styselmatriks gelei het. Dit het weer die interaksie tussen die ensiem en die substraat verkort en gevolglik subtraatafbraak sonder vertraging in flokkulerende selle vergemaklik. Die gebruik van flokkulering is ‘n belowende strategie om die uitdrukking van Lka1p in *Saccharomyces cerevisiae* vir rou styselhidrolise te benut.

In hierdie tesis word die strategieë om die molekulêre eienskappe wat by funksie van die *L. kononenkoae* α-amilases betrokke is, te ondersoek en die verbetering daarvan vir die effektiewe hidrolise van rou stysel beskryf. Hierdie studie lê die grondslag vir die ontwikkeling van amilolitiese *Saccharomyces cerevisiae*-rasse en hul potensiële gebruik in die een-stap, koste-effektiewe produksie van etanol vanaf goedkoop styselryke materiale.

This thesis is dedicated to my parents and to my aunt and uncle, Mrs Chitra and Mr Krishnamoorthy

ज्ञानं ज्ञेयं परिज्ञाता त्रिविधः कर्मचोदना ।
करणं कर्म कर्तेति त्रिविधः कर्मसंग्रहः ॥१८॥

“jnanam jneyam parijnata tri-vidha: karma-codana
karanam karma karteti tri-vidhah: karma-sangrahah”

“Knowledge, the object of knowledge and the knower
are the three factors which motivate action;
the senses, the work and the doer
comprise the threefold basis of action."

– *Bhagavad Gita*

BIOGRAPHICAL SKETCH

Nivetha Ramachandran was born in Coimbatore, Tamil Nadu, India on 12 November 1977. She matriculated from Vidya Vikasini Matriculation Higher Secondary School in 1992. Nivetha completed her Bachelors degree in Microbiology at Shri Nehru Maha Vidyalaya, which is affiliated to Barathiar University, Coimbatore, in 1997. She then enrolled for a Master's in Applied Microbiology at the P.S.G. College of Barathiar University, completing it in 1999. She joined the Institute for Wine Biotechnology at Stellenbosch University in 2001 to continue her studies towards a PhD.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

Dr Ricardo R. Cordero Otero, my research supervisor, for his guidance and all his efforts in developing my scientific skills.

Prof. Isak S. Pretorius, for providing the opportunity, and strength, inspiration and motivation throughout.

Prof. Melanie Vivier, Prof. Florian Bauer, Dr. P. Van Rensburg and Dr. M. du Toit, for invaluable moral support throughout my tenure at the Institute.

Dr Jonathan Arensburg, Patrick Govender, Dr Dewald van Dyk, Sharath, Jeremy Eksteen, Johan Kriel, Dr Philip Young, Dr Albert Joubert, John Becker, Dr Qin Qui Ma, Dr Shang Wu Cheng, Dr Elizabeth Rohwer and Jaco Minnaar, for their invaluable support in the techniques used in this study and for all their discussions.

Sharath, Sven Kroppenstedt, Vasudevan Gururajan, Annel Smit, Dr Dewald van Dyk, Adriaan Oelofse and all the colleagues at the Institute, for their friendship.

Stephanie Baard, Michelle Veenstra, Enzo D. Aguanno, Schalk Smit, Karin Vergeer, Tanya Ficker, Judy Cronje and all the administrative and support staff at the Institute, including **Judy Poole** and **Egon Februarie**, for their kind help throughout.

Dr Lydia Joubert, Department of Microbiology, and **Dr Martina Meincken**, Institute for Polymer Science, for their most appreciated help with the microscopy studies in this project.

The National Research Foundation, Winetech and the Harry Crossley Foundation, for financial support for this research project.

Dr Nilly Galyam and **Jawahar Sheriff**, for their care, affection and encouragement.

PREFACE

This thesis is presented as a compilation of six chapters. Additional information can be found in the appendix.

Chapter 1 **INTRODUCTION AND PROJECT AIMS**

Chapter 2 **LITERATURE REVIEW**
Molecular and cellular biocatalysts for starch hydrolysis

RESEARCH RESULTS

Chapter 3 Role of the N-terminal region of the *Lipomyces kononenkoae LKA1* and *LKA2* encoded α-amylases in substrate binding, thermal stability and substrate specificity

Chapter 4 Site-directed mutagenesis of Gln^{414} in the active site of *Lipomyces kononenkoae* α-amylase Lka1p

Chapter 5 Effect of flocculation on *Saccharomyces cerevisiae* expressing *Lipomyces kononenkoae* α-amylase Lka1p for raw starch fermentations

Chapter 6 **GENERAL CONCLUSION**

Appendix Amylolytic enzymes from the yeast *Lipomyces kononenkoae*

CONTENTS

CHAPTER 1. INTRODUCTION AND PROJECT AIMS

CHAPTER 2. MOLECULAR AND CELLULAR BIOCATALYSTS FOR STARCH HYDROLYSIS

CHAPTER 3. ROLE OF THE N-TERMINAL REGION OF THE *LIPOMYCES KONONENKOAE LKA1* AND *LKA2* ENCODED α-AMYLASES IN SUBSTRATE BINDING, THERMAL STABILITY AND SUBSTRATE SPECIFICITY

CHAPTER 4. SITE-DIRECTED MUTAGENESIS OF Gln^{414} IN THE ACTIVE SITE OF *LIPOMYCES KONONENKOAE* α- AMYLASE Lka1p

CHAPTER 5. EFFECT OF FLOCCULATION ON *SACCHAROMYCES CEREVISIAE* EXPRESSING *LIPOMYCES KONONENKOAE* α-AMYLASE Lka1p FOR RAW STARCH FERMENTATIONS

CHAPTER 1

INTRODUCTION AND PROJECT AIMS

1. INTRODUCTION

1.1 Bioethanol as an alternative fuel resource

Environmental pollution associated with fossil fuel use has led to a fundamental shift towards greater reliance on biomass in the world's energy system. With recent technological advances in the biotechnology sector, it is plausible to convert biomass into high quality energy carriers, such as electricity and liquid fuels (Larson *et al*., 1993).

Ethanol has been known for a long time and is perhaps the oldest product obtained through traditional biotechnology. It is a very attractive alternative fuel, not only because of its low contribution to the greenhouse gases compared to other fuels, but also because of its versatility and advantageous physical properties. It can be used in a low-level blend in unmodified automobiles and in high-level gasoline blends in modified vehicles. Similar to the high-level gasoline blends, it can also be used as neat ethanol and burnt in modified vehicles.

As ethanol has higher octane levels than gasoline, it is usually not necessary to add highly poisonous octane boosters. In addition to these advantages, ethanol is composed of oxygen. It therefore facilitates better combustion and reduces carbon monoxide emission. Ethanol can be blended with gasoline for use, but this initially increases the vapour pressure, which may lead to interference with the ozone production cycle. However, ethanol can be blended with gasoline in the form of ethyl tertiary butyl ether, which decreases the vapour pressure of gasoline, in turn decreasing the release of smog-forming compounds.

1.2 Conversion of starchy biomass to ethanol

The conversion of starchy biomass to ethanol has been an important focus during recent years. In principle, the process of making ethanol from cornstarch consists of two steps: breaking down complex starch into simple sugars, and fermentation; the ethanol can subsequently be refined by means of distillation. In general, there are two techniques available for the conversion of the sugar content of starch to ethanol. These are the acid hydrolysis of corn starch and the enzymatic breakdown process.

In the acid hydrolysis process, the components of starchy materials are broken down by the use of strong and concentrated aqueous solutions of mineral acids, such as hydrochloric acid and sulphuric acid, at temperatures lower than 100°C. The drawback of this method is that it requires high-grade stainless steel equipment, which could increase the capital investment considerably. The amylose fraction of starch is broken down earlier than the amylopectin and the sugars released at earlier time periods are exposed to harsh conditions, significantly lowering the production yields of the fermentation.

In the enzymatic process, the raw cornstarch material is pre-treated in order to increase accessibility to the starch-hydrolysing enzymes, such as α-amylases, glucoamylases and

pullulanases. During the pre-treatment process, the gelatinisation of cornstarch exposes the inner hydrophobic moiety, resulting in a helical form with contiguous hydrophobic surfaces, while the hydrophilic exterior becomes accessible to enzymatic attack. Since the conditions applied are mild and decrease the formation of other by-products, ethanol yields are higher. However, such enzymatic processing could be expensive if the addition of starch-degrading enzymes is required. The ideal situation would then be to use microorganisms capable of secreting these enzymes in the fermenter. Enzyme production is therefore a crucial process and it is necessary to produce efficient and robust enzymes capable of breaking down the cheap carbon sources available for these processes.

1.3 Artificial and natural biocatalysts

The role of microbes and their metabolites in governing the bioconversion processes that provide alternative energy is enormous. Microorganisms are found everywhere in nature and have been exploited to the benefit of mankind for thousands of years. They have been used in the production of beer, wine, cheese and many other products – all this even before we knew about the existence of microbes and before the term biotechnology was invented. With the emergence of biocatalysis as an important tool, the microbial cells and their enzymes have been used for the production of ethanol and other valuable products from cheaper raw materials. Traditionally, active biocatalysts have been obtained by screening a broad variety of microorganisms in nature, ranging from archaea to fungi, which are frequently isolated from extreme environments. These biocatalysts are used either as isolated enzymes or in the form of whole-cell preparations. Since the natural catalysts have several disadvantages to optimal use in our manmade processes, depending on the process parameters and adaptability in the environment in which we need to use them, recombinant systems were developed. Recombinant systems thus contain the gene encoding the desired enzyme, which is over expressed in a more limited set of industrially adapted microorganisms. The resulting “designer bugs” have an elevated level of the desired enzyme, as well as a low background of undesirable reactions catalysed by unwanted enzymes, because the genes coding for the latter enzymes are not transferred from the source microorganisms.

Dedicated efforts by researchers are being applied to explore biodiversity worldwide for novel enzymes at a genetic and functional level, which will further expand the arsenal of biocatalysts available for industrial applications. Modern technology is thus far being directed into two modes: a) engineering enzymes and proteins from nature that function as catalysts, and b) engineering whole cells for biocatalysis.

Enzymes from nature have been successfully modified or their activity has been enhanced with mutagenesis and protein engineering in which one or a few amino acid residues are rationally and directly replaced. Improvements by such directed protein engineering techniques have not always led to the desired result, and these methods can

also be time consuming. High-throughput screening and modern molecular biology techniques, in combination with tremendous improvements in genomics and bioinformatics, have led to the substantial availability of such modified enzymes. In addition, the use of enzymes that have been isolated is not always limited to the production of compounds that are similar to the natural substrate that the enzyme was made to convert. Engineering therefore includes increasing the scope of enzymes to a broader range of natural starting materials and products. Consequently, directed evolution has opened the path to biocatalysts with broader substrate ranges, as well as enzymes dedicated to a single, specific transformation.

In parallel, the operational stability of biocatalysts applied in industrial processes is also an essential criterion. While protein engineering focuses on mimicking the natural diversity in the tailored enzyme, the modification of whole cells secreting these enzymes has been advantageous due to the versatility of whole cells to adapt and respond to various processing conditions. This is achieved through genetic engineering, heterologous protein expression, cell-surface engineering and immobilisation for the improved processing of complex substrates used in the fermenter. Apart from these strategies, the development of systems biology and metabolic engineering has opened a new gateway for improving catalysis by means of microorganisms. By channelling the metabolic pathways in microorganisms towards a desired metabolite through the rational introduction and removal of genes – known as metabolic engineering – a new range of products can be produced.

In nature, bacteria, fungi and some yeast species produce an arsenal of starch hydrolases, which include glucoamylases, α-amylases, isoamylases, glucosidases, etc. Among the yeast species, *Lipomyces kononenkoae* is known to possess a highly efficient amylolytic system and is capable of degrading 98% of the raw starch supplied (Horn *et al*., 1988). Despite its efficiency in hydrolysing raw starch, its poorly characterised genetics and low ethanol tolerance make this organism unsuitable for fermentations. *Saccharomyces cerevisiae,* on the other hand, lacks the amylolytic system to degrade starch, but is the most exploited organism for fermentation because of its well-understood genotype and its high ethanol tolerance. The *LKA1* and *LKA2* genes, which encode the raw-starch-degrading Lka1p and Lka2p α-amylases, were cloned from *L. kononenkoae* strain IGC4052B and expressed in *S. cerevisiae* (Steyn *et al*., 1995; Eksteen *et al*., 2003a). Raw starch-degrading α-amylases have been of special interest in starch breakdown because of their specificity towards α-(1,4) and α-(1,6) linkages. Lka1p was functionally characterised, showing that it has the properties of an endo-acting enzyme with specificity to the α-1,4 and α-1,6 linkages of the glucose polymer. Lka2p α-amylase has specificity towards starch and dextrin. The expression of the *LKA1* and *LKA2* under the control of the phosphoglycerate kinase gene (*PGK1*) promoter-terminator *vs.* native promoters was studied previously. The expression of *LKA1* and *LKA2* singly and in combination showed that their expression in combination

resulted in a synergistic effect on starch degradation. Thus, previous studies have aptly demonstrated the efficiency of these enzymes in the *S. cerevisiae* host system (Eksteen *et al*., 2003b). This is further useful to exploit recombinant *S. cerevisiae* for starch bioconversion.

The scope of this thesis lies in the understanding, design and application of biocatalysts for starch hydrolysis, with emphasis on molecular characterisation and the development of improved α-amylases from *Lipomyces kononenkoae*.

2. PROJECT AIMS

The aims of this thesis are to characterise the molecular structural organisation of the α-amylases, Lka1p and Lka2p, from the yeast *Lipomyces kononenkoae* and to further understand the role of the structural domains and specific amino acids in protein function. This study aims to throw light on the structure-function relationship of the *L. kononenkoae* amylases, which could provide us with valuable information for the design of a novel hybrid enzyme containing the desirable properties of Lka1p and Lka2p. Apart from attempts to improve enzyme function by mutagenic techniques, the work also aimed at the expression of α-amylase Lka1p in a flocculent genetic background of *Saccharomyces cerevisiae* in order to develop whole-cell flocculent biocatalysts for raw starch fermentations.

The purpose of this study was two-fold:

(i) the molecular characterisation of Lka1p and Lka2p to further understand the structure-function relationships and role of specific amino acids in protein function, and the possible improvement of their substrate specificity in raw starch hydrolysis; and

(ii) the determination of the effect of yeast cell flocculence on the efficiency of starch fermentation and the possible development of high-flocculating, *LKA1*-expressing *S. cerevisiae* strains as 'whole-cell biocatalysts' and the production of high yields of ethanol from raw starch.

3. REFERENCES

Eksteen JM, Van Rensburg P, Cordero Otero RR, Pretorius IS. 2003b. Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the alpha-amylase and glucoamylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*. *Biotechnol Bioeng* **84(6)**:639-646.

Eksteen JM, Steyn AJC, Van Rensburg P, Cordero Otero RR, Pretorius IS. 2003a. Cloning and characterization of a second α-amylase gene (*LKA2*) from *Lipomyces kononenkoae* IGC4052B and its expression in *Saccharomyces cerevisiae*. *Yeast* **20**:69-78.

Horn CH, De Kock A, Du Preez, JC, Lategan PM. 1988. A comparative study of the amylolytic ability of *Lipomyces* and *Schwanniomyces* yeast species. *System Appl Microbiol* **10:**106-110.

Larson ED. 1993. Technology for electricity and fuels from biomass. *Ann Rev Energy Env* **18:**567-630.

Steyn AJC, Pretorius IS. 1995 Characterization of a novel α-amylase from *Lipomyces kononenkoae* and expression of its gene (*LKA1*) in *Saccharomyces cerevisiae*. *Curr Genet* **28:**526-533.

CHAPTER 2

LITERATURE REVIEW

Molecular and cellular biocatalysts for starch hydrolysis

1. INTRODUCTION

Starch is a major storage product and the bioconversion of starch into commercial products of interest is achieved by using starch hydrolases, which break down the starch substrate by hydrolysing the glycosidic linkages of the polymer. Since the market for starch hydrolases is growing rapidly and is expected to increase further, such enzymes have been the first targets of protein engineering. In parallel, the scientific community has made a great contribution towards the development and engineering of cellular biocatalysts. Most often, a natural enzyme catalyst is heterologously expressed in a microorganism and the whole recombinant cells could be further engineered for the optimal processing of raw materials. In this chapter, we will discuss the structure and chemical composition of raw cornstarch and the pre-treatment before catalysis by enzymatic attack. The different members of the starch hydrolases, their catalytic mechanism and their unique properties of interest for the starch industry from a protein-engineering perspective are reviewed. In addition, the use of whole cell biocatalysts and various strategies followed for optimal bioprocessing are also discussed in detail. This background information is relevant to our research, which deals with the understanding of the structure-function relationship of α-amylases LKA1 and LKA2 from *Lipomyces kononenkoae* in the context of designing novel enzyme catalysts for raw starch hydrolysis in the future. The review also provides a background study of strain development at the application level and of the optimal/economic use of the secreted enzymes in the bioprocessor.

2. CORN STARCH – STRUCTURE AND COMPOSITION

Corn contains 60-68% starch and is the most widely used cereal in dry milling operations. It is easy to process from cooking through to fermentation. A quantity of 0.025 tonnes of corn generally contains about 14.5 Kg of starch, which is present in the endosperm portion of the kernel in the form of granules. When hydrolysed, this starch yields about 16.5 of glucose. The mass of cornstarch increases as water is taken up during hydrolysis. The pie chart below (Figure 2.1) shows the abundance of cornstarch compared to other sources of starch.

Starch consists of two types of molecules, amylose, which generally makes up 20-30%, and amylopectin, which generally constitutes 70-80% (Figure 2.2). Both these types of molecules consist of polymers of α-D-glucose units in the $^4C_1$ conformation. In amylose, these are linked α-(1 ,4), with the ring oxygen atoms all on the same side, whereas in amylopectin, about one residue in every twenty or so is also linked α-(1,6), forming branch points. The relative proportions of amylose to amylopectin and α-(1,6) branch points depend on the

source of the starch, e.g. amylomaizes contain over 50% amylose, whereas maize starch contains as little as 3%. The amylose fraction of starch contributes to the viscosity of the solution after pre-treatment.

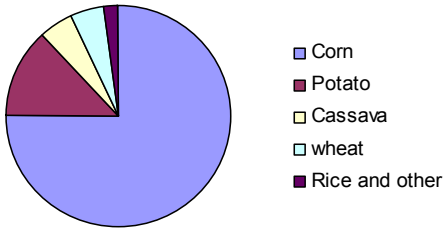
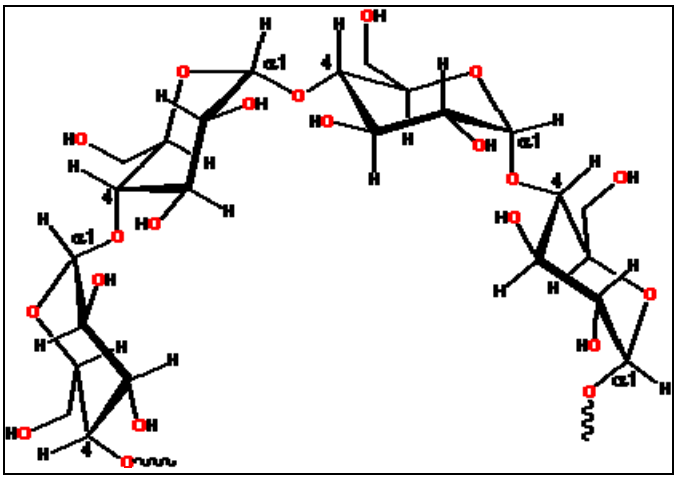


Fig. 2.1 Starch distribution worldwide by raw material: Statistics (1999-2001) from the International Starch Research Institute, Denmark

a)



b)

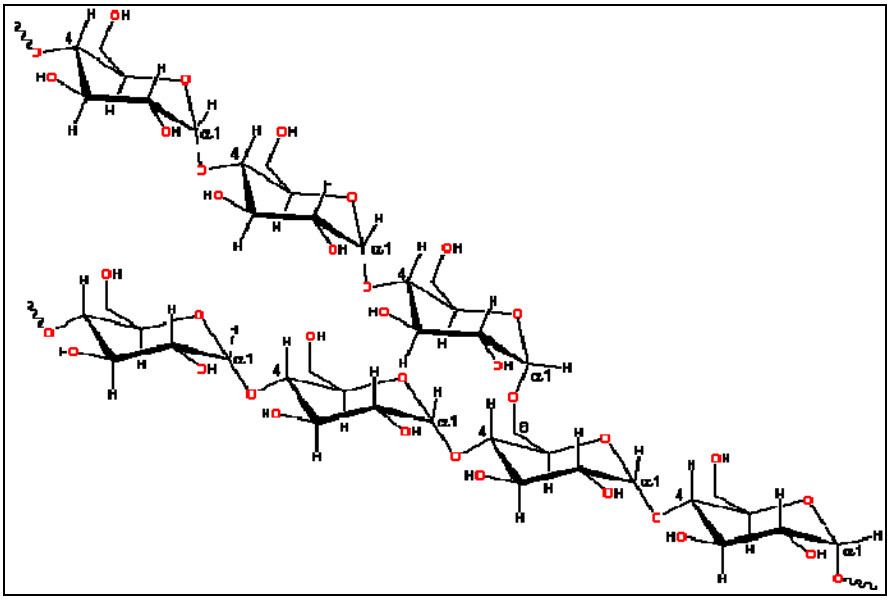


Fig 2.2. Structure of the principle components of starch: **a)** amylose **b)** amylopectin

2.1 Molecular structure of starch

Amylose and amylopectin are inherently incompatible molecules: amylose has a lower molecular weight with a relatively extended shape, whereas amylopectin has huge but compact molecules. Most of their structure consists of α-(1,4)-D-glucose units. Although the α-(1,4) links are capable of relatively free rotation around the (ϕ) and (φ) torsions, hydrogen bonding between the O3' and O2' oxygen atoms of sequential residues tends to encourage a helical conformation. These helical structures are relatively stiff and may present contiguous hydrophobic surfaces.

Amylose

Amylose molecules consist of single, mostly linear chains with 500-20 000 α-(1,4) D-glucose units, dependent on the source. A few α-(1,6) branches and linked phosphate groups may be found, but these have little influence on the molecule's behaviour. Amylose can form an extended shape (hydrodynamic radius 7-22 nm), but generally tends to wind up into a rather stiff left-handed single helix (Immel et al., 2002) or form even stiffer parallel left-handed double helical junction zones (Figure 2.3a). Single helical amylose has hydrogen-bonding O2' and O6' atoms on the outside surface of the helix, with only the ring oxygen pointing inwards. The aligned chains may then form double-stranded crystallites that are resistant to amylases. These possess extensive inter- and intra-strand hydrogen bonding, resulting in a fairly hydrophobic structure of low solubility.

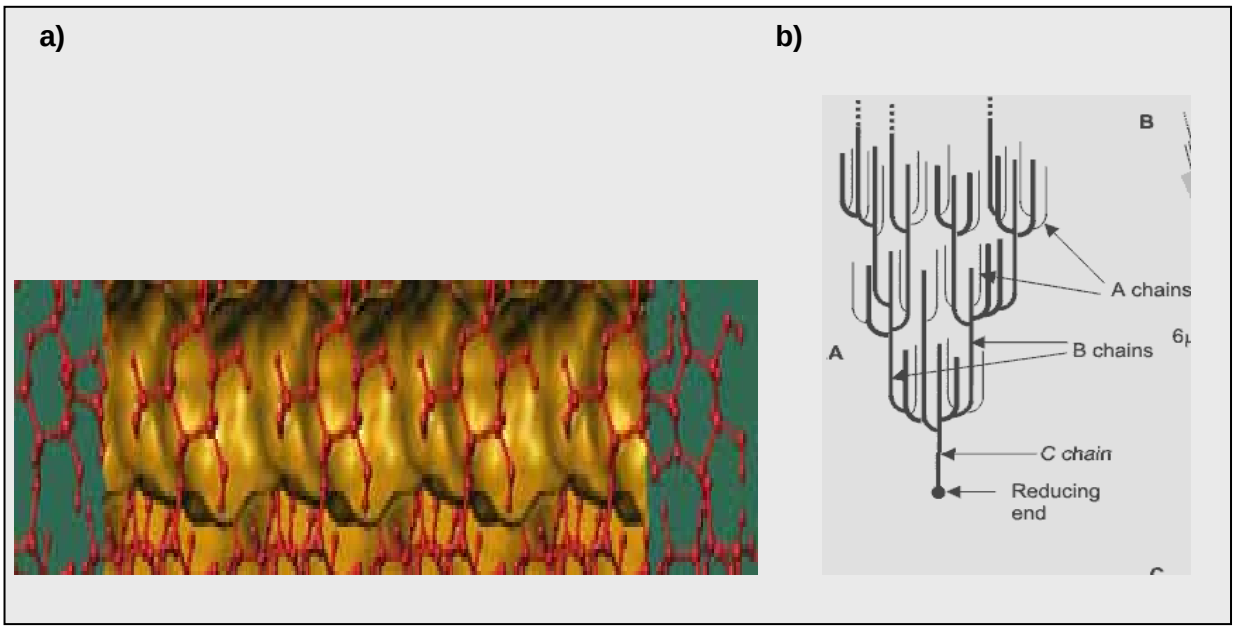


Fig. 2.3 Structure of **a)** amylose and **b)** amylopectin

Amylopectin

Amylopectin is formed by the non-random α-(1,6) branching of the amylose-type α-(1,4) D-glucose structure. This type of branching is determined by branching enzymes that leave each chain with up to 30 glucose residues. Each amylopectin molecule contains a million or so residues, about 5% of which form the branch points. There are usually slightly more 'outer' unbranched chains (A chains) than inner branched chains (B chains). There is only one chain (the C chain) that contains the single reducing group (Figure 2.3b). Each amylopectin molecule contains up to two million glucose residues in a compact structure with a hydrodynamic radius of 21-75 nm The molecules are oriented radially in the starch granule and, as the radius increases, so does the number of branches required to fill up the space, with the consequent formation of concentric regions of alternating amorphous and crystalline structure.

In simpler terms, starch principally made up of amylose and amylopectin is a polymer of glucose linked by C1 oxygen, known as the glycosidic bond. This bond is stable at a high pH, but hydrolyses at low pH. A latent aldehyde group is present at the end of the polymeric chain. The breakdown of starch is accomplished by chemical and biological means. In industrial sugar platforms, starch from corn is hydrolysed to release glucose for the production of fuel ethanol and other chemicals. In the early 19th century, starch processing was done by chemical methods such as acid hydrolysis, but today this has been replaced by biocatalysts that can hydrolyse glucose from starch much more efficiently and cost effectively. When compared to their chemical counterparts, biocatalysts are exquisitely selective and highly reactive over a broad range of operating conditions.

3. PRE-TREATMENT PROCESS

Cornstarch is converted to ethanol by dry or wet milling, followed by saccharification and fermentation. This is the foremost bioreactor fermentation technology in use today. In order to use starch for enzymatic breakdown and fermentation, it is necessary to perform pre-treatment procedures that make this complex substrate more accessible to biological attack. The industrial processing of corn starch is consolidated in Figure 2.4. Dry or wet milling of the corn kernels allows the breakdown of cereal grains to a particle size that is as small as possible to facilitate the subsequent penetration of water during the cooking process.

After the grains are milled, the cooking of the cereal stock begins by mixing water with the stock. The purpose of cooking and saccharification is to achieve the hydrolysis of starch to fermentable sugars. In order for the enzymes to bring about the hydrolysis of starch to dextrins, the granular structure of starch must first be broken down in a process called

gelatinisation. When the slurry of cornmeal and water are cooked, the starch granules start to adsorb and swell. They gradually lose their crystalline structure until they become large, gel-filled sacs that tend to fill all of the available space and break with agitation and abrasion. The peak of gelatinisation is when the mash reaches a maximum viscosity.

3.1 Premixing, cooking and liquefaction

A variety of batch and continuous cooking systems are available to process the milled grains. In the batch system, only one tank is used, which serves as a slurrying, cooking and liquefaction vessel. Steam jets are installed in the vessel to bring the mash to boiling temperature, along with cooling coils to cool the mash for liquefaction. In the batch cooking system, a small quantity of α-amylase is traditionally added at the beginning of the process to facilitate agitation in the high-viscosity stage of gelatinisation. Generally, this requires the α-amylase to be highly thermostable. After boiling, the mash is cooled to a temperature of 75-90°C and the second addition of α-amylase is done. Liquefaction takes place, usually over a holding period of 45-90 minutes. The pH range for efficient α-amylase usage is 6-6.5. The mash pH should therefore be controlled in this range from the addition of the first enzyme until the end of liquefaction. Practically, an α-amylase with activity at a wider range of pH conditions would make the process more efficient, considering the dynamics and chemical composition of the raw materials. The glucoamylases usually have a pH range of 4.0-5.5, thus it is necessary to adjust the pH after liquefaction by using sulphuric acid.

In continuous cooking processes, meal, water and nutrients are continually fed into the premix tank at a temperature just below that of gelatinisation. The mash is pumped continuously through a jet cooker, where the temperature is instantly raised to 120°C. With plug flow, the mash moves down through a vertical column for 20 min and then passes into the flash chamber for liquefaction at 80-90°C. High temperature-tolerant α-amylase is added to the vessel to bring about liquefaction. The retention time in the liquefaction/flash chamber is usually 30 min. The pH from slurrying through the liquefaction vessel must be controlled within the 6.0-6.5 range. The greatest advantage of this system is that no enzyme is needed during the slurrying stage, accounting for significant savings in enzyme usage. However, an appropriate α-amylase that is highly thermostable is an essential criterion. From the liquefaction chamber, the mash is pumped through a heat exchanger to be cooled for saccharification and fermentation.

One purpose of the cooking process is to cleave the hydrogen bonds that link the starch molecules, thus breaking the granular structure and converting the molecules into a colloidal suspension. However, most hydrolysis takes place during the liquefaction stage and all cooking systems employ enzymes during this stage. For the simultaneous saccharification

process (SSF), a saccharifying enzyme is added directly to the process. The use of Rhizozyme^{TM} , an enzyme found to be particularly favourable in SSF processes, has industrial advantages, such as optimal activity in the fermenter and side activities that assist in releasing simpler sugars from the complex substrates provided. α-Amylases are extracted from animal (usually pancreatic or salivary), cereal (usually wheat or barley) (Thoma *et al*., 1971), fungal (usually derived from large-scale fermentations of *Aspergillus* species), and bacterial (derived from similar fermentations of *Bacillus* species) (Takagi *et al*., 1971). In many countries, the plant amylases are used for conversion of barley malt into simple sugars. During saccharification, the mash, which has been cooled to 60-65°C, is transferred to a liquefaction vessel, where glucoamylase is added. This exo-enzyme starts hydrolysing the dextrin from the non-reducing end of the molecule, progressively releasing glucose. Liquefaction occurs at a temperature near 90°C and a pH range of 6-6.5, which is not applicable for saccharification. The pH for saccharification must be between 4.0 and 5.0 and the optimum temperature for the activity of glucoamylase is 75°C, varying according to the stability of the native enzymes used.

The hydrolysis of starch itself is a dynamic process and the choice of catalysts is therefore crucial in process development. In the following sections, details of molecular and cellular biocatalysts used for starch hydrolysis and strategies for the improvement of these catalysts for starch hydrolysis are presented.

4. WHOLE-CELL BIOCATALYSTS IN STARCH FERMENTATIONS

Conversion technologies for the production of energy from starchy biomass can be classified as biological (fermentation) or thermal (burning, pyrolysis, gasification). In the present day, the fermentation of sugars to ethanol is the best established process for the conversion of biomass to energy. Most yeasts are capable of converting hexoses via glycolysis into pyruvate and, subsequently, decarboxylate pyruvate into acetaldehyde. To maintain a redox balance, acetaldehyde is further reduced to ethanol. The regulation of fermentation metabolism on the one hand, and respiration on the other, is diverse and complex; the main determinants are the concentration of oxygen and the fermentable carbon source.

The industrial biotechnology processes using microorganisms are based on the exploitation of the cells in the fermentation medium during the process. The classical fermentation processes suffer from various constraints, such as low cell density, nutritional limitations and batch mode operations with longer downstream processing times. There are several microbes, bacteria, fungi and yeasts available in nature that secrete enzymes that catalyse the hydrolysis of starch. However, the microbes that function extremely well in their natural habitat are not always suitable in the fermentation conditions that are devised in the industry.

Both the physicochemical properties of the cells and the physicochemical environment in the fermenter are of crucial value in determining a productive bioconversion process.

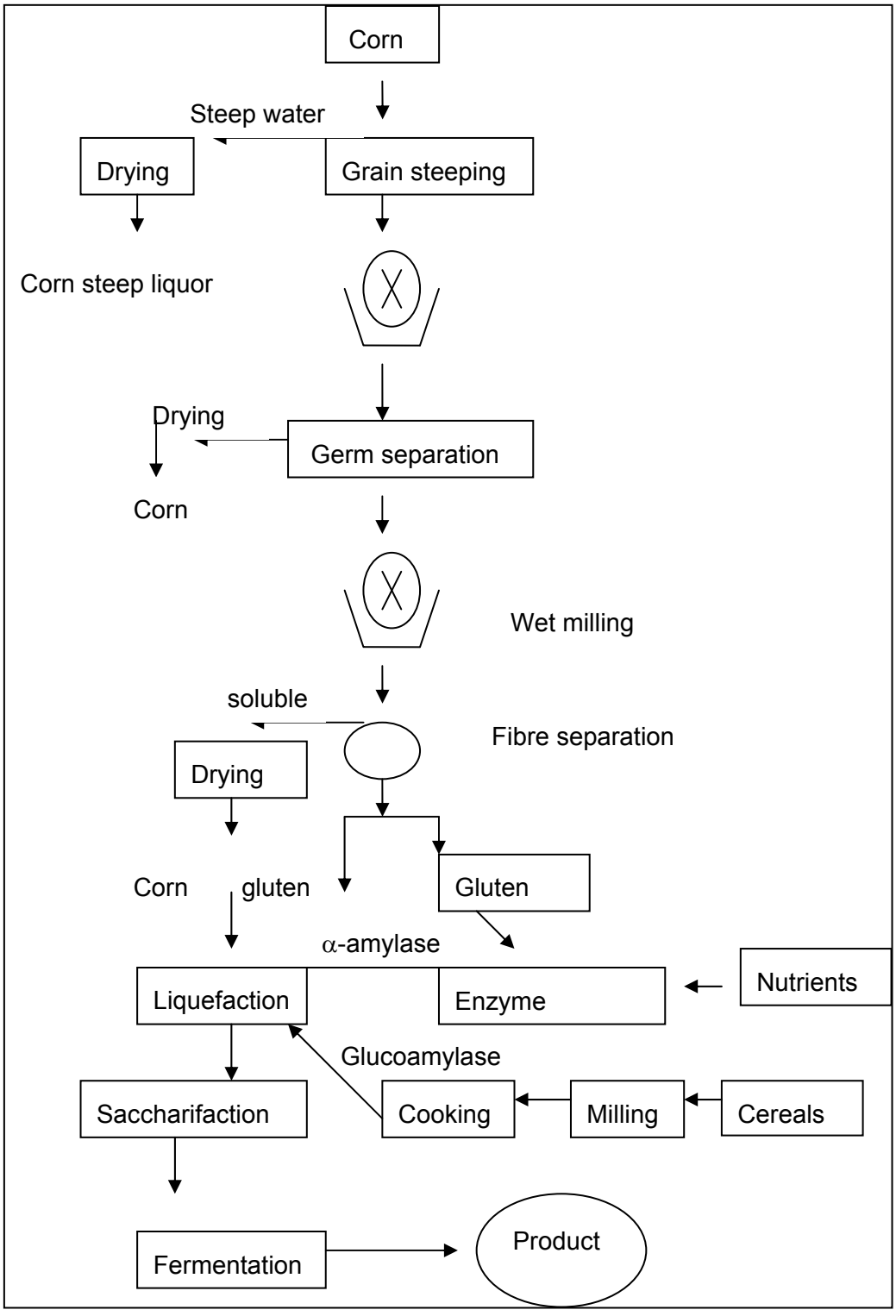


Fig. 2.4 Schematic representation of the industrial processing of cornstarch

It has also been difficult to isolate microorganisms that are active at the temperature, pH and chemical conditions used in industrial bioprocesses. If an organism was isolated from a psychrophilic or thermophilic environment, or from a halophilic or osmophilic environment, there is a chance that the use of this microbe could be unsuccessful, since adaptability is not always an easily understood phenomenon. As an alternative, the amylolytic enzyme that is produced is extracted and purified to be used in the fermentation, but this is very expensive, time consuming and tedious. Consequently, there are numerous limitations to using naturally occurring microbes directly in industrial processes. Often, the microbe that is used produces not only the desired product, but also many other by-products and metabolites, which might not be conducive in upstream and downstream processing in the fermentations. However, with the advent of recombinant DNA technology and genetic engineering, it has been possible to isolate genes from the natural amylolytic organism and to clone in a desired host, such as *Saccharomyces cerevisiae*, which has suitable fermentation properties.

4.1 *Saccharomyces cerevisiae* – advantages and disadvantages in sugar conversion

Saccharomyces cerevisiae has been the most conventional and widely used fermentation microorganism. It offers several advantages over other yeasts in the bioconversion of sugars. Under excess carbon conditions, its metabolic flux to ethanol is hardly affected by the presence of oxygen (Lagunas, 1979), it is able to grow under strict anaerobiosis (Visser, 1995) and it has a high ethanol tolerance, amounting to 150 g/L ethanol. Apart from characteristically well understood fermentation properties, it is the best-studied eukaryote at the molecular level and, over the past few decades, a wealth of expertise has been accumulated, both in the fermentation technology and the basic genetics of this organism. Further advantages of using *S. cerevisiae* as a host for gene cloning and gene expression lie in its non-pathogenic character, its secretion proficiency, its glycosylation potential and its usefulness as a eukaryotic model organism to study features such as cell cycle division, mitosis, transcription factor activity, etc. On a result of the aforementioned reasons, the production of ethanol by *Saccharomyces cerevisiae* is unchallenged by other yeasts, fungi and bacteria. However, saccharolytic and amylolytic properties found in other ethanol-producing microbes may result in efficient ethanol production. Nevertheless, it is to be noted that the ethanol production rate of other yeast species, such as *Zymomonas mobilis*, *Pichia stipitis*, *Candida shehatae*, etc., with glucose as substrate is at least five times lower than observed in *S. cerevisiae* (Hahn-Hagerdal *et al*., 1994).

The main limitation to the use of *S. cerevisiae*, however, is its inability to convert relatively inexpensive polysaccharide-rich substrates, such as starchy biomass, to industrially important commodities owing to the lack of amylolytic enzymes. In addition they are not

useful in pentose conversion. To overcome this drawback, genes encoding amylolytic enzymes have been cloned to create recombinant amylolytic *S. cerevisiae* strains.

4.2 Heterologous expression of amylolytic enzymes in *Saccharomyces cerevisiae*

The production of amylolytic enzymes from naturally-occurring microbes is usually applied for for commercial exploitation. However, impressive improvements using biological tools have led to process development for higher yields in commercial production. Molecular biology thus provides an unparalleled ability to manipulate genes individually or in combination. The cloning and expression of heterologous amylolytic enzymes in *S. cerevisiae* involves the identification of amylolytic enzymes, the isolation of genes encoding these enzymes, and their cloning and expression. Such methods have led to the amplification of the useful characteristics of *S. cerevisiae* and the annihilation of unfavourable ones to modify the genetic makeup of the production strains. Thus, the bioconversion of starch by *S. cerevisiae* through the recruitment of heterologous sequences specifying amylolytic enzymes has been the subject of several studies.

Among starch hydrolases, two major enzymes, α-amylases and glucoamylases, in conjunction can efficiently hydrolyse starch. While α-amylases act on the endo-linkages of starch, the glucoamylases release glucose from the reducing end, and this can then be further processed by *S. cerevisiae*. The optimal expression and secretion of these two enzymes in *S. cerevisiae* is therefore central to the efficient one-step conversion of starch to ethanol. Many such recombinant strains have been constructed, and the co-expression of *Bacillus amyloliquefaciens* α-amylase and the *Saccharomyces diastaticus* glucoamylase has shown to improve the amylolytic efficiency of the recombinant strains (Shibuya *et al*., 1992b). The heterologous expression of α-amylases and starch-degrading enzymes from other yeast species, such as *Lipomyces kononenkoae* (Lka1p and Lka2p) and *Saccharomycopsis fibuligera* (Sfa1p and Sfg1p), has been reported (Eksteen *et al*., 2003). Co-expression of the α-amylase and glucoamylase genes was shown to enhance starch degradation additively in *S. cerevisiae*. The strain expressing *LKA1* and *LKA2* resulted in the highest levels of α-amylase in liquid media and, when used in small-scale batch fermentations, utilised 80% of the available starch, producing 0.61 g/100 mL of ethanol after six days of fermentation (Eksteen *et al*., 2003).

However, the conversion of complex raw starch requires catalytic activity on both α-1,4 and α-1,6 linkages to hydrolyse the branching points in these molecules. The α-amylase gene, *AMY1*, from *Bacillus amyloliquefaciens,* the glucoamylase gene, *STA2*, from *S. cerevisiae var. diastaticus*, and the pullulanase gene, *PUL1*, from *Klebsiella pneumoniae* individually and jointly expressed in (Pretorius *et al*., 1986; Pretorius and Marmur, 1988; Steyn and

Pretorius 1991; Janse and Pretorius, 1993, 1995) industrial strains of *S. cerevisiae* and the additive effect of these enzymes led to the efficient digestion of soluble and raw starch.

The expression of a bifunctional protein from a fusion of the complete reading frames of both α-amylase and glucoamylase cDNAs from the filamentous fungus *Aspergillus shirousamii* in *S. cerevisiae* was a more effective approach (Shibuya *et al.,* 1992a). This strain displayed a higher level of activity on raw starch substrate than the mixture of the two native enzymes. Later, De Moraes *et al*. (1995) prepared eight different constructs, including strains that produce *Bacillus subtilis* α-amylase (BSAAase), mouse pancreatic α-amylase (MAAase) or *Aspergillus awamori* glucoamylase (GAAse) either singly or in combination, as well as strains that produce either BsAAase/Gaase or MAAase/GAAse fusion enzymes.

In order to achieve high productivity levels efficient expression and considerably high levels of heterologous protein yield are essential. The choices of expression in the host system for improved protein yields include multicopy episomal expression or multiple integrations into the genome, which offers higher mitotic stability than the former. Lopes *et al*. (1996) described a yeast integrative vector that has the favourable properties of high mitotic stability and high copy number. The plasmid contains a portion of ribosomal DNA from *S. cerevisiae* that allows it to be targeted to the genomic rDNA locus. Using this approach, *Aspergillus awamori* glucoamylase gene *(glu)* was expressed in *S. cerevisiae* and the integrant, G23-8, could consume 82% of the soluble starch supplied (Lin *et al*., 1998). In a similar study, the recombinant yeast producing isoamylase and glucoamylase by multiple integration of the *Pseudomonas amylodermosa iso* gene into the integrant G23-8 chromosome led to the utilisation of 95% of the soluble starch provided (Ma *et al*., 2000).

Higher expression levels do not always result in higher protein yields in the extracellular media. One of the most obvious reasons for lower yields is the abundant production of proteases by *S. cerevisiae* that target the heterologous proteins. A major improvement addressing this problem is the “carrier approach”, which is based on the idea that a secreted protein can act as a carrier for the more efficient secretion of a heterologous protein (Punt *et al*., 2002). Thus, the use of yeast mating pheromone MFα has been used as a secretion signal for heterologous proteins in *S. cerevisiae.* There are many examples that illustrate the successful use of an MFα secretion signal: our own research (this thesis) in which MFα with a synthetic spacer peptide (Kjeldsen *et al*., 1996) was used for the effective secretion of the α-amylase LKA1 in *S. cerevisiae* affirms this to be a promising strategy.

Further, in an attempt to exploit the secreted enzymes efficiently for hydrolysis and the production of ethanol, various strategies, such as the cell surface display of amylolytic enzymes on flocculent yeast cells, mixed culture fermentations and the use of immobilised whole cell biocatalysts, have been applied in starch fermentations. One of the approaches

used to deal with the problem of heterologous expression is the co-culturing of an amylolytic microorganism with *S. cerevisiae*. Thus, a simultaneous single-step system for the enhanced fermentation of starch to ethanol could be achieved by using symbiotic co-cultures of *Aspergillus* species, which hydrolyse starch to glucose, and *S. cerevisiae*, which is non-amylolytic but efficiently ferments glucose to ethanol. Since *S. cerevisiae* utilises the breakdown products from starch at a faster rate, feedback inhibition of the enzyme by the products is lowered in this process.

Alternative methods, such as the use of raw substrate and an enzyme with *S. cerevisiae*, have been devised. A simultaneous saccharification and fermentation process (SSF) was used on raw wheat flour and the production of ethanol was attained by using amyloglucosidase and *S. cerevisiae* in combination (Montesinos and Navarro, 2000). This processing effectively reduced the operational costs during upstream processing.

4.3 Flocculent biocatalysts

Among the novel methodologies used in consolidated bioprocessing, flocculation is an interesting phenomenon for ethanol production, because the time of occurrence and intensity of flocculation can affect the degree of fermentation and, therefore, the alcohol level (Masy *et al*., 1992). It is also one of the important characteristics of yeast strains used in industrial applications such as brewing and the production of recombinant proteins, because the flocculent yeast can easily be separated during downstream processing (Kondo *et al*., 2002).

Kondo *et al*. (2002) demonstrated the use of flocculent yeast cells displaying glucoamylase for efficient, direct ethanol production from soluble starch, essentially due to the high ethanol production rate (0.71 g h^{-1} l^{-1}) obtained. Ideally, a large cell mass should be obtained by high cell density culture under aerobic conditions, and the cells harvested could be used for ethanol fermentations. Immobilising an enzyme on the surface of the yeast has been carried out so that whole cells can be used directly for ethanol fermentations. Successful reports have been made in this context for example, cell surface display (Shigechi *et al*., 2004) of glucoamylase on flocculent *S. cerevisiae* (Kondo *et al*., 2002) and cell surface display systems using Flo1p of different anchor lengths (Sato *et al*., 2002).

In our current study, it was clear that the flocculent phenotype interacts with starch to better reduce the dilution of the enzyme produced into the surrounding media and thereby leading to earlier hydrolysis. However, immobilisation of the flocculent cells onto a solid surface was found to offer higher productivity levels.

4.4 Cell immobilisation

As in flocculation, the use of immobilised cells often eliminates the tedious, time-consuming and expensive steps involved in the isolation and purification of intracellular enzymes. It also tends to enhance the stability of the enzyme by retaining its natural catalytic surroundings during immobilisation and the subsequent continuous operation. The ease of conversion of batch processes into continuous mode and the maintenance of a high cell density without washout conditions, even at very high dilution rates, are among the many advantages of immobilised cell systems. Immobilisation thus creates a favourable microenvironment for the enzyme to act on the substrate by reducing the time of diffusion, as well as the dilution rate of the enzyme.

Immobilisation is usually achieved by the use of high molecular hydrophilic polymeric gel, such as alginate, carrageenan, agarose, etc. In these cases, the cells are immobilised by entrapment in the pertinent gel by a drop-forming procedure. When traditional fermentations are compared to immobilised cells, the productivity obtained is considerably higher in the latter, obviously due to the high cell density and immobilisation-induced genetic modifications. The use of passively immobilised cells of flocculent yeast *S. diastaticus* in batch fermentations resulted in 70% higher productivity than in the batch fermentations without immobilisation. Alternatively enzyme immobilisation is also advantageous. It evident that the use of co-immobilised β-amylase and pullulanase led to a reduction in the saccharification time of starch and an increase in maltose yield (Atia *et al*., 2003). Recent reports on the enhanced plasmid stability of genetically modified microorganisms under immobilised conditions and the viability of microbes for longer time periods under entrapped conditions are among the advantages for many potential new applications of immobilised cells. Thus, whole-cell immobilisation has been used as a tool to intensify microbiological processes that are well established.

For application in biocatalysis, process methods may be performed in three different ways:

a) Isolated enzymes that are soluble can be partially purified from cells and used in a reaction mixture as free enzyme or immobilised on a solid carrier, which allows easy recovery for repeated usage.
b) Whole cells that are intact cells removed from the growth medium but that still maintain most of their metabolic activities. This is the most useful during the use of enzymes not amenable to isolation or if the use of cells would make the process cost effective.
c) Active fermentation, whereby actively growing cells are sometimes used for enzyme reactions. This system is only used if the components in the fermentation medium do not interfere with the enzyme reaction. The addition of substrate to the fermentation broth is the most direct means of biocatalysis.

5. ENZYMATIC SACCHARIFICATION OF STARCH

In general, enzymes have been used in industrial processes for a long time. In comparison to other biomolecules that function by binding, e.g. for antibodies or receptors, enzymes are amazingly active. They overcome energy barriers of more than 24 Kcal/mole and have a complex mechanism of action (Griffiths and Tawfik, 2000). In the context of starch conversion, there has been a shift during the past few decades from acid hydrolysis to the use of starch-converting enzymes in the production of maltodextrins, modified starches, glucose syrups, sweeteners, etc. Currently, these starch-converting enzymes from nature comprise about 30% of the world's enzyme production. Besides starch conversion, these enzymes are also used in a number of industrial applications, such as laundry and porcelain detergents or as anti-staling agents in baking (Van der Maarel *et al*., 2002).

The engineering of enzymes has been possible in principle since the advent of genetic engineering, and studies involving the site-specific mutagenesis of enzymes for both reengineering and purely investigatory purposes have produced significant data for the development of improved starch hydrolases. This family of enzymes consists of over 30 different specificities and this feature makes the protein-engineering studies of these enzymes highly interesting and productive from an industrial perspective. The following sections summarise a detailed overview of the enzymes represented in the α-amylase family, and some developments in improving the properties of α-amylases.

5.1 The α-amylase family: members, structure and catalytic mechanism

α-Amylases were originally recognised as a group of starch hydrolases and related enzymes that exhibit clear sequence similarities and a predicted common super secondary fold, a parallel $(\beta/\alpha)_8$ barrel (Farber and Petsko, 1990; Faber, 1993). The α-amylase family, GH-H clan of glucosyl hydrolases, is the largest family of glucoside hydrolases, transferases and isomerases, comprising over 30 different specificities. Takata *et al*. (1992) explained the concept of the α-amylase family and confined this classification to enzymes that satisfy the following requirements: 1) they act on α-glucosidic linkages; 2) they hydrolyse or form α-glucosidic linkages by transglycosylation reactions; 3) their amino acid sequence consists of four conserved regions; and 4) they contain Asp, Glu and Asp residues corresponding to the Asp^{206}, Glu^{230} and Asp^{297} of Taka-amylase A.

Currently, the α-amylase family constitutes the clan of GH-H families 13, 70 and 77, and the enzymes of these different families can operate on α-1,1-, α-1,2-, α-1,3- and α-1,5-linkages, as well as on α-1,4- and α-1,6-glucosidic linkages. Many specialised reports are available that focus on the protein engineering of the individual enzymes and certain groups of this

family, and on the possibilities of engineering their specificities and properties to meet industrial needs. It is possible to categorise starch-converting enzymes into four broad groups: 1) endoamylases; 2) exoamylases; 3) debranching enzymes and 4) transferases.

Endoamylases: Endoamylases are able to cleave α-1,4-glycosidic bonds that are present internally in the amylose or amylopectin chain. Classically, α-amylase (EC 3.2.1.1) belongs to this group of enzymes. The end products of α-amylase action are oligosaccharides of varying length with an α-configuration and α-limit dextrins that constitute branched oligosaccharides.

Exoamylases: Enzymes belonging to the group of exoamylases either cleave α-1,4 bonds such as β-amylase (EC 3.2.1.2), or α-1,4 and α-1,6 bonds such as amyloglucosidase or glucoamylase (3.2.1.3) and α-glucosidase (EC 3.2.1.20). Exoamylases thus act only on the external bonds and therefore liberate only glucose residues (as in the case of glucoamylases and α-glucosidases), or maltose and β-limit dextrin (β-amylases). In addition, β-amylases and glucoamylases can also convert the anomeric form of maltose from α to β. Glucoamylases prefer long-chain polysaccharides, while α-glucosidase are known to function best on short-chain saccharides. The enzymes catalysing transglycosylation reactions, namely cyclodextrin glycosyltransferases (EC2.4.1.19), maltogenic amylases (referred to as glucan-1,4-α-D-glucanohydrolase (EC 3.2.1.133)) and some maltooligosaccharide-forming α-amylases (EC 3.2.1.60 and EC 3.2.1.98), are also grouped as exoamylases.

Debranching enzymes: Debranching enzymes exclusively break the α-1,6 glycosidic linkages. Isoamylases (3.2.1.68) and pullulanases (3.2.1.41) fall under this category. The major difference between isoamylase and pullulanase is that pullulanases can hydrolyse the α-1,6 bonds of pullulan and amylopectin, while isoamylase can only hydrolyse the α-1,6 bonds of the amylopectin chain. There are also a number of pullulanase-type enzymes that hydrolyse α-1,4 and α-1,6 linkages. These enzymes, which are most commonly referred to as amylopullulanases, produce maltose and maltotriose as major end products. A unique enzyme belonging to this group is neopullulanase, which can also transglycosylate to form α-1,4 and α-1,6 bonds (Takata *et al*., 1992).

Transferases: Transferases cleave an α-1,4 glycosidic bond of the donor molecule and transfer part of the donor to a glycosidic acceptor, with the formation of a new glycosidic bond. Enzymes such as amylomaltases (EC 2.4.1.25) and cyclodextrin glycosyltransferases (EC 2.4.1.19) form α-1,4 glycosidic bonds, whilst branching enzymes (EC 2.4.1.18) form α-1,6 glucosidic bonds. Amylomaltases perform transglycosylation to form a linear chain product, while cyclodextrin glycosyltransferases catalyse a similar reaction to form cyclic products.

The enzyme commission classification of enzymes, on the other hand, is based on naming the reactions they catalyse. Each member indicated above is assigned an EC number that uniquely identifies the reaction. For the α-amylases, this number is 3.2.1.1 and the reaction is described as the "endohydrolysis of α-1,4-glucosidic linkages in oligosaccharides and polysaccharides". Enzymes that catalyse similar reactions are classified under different EC numbers (Table 2.1). For example, the cyclodextrin glucanotransferases for have 2.4.1.19 as their EC number, although they are structurally and enzymatically similar to the α-amylases, α-glucosidases (3.2.1.20), maltogenic α-amylases (3.2.1.133) and maltotetrose-forming amylases (3.2.1.60). These enzymes are not classified with α-amylases because they possess properties of exoamylases. The Carbohydrate Active Enzymes database (CAZy) (Coutinho and Henrissat, 1999) consists of enzymes with α-amylase activity classified into two structurally different glycosyl hydrolases, namely families 13 and 57. Family 13 consists of 514 sequences and 19 different enzyme activities. Among these are the CGTases, the α-glucosidases and the maltotetrose-forming amylases. Family 57, on the other hand, has 13 sequences and only two different enzyme specificities, namely α-amylase activity (EC 3.2.1.1) and 4-α-glucanotransferase activity (EC 2.4.1.-). None of the members have X-ray crystallographic structures and they are less characterised than family 13.

5.2 Structural architecture of α-amylases

The X-ray crystallographic structures determined for α-amylases from various origins show that most α-amylases possess a multidomain architecture with three major domains, commonly denoted as A, B and C. A central $(\beta/\alpha)_8$ barrel domain, which forms the core of the molecule, is referred to as domain A. This domain consists of $(\beta/\alpha)_{8-}$ or a TIM (Triose phosphate isomerase) barrel catalytic domain with a highly symmetrical fold of eight inner parallel β-strands surrounded by eight helices (Svensson, 1994), a motif recognised in at least nine families of glucosyl hydrolases (Davies *et al*., 1997).

Table 2.1 Summary of different members of the α-amylase family (MacGregor *et al*., 2001)

Structural information on the α-amylases family of enzymes	
Enzyme	EC number
a) Enzymes of known three-dimensional structure	
α-Amylase	3.2.1.1
Oligo-1,6 glucosidase	3.2.1.10
Maltotetraohydrolase	3.2.1.60
Isoamylase	3.2.1.68
Maltogenic amylase	3.2.1.133
Neopullulanase	3.2.1.135
Malto-oligosyltrehalose trehalohydrolase	3.2.1.141
Amylosucrase	2.4.1.4
Cyclodextrin glucanotransferase	2.4.1.19
Amylomaltase	2.4.1.25
b) Enzymes predicted to belong to the α-amylase family	
α-Glucosidases	3.2.1.10
Pullulanase (Limit dextrinase)	3.2.1.41
Amylopullulanase	3.2.1.1/41
Cyclomaltodextrinase	3.2.1.54
Dextran glucosidase	3.2.1.70
Trehalose-6-phosphate hydrolase	3.2.1.93
Maltohexohydrolase	3.2.1.98
Maltotriohydrolase	3.2.1.116
Maltopentohydrolase	3.2.1.-
Branching enzyme	2.4.1.18
Glucan debranching enzyme	2.4.1.25/3.2.1.33
Maltosyl transferase	2.4.1.-
Dextran sucrase/ Alternansucrase	2.4.1.5/2.4.1.140

Despite the diversity of its catalytic actions, the active site is always located at the C-terminal end of the barrel structures (Farber and Petsko, 1990). Domains B and C are located roughly at opposite ends of this barrel. The loops that link the β-strands of the adjacent helices usually carry the active site amino acids; some of these loops may be long enough to be considered as domains and thus, in most cases, domain B is formed by the protrusion between the third strand and the third helix of the TIM barrel (Farber and Petsko, 1990). This loop has an irregular structure that varies from enzyme to enzyme. In most, but not all, α-amylases, the catalytic domain (domain A) typically occurs at the N terminus of the protein. However, in some distinct members, the catalytic domain is preceded by an extra sequence – a domain whose role is uncertain. An overview of structural organisation in α-amylases is shown in Figure 2.5 and the structure of individual domains is presented in Figure 2.6. Several members of the family also contain more than one N-terminal domain preceding the barrel (Jespersen *et al*.,1991). These domains have occasionally been referred to as the N domain, although they are not structurally related in all the enzymes that possess them.

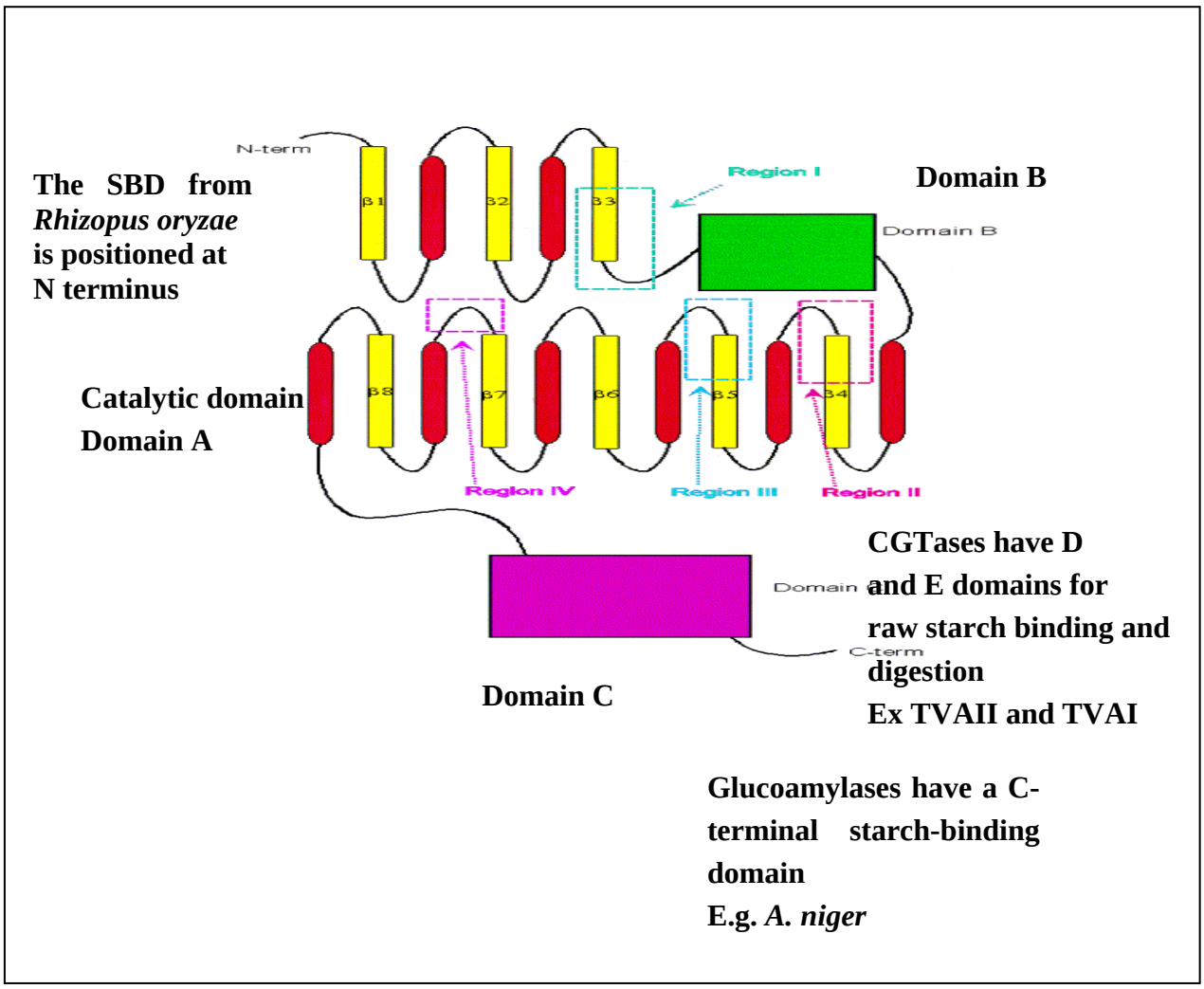
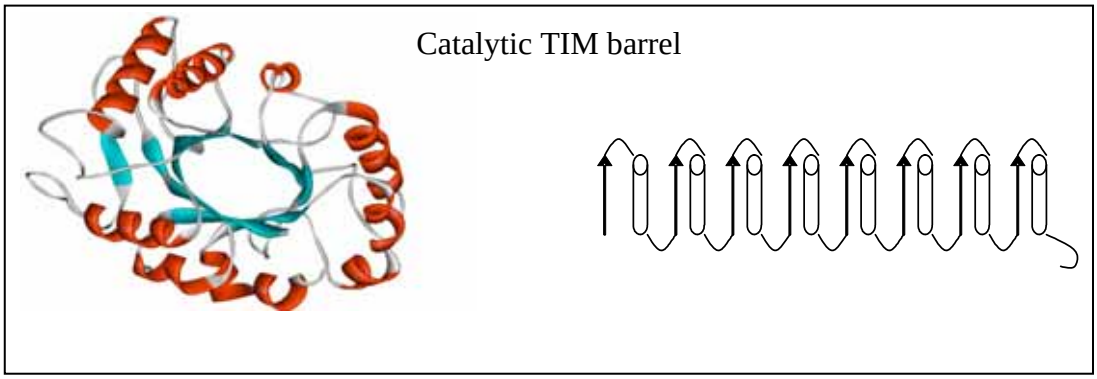


Fig. 2.5 Structural organisation in the α-amylase family

Domain C forms the C-terminal part of the protein and is a β-sandwich domain containing a Greek key motif and is thought to stabilise the catalytic TIM barrel by shielding the hydrophobic residues of domain A from the solvent. This beta-sandwich structure, which is characteristic of this family, is absent in the amylomaltase from *Thermus aquaticus* (Przylas *et al*., 2000). Some enzymes have domains D and E following domain C. If the enzymes possess domain D and E, they do not normally contain domain N. While the function of domain D is not certain, domain E, specifically in cyclodextrin glucanotransferases, has received much attention due to its raw starch-binding function, which facilitates the degradation of starch granules by enzymes possessing this domain. However, in a few cases, the starch-binding domain is present at the N terminus of the protein, or the so-called domain N functions as the raw starch-binding domain (Ashikari *et al.,* 1986; Abe *et al*., 2004).

5.3 Raw starch-binding domains

The starch-binding domain (SBD) is a functional domain that can bind to raw starch and thus enhance the ability of the enzyme to digest granular starch at the catalytic site. The SBD is present in a number of amylolytic enzymes of the glycoside hydrolase families. The presence of SBD in an amylolytic enzyme is closely connected to the enzyme origin.



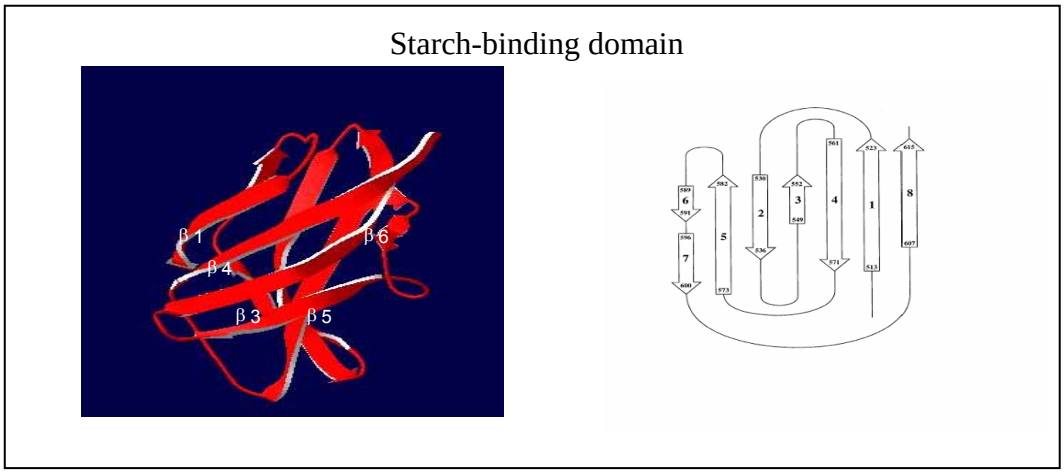


Fig. 2.6 Structure of the different domains of α-amylase

Only microorganisms, in particular filamentous fungi, gram-positive bacteria (firmicutes), proteobacteria, archaea and actinomycetes are known to produce α-amylases containing an SBD (Janecek *et al*., 2003). The SBDs are approximately 100 amino acids long and affixed at the C-terminal end, with the exception of an N-terminal SBD in *Rhizopus oryzae* glucoamylase. The well-studied starch-binding domain of *Aspergillus niger* glucoamylase consists of an $(\alpha/\alpha)_6$ barrel-fold structure (Sorimachi *et al*., 1997). Clustal W analyses of C terminal regions of α-amylases reveal the conserved amino acids of the SBDs located in the C-terminal region of α-amylases. The N-terminal SBD from *R. oryzae* and the *Lipomyces kononenkoae* α-amylase LKA1's N-terminal region involved in raw starch adsorption and digestion were also BLAST with the C-terminal SBD sequences. There is, however, no homology between the N-terminal and C-terminal SBDs. The occurrence of starch binding

domains in glucosyl hydrolase family members is shown in Figure 2.7. The SBD requirements of enzymes vary. In the case of *Aspergillus niger* glucoamylase, removal of the SBD results in lowered activity on raw starch, although the activity on soluble starch is retained. However, in the case of some CGTases, the truncation of the SBD or of domain E results in a loss of catalytic function (Penninga *et al*., 1996). However, the initial role of the SBD is to bind the starch substrate. In the *A. niger* SBD, which has been extensively studied, it has been demonstrated that two tryptophan residues, W^{590} in site 1 and W^{563} in site 2, are essential for binding (Williamson *et al*., 1997). Site 1 was proposed to be involved in the initial recognition of the substrate, while site 2 facilitates tighter binding, preparing the substrate for catalysis. The function and structure of this SBD suggest that the role of the SBD is merely to bind to raw granular starch, thereby increasing the local concentration of the substrate at the active site of the enzyme and dramatically increasing the rate of catalytic reaction (Cornett *et al*., 2003). Although the various studies clearly indicate the function of the SBD in raw starch binding and adsorption, the cooperation of the SBD and the catalytic domain is still not clearly understood.

5.4 Sequence similarity

Sequence similarity among α-amylases is considerably low, even if only enzymes from EC 3.2.1.1 are considered. But the availability of various sequences led to the present status of only three amino acid residues that are invariant throughout the family. The three catalytic domain residues, Asp^{206}, Glu^{230} and Asp^{297} in TAKA amylase, are well conserved throughout the family. The two invariant histidine residues, H^{122} and H^{296}, in addition to the above-mentioned catalytic residues, form the basis of the four major conserved regions in this family. It is clear, however, that the two histidines are not conserved throughout the family (Janecek, 2002). These histidines are fully conserved in the α-amylases, cyclodextrin glucanotransferases, oligo-1,6-glucosidase, maltotetraohydolase, isoamylases, neopullulanases and maltogenic α-amylases. These four conserved sequences are essentially positioned at the C terminus of the strands α_3, α_4, α_5, α_7 of the catalytic TIM barrel. Although the presence of four conserved sequences was established by Nakajima *et al.* in 1986, it was later proposed by MacGregor and Svennson (1989) that there could be additional sequence similarities within these enzymes. Three additional sequence conservations were proposed by Janecek (1995) on the basis of exhaustive analyses of a large number of amino acid sequences of α-amylases and related enzymes. Thus, the fifth conserved region is located near the C terminus of domain B, around the calcium-binding site Asp^{175} of TAKA amylase. The sixth and seventh conservations are found at the strands β_2 and β_8 of the catalytic barrel. The seven conserved regions that have been reported are shown in Table 2.3 (Janecek, 2002).

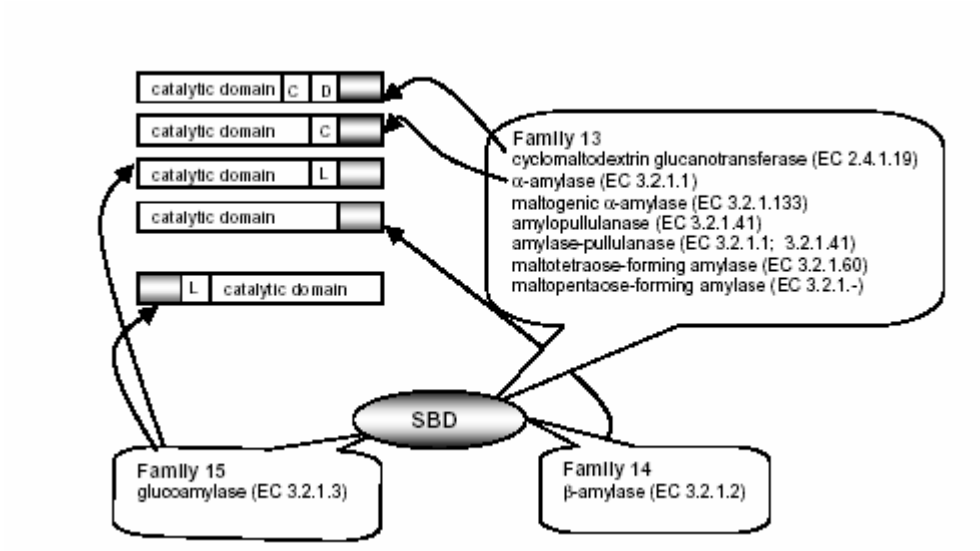


Fig. 2.7 Occurrence of starch binding domains within starch hydrolases (Juge *et al*., 2002)

5.5 Common catalytic mechanism

As many as 21 different reaction and product specificities are known for the family of α-amylases and, currently, 25 three-dimensional structures of a few members of the family have been determined by X-ray crystallography. Three amino acids, aspartic acid (D^{206}), glutamic acid (E^{230}) and aspartic acid (D^{297}), as already mentioned, have been identified as the catalytic sites in the α-amylase family members based on X-ray crystallographic analyses, chemical modification and site-specific mutagenesis (Matsuura *et al*., 1980, 1984; Vihinen *et al*., 1990; Swift *et al*., 1991; Nakamura *et al*., 1992; Hasegawa *et al*., 1999).

Table 2.3 The seven conserved regions of the catalytic domain in α-amylases (Janecek, 2002)

TAKA amylase	DVVANH……GLRIDT…….		GEVLD……..VENHD	
	REGION I	**REGION II**	**REGION III**	**REGION IV**
	β_3	**β_4**	**β_5**	**β_7**
TAKA amylase	-LPDLD	GFTAIWIT-P	GIPIIYAGQ	
	REGION V	**REGION VI**	**REGION VII**	
	C terminus of domain B	**β_2**	**β_8**	

The catalytic mechanism of α-amylases is illustrated in Figure 2.8. The reactions catalysed by α-amylases proceed with the retention of the substrate's anomeric (α-) configuration. Since each substitution at the chiral centre results in an inversion of the configuration,

catalysis must proceed through a double displacement reaction. During the first displacement, an acid group on the enzyme protonates the glycosidic oxygen, bringing about scission of the Cl-O bond and formation of an oxocarbenium ion transition state. Further, a nucleophilic acid group of the protein attacks at the sugar's anomeric centre to give a β-glycosyl enzyme intermediate, while the aglycone of the substrate leaves the active site. During the double displacement reaction in hydrolysis, the process just described is essentially reversed by the carboxylate form of the former proton donor. The second stage of the reaction proceeds via an ion-like transition state, as before, to yield a product with an α-anomeric configuration and reprotonation of the original acid group. Transglycosylation occurs if the attacking group in the second displacement reaction is a free hydroxyl of a sugar residue rather than water. In TAKA amylase A, the glutamic acid (E^{230}) is thus believed to be the proton donor, while the first of the two aspartic acid residues (D^{206} and D^{297}) is thought to act as a nucleophile. The role of the second aspartic acid is less certain, however, and has been suggested to be involved in the stabilisation of the oxocarbenium ion-like transition state and also in the maintenance of glutamic acid in the correct state of protonation for activity.

In the conserved structure of α-amylases, the arginine R^{204} imino group is always hydrogen-bonded to the side chain of the essential carboxylate aspartic acid D^{206}. Since the R^{204} imino group carries a cation, the side chain of D^{206} must be anionised. This event causes resistance to the deprotonation of E^{230}, which lies close to D^{206}. The side chain of E^{230} is also considered to be in a hydrophobic environment, as a hydrophobic residue is normally found adjacent to this residue and overhangs the side chain of E^{230} in the three-dimensional structure. The hydrophobic environment thus enhances the interaction between the D^{206} and E^{230} side chains. This leads to an increase in the pKa of E^{230}, probably to near neutral pH. The side chain of D^{297} is in the hydrophilic environment and thus is more likely to be ionised. As a result of these events, the side chain of non-ionised carboxylate E^{230} is ready to work as an acid catalyst, giving a proton to the glycosidic bond of the substrate, leading to cleavage (Matsuura, 2002).

The active site of the enzymes from the α-amylase family is comprised of a number of subsites, with each subsite being capable of interaction with the aglycone of the substrate. The subsites themselves are composed of the side chains of amino acids situated in the loops connecting the β-strands and α-helices of the catalytic domain. Since the architecture of the β-α loops vary from enzyme to enzyme, the number and nature of the subsites at the active sites vary and thus are a characteristic of a particular enzyme. Enzymes of the α-amylase family act on different substrates, but the binding of the substrates usually occurs at the -1 subsite. There is thus a strong conservation of amino acids constituting this subsite.

The large diversity of specificity and different types of reaction catalysed by enzymes in glycoside hydrolase families 13, 70 and 77 or clan GH-H invites the rational engineering of enzyme specificity. Early mutational analyses of these enzymes investigated the structure-function relationships and protein engineering addressed important industrial goals, such as the improvement of thermostability or changing pH activity dependence. Similarly, the modification of product specificity has major industrial advantages. The following sections highlight research done towards engineering these proteins and understanding the structure-function relationships in α-amylases.

6. PROTEIN ENGINEERING: SUBSTRATE SPECIFICITY OF α-AMYLASES

The enzymes of the α-amylase family have varied specificities, although they share a common catalytic mechanism; some are active only on the α-1,4 linkages of the glucose polymer, and some only on the α-1,6 linkages, while others have specificity towards both α-1,4 and α-1,6 bonds. The active site cleft of α-amylases is located at the interface between domain A and domain B and is found at the C-terminal end of the β-strands in the TIM barrel. Davies *et al*. (1997) proposed a subsite nomenclature for glycosidases. The subsites are numbered according to the location of the scissile/cleavage bond, with negative subsite numbers on the non-reducing end of the scissile bond (Figure 2.9)

In α-amylases, there are two or three subsites present at the reducing end of the scissile bond (subsites +1, +2 and +3), whereas the number of subsites at the non-reducing end varies from two to seven. For all the enzymes, activity involves the binding of a glucose residue at subsite –1(the site in the substrate where the bond cleavage occurs), while the nature of substrate binding at subsites +1 and +2 varies with the specificity of the enzyme, particularly with respect to the type of bond cleaved (MacGregor *et al*., 2001). From structural and mutagenic studies done on modified enzymes, substrates and inhibitors, it appears that the conserved segments II and III constitute subsites +1 and +2 (Matsuura *et al*., 1984). The amino acids 209, 210, 231, 232, 233 and 234 (TAA numbering), which occur in loops 4 and 5 of the $(\beta/\alpha)_8$ barrel, may confer specificity characteristics to individual enzymes. Table 2.4 shows different amino acids and mutagenic studies done to understand the functional relevance for these amino acids in full-length peptides.

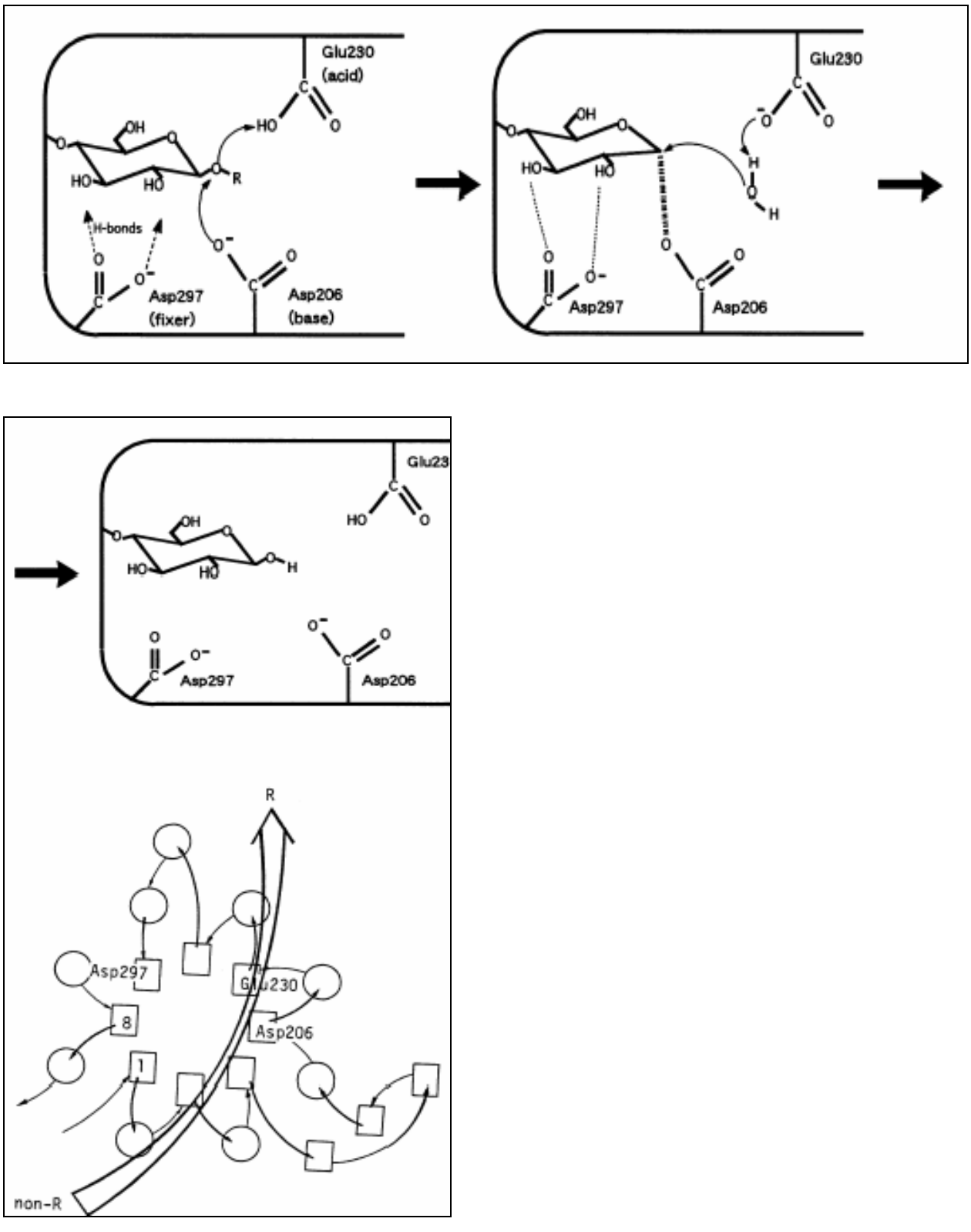


Fig. 2.8 Possible mechanism of catalysis involving the three essential residues of α-amylases (TAA numbering); direction of binding substrate (arrow) with respect to $(\beta/\alpha)_8$ catalytic barrel.

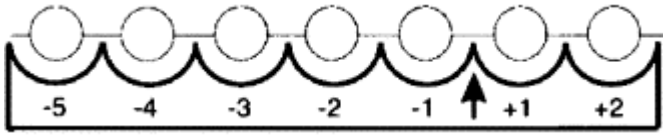


Fig. 2.9 Active site cleft subsite nomenclature for glycosyl hydrolases as defined by Davies *et al*. (1997). The non-reducing end is towards the left. The arrow indicates the scissile bond.

Table 2.4 Site-specific mutagenesis of key amino acids in α-amylases and their functional significance

Mutations	Numbering system	Significance	Enzyme	Reference
D^{206} N/Q	TAA	Catalytic nucleophile	*Aspergillus* α-amylase	Nagashima *et al*., 1992
Q^{230} E/D	TAA	Proton donor		Uitdehaag *et al*., 2000
D^{297} N/Q	TAA	Catalysis		Nielsen and Borchert, 2000
K ^{209} F/R	TAA	Catalysis		Nagashima *et al*.,1992
H ^{296} E/N	TAA			
H^{122}	TAA	Transition-state stability		Svensson,1994
K^{209}	TAA			MacGregor *et al*., 2001
H^{210}	TAA			
V^{231}	TAA			
L^{232}	TAA			Neilsen and Borchert, 2000
D^{233}	TAA	Form the subsites +1 and +2, contribute to specificity		
G^{234}	TAA			
D^{176} N	D^{206} TAA	Reduced activity on starch	*B. stearothermophilus* α- amylase	Takase, 1992
H ^{180} N	H^{210} TAA	Reduced 17-fold activity		
E ^{203} Q	E^{230} TAA			
E ^{208} Q		Reduction in activity 1-fold		
D ^{269} N	D^{297} TAA	Results in poor activity		
R ^{232} K	R^{204} TAA			
H ^{238} D	H ^{210}	Catalytically important residues		
H ^{238} N				
H ^{238} Y				
E ^{264} V	E ^{230}			
Y ^{265} F	V ^{231}			

Y ^{265} S D ^{331} A D ^{331} E	D ^{297}			
Y ^{83} F Y ^{83} L Y ^{83} N Y ^{83} W W ^{84} F W ^{84} L W ^{84} Y	A^{102} L^{103}	Promotes transglycosylation	*Saccharomycopsis* amylase	Svensson,1994
H ^{247} E D ^{328} H D ^{328} N E ^{332} H E ^{357} H	H ^{122} TAA D ^{206} H ^{210} E ^{230}	Reduced hydrolysis of α-1,4 and 1,6 bonds	*Bacillus stearothermophilus* neopullulanase	Kuriki *et al*., 1991
Y ^{265} F Y ^{265} S	V ^{231} TAA	Moderately active Fully active	*Bacillus stearothermophilus* α-amylase	Holm *et al*., 1990
W ^{254} V	V ^{231}	Discriminate cyclisation over the hydrolytic reactions		
F ^{255} I	L ^{232}	Lost capacity to cyclise but there was a 2.5-fold increase in saccharifying capacity	*Bacillus stearothermophilus* CGT ase	
K ^{209} N	Y	Reduced affinity		
K ^{210} R		Lower activity on shorter substrates than longer substrates, increased affinity improving binding to 1st and 2nd glucose of aglycone part of the substrate	*Saccharomycopsis fibuliger* α - amylase	Matsui *et al*., 1992
H ^{233} N	H^{210} TAA	Suppressed production of α-cyclodextrins	CGTases	
H ^{144} N H ^{327} N	H^{122} TAA H ^{296} TAA	Enhance the formation of β-cyclodextrins		Nakamura *et al*., 1993
H ^{210} N		Enhanced maltase activity	Human pancreatic α-amylase	
E ^{332} H	H ^{210}	Enhanced 1,6 specificity	Neopullulanase	

		and reduced 1,4 cleavage		
N ^{329} K N ^{329} V N ^{329} D	N ^{295} TAA	To mimic pullulanase To mimic isoamylase	*Bacillus stearothermophilus* α-amylase	Takase *et al*., 1993
N ^{329} V N ^{329} D		Twice the activity of wild type at 15°C		Svensson,1994
H ^{607} A H ^{607} N	H ^{122}	Catalytic site and conserved region 1, stability and catalysis	*Klebsiella aerogenes* pullulanase	Yamashita *et al*., 1997
D ^{677} S				
H ^{633} A H ^{633} N	H ^{298}	Catalysis		
Y ^{559} A Y ^{559} F Y ^{564} A Y ^{564} F	Y^{75} and Y^{79} TAA	Important for 1,6 activity substrate-binding capacity was retained	*Klebsiella aerogenes* Pullulanase	Yamashita *et al*., 1997

To desirably alter substrate specificity, the mutations are proposed specifically in the β/α loops of the catalytic barrel. The three-dimensional structures known thus far also facilitate the understanding of the structure and the amino acids involved in binding to the substrate. In addition to the four conserved sequences noticeable in all α-amylases, the domain B protruding through the catalytic barrel might be involved in substrate interactions or conformations facilitating the active site to accommodate a broad range of substrates. Thus, the β-α connecting segments of the catalytic barrel create the substrate-binding sites and the catalytic site. A vast majority of GH-H enzymes have a long β-α 3 segment (referred to as domain B) containing secondary structure elements (Janecek *et al*., 1997). It starts with a consensus sequence motif containing essential residues, which, together with β-strands 4, 5 and 7 and their immediate extensions, reflect enzyme specificity (Janecek *et al*., 1997; MacGregor *et al*., 2001). These four segments are important in certain substrate-binding subsites. The length of the subsites varies in individual enzymes. α-Amylases are generally considered as endo-acting enzymes and it is not clearly understood which amino acid residues of exo-acting enzymes ensure exo-activity over endo-activity, although they are likely to be in the loops 1, 2, 3, 7 and 8. These loops contribute to the binding glucose

residues of the substrate on the glycone side of the bond on which the enzyme acts (Strokopytov *et al*., 1996; Yoshioka *et al*., 1997; Fujimoto *et al*., 1998; Kadziola *et al*., 1998). It is evident from the mutagenic studies, however, that the presence of lysine or arginine at position 209 is essential for α-1,4 activity, since these residues are conspicuously absent in enzymes with specificity towards other glycosidic bonds. In addition, histidine or glycine at position 210 has some relevance to the α-1,4 activity of the enzymes. In the case of enzymes acting on α-1,6 linkages, a Ser or Gly is most commonly found at position 209. However, α-1,6 activity has been observed in the presence of glycine in some α-glucosidases. In addition, a hydrophobic amino acid, such as valine, leucine or methionine, is found at position 207. Specificity for α-1,6 linkages also demands the presence of a hydrophobic or aromatic amino acid at position 210.

Some α-amylases, neopullulanases and amylopullulanases act on α-1,4 and α-1,6 linkages at a single active site. These enzymes mainly possess asparagine at position 209 and glutamine at position 210, although many other possibilities are known.

6.1 Mapping subsite amino acids in α-amylases

X-ray crystallographic analysis of proteins in the crystalline state, free or complexed with the substrate analogue, is a powerful method for mapping the active site of alpha amylases. However, these data can vary according to the crystalline varieties and, in many cases, are not available at all. The use of low molecular mass substrates could therefore be an effective way to elucidate the number of subsites of the active site in these enzymes. In addition, mutagenic and kinetic analyses summarise data regarding the requirements of certain amino acids and structure-function implications in these molecules. The process of quantifying a subsite model is referred to as subsite mapping. The methods used for subsite mapping originate from the early 1970s. Quantitative theories of the action pattern of amylase in terms of subsite affinities were proposed independently by Hiromi (1970) and Allen and Thoma (1976), and later by Suganama *et al*. (1978). Hiromi *et al*. (1970) proposed a method based on kinetics for evaluating subsite affinities from the dependence of the hydrolytic rate on the degree of polymerisation (DP) of the substrates. A product-based method was developed by Allen and Thoma, which is feasible for endo-acting enzymes but not for exo-enzymes. Suganama *et al*. (1978) devised a method based on the calculation of bond cleavage frequency data at a sufficiently low substrate concentration whereby the secondary attacks on the substrate can be ignored. The active site cleft in α-amylases encompasses a varying number of consecutive subsites interacting with the substrate glycosyl residues. Enzymatic subsite mapping procedures thus assist in characterising the number of glycosyl residues recognised by the enzyme, the binding affinity of individual subsites and the position of the bond to be cleaved. The spatial distribution of binding forces illustrates how particular

subsites of high or low affinity along the cleft control the productive binding modes of the substrates.

Electron density studies of the TAKA amylase molecule complexed with the inhibitor acarbose, which is a pseudotetrasaccharide unit, and kinetic characterisation studies suggested the presence of at least six ordered saccharide units (Figure. 2.9) (Brzozowski and Davies, 1997; Gilles *et al*., 1996; Qian *et al*., 1994; Strokopytov *et al*., 1996; Nitta *et al*., 1971). Clearly connected density for six saccharide units was found in the active site of TAKA amylase.

The active centre of TAKA amylase is located in a deep pocket on the C-terminal side of the (β/α)8 barrel domain. Crystallographic data also suggest that the active site is a V-shaped depression formed by the interface between the A and B domains. The sugars at the non-reducing end of the chain enter via a steep couloir until they reach the -1/+1 subsites at the catalytic centre, where the sugar gets cleaved. There are four highly conserved sequences and the effective role of amino acids found in these regions in α-amylases has been studied through mutagenic and structural evaluations. Despite the conservation, there might be some amino acids variant in these regions that confer certain substrate specificity properties to different enzymes of this family. The amino acids involved in catalysis and subsite formation in TAKA amylase, the structural interactions at the subsites and their specific roles are summarised in Table 2.5.

The isoenzymes of barley alpha amylase AMY1 and AMY2 have 10 consecutive substrate-binding subsites, -1 to -6, towards the non-reducing end of the substrate from the catalytic site and +1 to +4 towards the reducing end (Mori *et al*., 2002). The structure of *Bacillus amyloliquefaciens* and *Bacillus licheniformis* α-amylases with maltodecaose analogue derived from acarbose accommodated sites -7 to +3 (Machius *et al*., 1995; Brzozowski *et al*., 2000).

Table 2.5 Subsite-forming amino acids of TAKA amylase and their structural significance

Subsite	TAKA amylase	Structural significance
-3	Tyr^{75}	The binding of non-reducing sugar at subsite -3 is the aromatic stacking with Tyr^{75}
-3	His^{80}, Gln^{35}	The sugar makes a direct H bond to the protein from O3 to the amide nitrogen of Gln^{35}. A water- mediated H bond then bridges His^{80} to the O2 of the sugar hydroxyl

-2	Asp[340], Arg[344], Trp[83]	Hydrogen bonding of the O3 atom with the carboxylate of Asp[340]; O2 with the guanidine head group of Arg[344]; O6 with the NEI of Trp[83]
+2	Lys[209], Leu[232]	Direct interactions exist only with the +2 sugar O2 and O3 of terminal amino of Lys[209] The pyranoside ring in the +2 stacks above the hydrophobic moiety of Leu[232] (Svennson, 1994)
+3	Gly[234]	Amide of Gly[234] bond with+3 O3 of the sugar
-1	Asp[297], Glu[230], His[122], Asp[206]	O3 and O2 hydroxyls hydrogen bond with the Asp[297], conformation of subsite -1 mimics the transition state , the hydrogen bonding with Asp[297] facilitates water-mediated h bond with Glu[230]; O6 group interacts with His[122] and Asp[206] carboxylate; -1 subsite also makes an important stacking with Trp[82]: orientation of this Tyr is maintained by hydrogen bond between the Tyr OH and NE2 of His[296]
+1	His[210], Glu[230]	O of Glu[230] hydrogen bonds with the+1; His [210] hydrogen bonds with O2 hydroxyl of the sugar, all other interactions at +1 are through solvent molecule

In porcine pancreatic α-amylase, enzymatic studies revealed the presence of an array of five consecutive binding subsites (-3, -2, -2, 1, 2), each of which interacts with a single substrate glucosyl residue (Kandra *et al*., 1997; Brayer *et al*., 2000). The catalytic site is located between the second (1) and the third (-1) subsite from the reducing end. The other widely studied enzyme is the human salivary α-amylase (HAS). The action pattern of HSA on model substrates revealed that the binding region in HSA consists of at least six subsites; four glycone-binding sites (-4, -3, -2, -1) and two aglycone-binding sites (+2, +1) (Kandra and Gyemant, 2000). Detailed information of the subsite model of enzymes would be useful to alter the substrate specificity of an enzyme by mutagenic and recombinant techniques.

6.2 Subsite mutagenesis for substrate specificity engineering

Subsite-forming amino acids interact directly with the substrate and hence have a significant influence on the substrate preference properties of enzymes. Barley α-amylase is one of the few members on which mutational, kinetic and structural analyses of subsites have been done thus far and some mutant enzymes with improved properties have been generated successfully. Site-directed mutagenesis studies in barley α-amylase AMY1 revealed that the mutation of the -5 /-6 subsites (Cys[95]-Ala) significantly elevated the activity to 150-fold higher

than that of the wild type towards insoluble starch; the double mutant Cys^{95}-Ala/Met^{298}-Ala retained the above enhanced enzymatic activity. The corresponding single mutant, Met^{298}-Ala, exhibited only 10-fold higher relative activity, although its catalytic activity was the highest on amylose DP_{17} (Degree of polymerisation 17). Thus, the double subsite mutant favourably selects for large and insoluble substrates over a short substrate. It has been suggested that, to stabilise the productive enzyme-substrate complexes, there was wider contact for the mutants with the substrate aglycone part, reflecting that certain interactions were either lost or favoured as a consequence of these mutations. The mutants thus overcome the impact of perturbation near the catalytic site through several distant substrate-protein interactions, as illustrated by the higher activity on larger oligosaccharide substrates.

In a study by Mori *et al*. (2002), the central role of Met^{53} from a substrate glycon-binding motif at the high-affinity subsite -2 was emphasised by the mutational and kinetic analysis of various substrates. The action pattern of the Met^{53} mutants of loop 2 of the TIM barrel showed that modification of subsite -2 could influence the utilisation of outermost subsite, -6, and was necessary for wild-type kinetic properties such as affinity and action on maltooligosaccharides and maltodextrins. Furthermore, local random mutagenesis was performed to study the functional significance of Arg^{183}-Gly^{184}-Tyr^{185} localised in the fourth β-α loop in the catalytic domain of AMY1. In these amino acid positions, other family members with different specificity possess Arg/Lys-Gly-aromatic in plant α-amylases and Lys-His-Z (hydrophobic/aromatic residue) in animal or microbial α-amylases. KGY-AMY1 mimicking plant α-amylases resulted in a three-fold increase in activity on G_6PNP and G_7PNP. The activity of SGM-AMY1 increased with increasing substrate length. The strong dependence of this mutant activity on substrate length was explained to be the result of increased K_{int}, the interaction constant, and decreased affinity at one or more subsites near the catalytic site. These mutations clearly illustrated that the long chain of Arg^{183} contributes to binding at subsite +1 and +2, but also confers constraints on substrate processing. It was proposed that a combination of mutations in the subsites may accentuate the suppression of activity on shorter oligosaccharides and further develop enzyme specificity. The various mutations performed in barley α-amylase to evaluate the importance of subsite amino acids and susbtrate preference properties are listed in Table 2.6. The studies on barley α-amylase thus offer promising prospects for activity improvement and action pattern modulation, and similar strategies can be applied by detailed analyses of other members. It is often possible to compare enzymes with different specificities and engineer an enzyme with desired specificity.

Table 2.6 Mutations of subsite-forming amino acids for substrate preference engineering

Amino acids and protein source	Mutations	Location	Significance
R ^{183}-G ^{184}- Y^{185}	S ^{183}-G ^{184}- M^{185}	Fourth α-β loop; subsite +2	Decreased affinity for calcium ions; 2-fold increased activity on amylose
Barley α-amylase	N ^{183}-G ^{184}- Y^{185}		Slightly lowered activity on G_7PNP, but reduced activity on smaller oligosaccharides
	K ^{183}-G ^{184}- Y^{185}		3-fold elevated activity on oligosaccharides
	T ^{183}-G ^{184}- L^{185}		Decreased calcium ion activity and modest activity on all substrates
C ^{95}; M^{298}	C^{95}-A	Domain B -5/-6 subsite	Elevated activity on insoluble starch
Barley α-amylase	C^{95}- A ; M^{298}- A		Elevated activity of C-A maintained;10-fold higher activity and highest catalytic efficiency on amylose
M^{53}	M^{53}-E; M^{53}-A; M^{53}G/D;M^{53}-Y; M^{53}-W	High affinity subsite -2; C-terminal of α-β_7	G/D displayed 117/ 90%; A/S/G resulted in 76/58/38%; Y/W 0.9/0.1% towards insoluble substrate compared to wild type
Y^{102} and T^{212}	Y-A		140,15 and <1% activity on starch, amylose and maltoheptaoside
Barley α-amylase	T-A		32, 370 and 90% activity on starch, amylose and maltoheptaoside
	Y^{105} A and T^{212}Y/ W		Low activity on insoluble starch but selective improvement of binding on soluble substrates; Y^{105} crucial for binding at subsite -6
D^{197}	N	Catalytic cleft	10^6-fold decrease in catalytic activity
E^{233}	Q		10^3-fold decrease in catalytic activity
D^{300}	N	Forms part of +2 subsite	Shift the point of cleavage closer to the reducing end by one subsite, reduced affinity
Y^{151}	M		

Although the reliable design of properties is not feasible from the present limited number of mutants and information available, such studies have been applied in the engineering of corresponding subsites in related glucosyl hydrolase members. Protein engineering of amylolytic enzymes by structure-guided evolution and rational design can exploit the knowledge gained in subsite engineering to meet specificity requirements and to improve insight into enzyme-substrate relationships.

7. STABILITY ENGINEERING

α-Amylases are very useful in industry and most of the industrial processes involve the use of these enzymes under extreme conditions with respect to temperature and pH. There are many reasons for the lack of performance by enzymes on an industrial platform, one of them being incubation at high temperatures, which primarily leads to unfolding and thus a loss of activity. On the other hand, exposure to extreme pH conditions, metal ion chelators and oxidizing agents can also result in the denaturation of enzymes. Thus, the overall stability of these enzymes is of primary interest under industrial conditions.

7.1 Thermostability of α-amylases

α-Amylases from *Bacillus* are industrially important due to their thermostability and have been studied in detail using biochemical and protein engineering methods. *Bacillus licheniformis* α-amylase (BLA), *Bacillus stearothermophilus* (BStA) and the α-amylase from *B. amyloliquefaciens* (BAA) have been investigated and serve to identify some plausible mechanisms of thermostability. Of these three enzymes, BLA is the most stable, followed by BStA and BAA. A comparison of the effects of temperature on these enzymes was done and it was proposed that the primary cause of inactivation of BLA was the deamidation of asparagine and glutamine residues in the catalytic site of the enzyme. However, Tomazic and Klibanov (1988) proposed that the enhanced stability of BLA is primarily due to the additional salt bridges involving the lysine residues, K^{88}, K^{253} and K^{385}. On the basis of this, a number of researchers generated hybrid molecules between α-amylases as an elegant way of pinpointing the regions of particular importance for different properties of individual enzymes. Two regions were identified as being important for the thermostability of BLA (Suzuki *et al*., 1989) and, three stabilising mutations were proposed on the basis of amino acid differences: the deletion of R^{176} and G^{177} and the substitution of $K^{269}A$. These mutations led to enhanced and additive thermostability in BAA; similar effects to these mutations were observed in other α-amylases (Bisgaard-Frantzen *et al.*, 1999; Igarashi *et al*., 1998; Matthews *et al*., 1987). Furthermore, Conrad *et al*. (1995) pointed out the significance of four

regions of particular importance in the thermostability of BLA: 34-76, 112-142, 174-179 and 263-276. However, these findings did not support the hypothesis of Tomazic and Klibanov (1988) on the stabilising effect of lysine residues. On the contrary, the removal of lysine led to a stabilising effect in the enzymes.

One of the most common protein engineering concepts for thermostability is the introduction of prolines. It is supposed that the stabilising effect is contributed by lowering the entropy of the unfolded state of an enzyme more than the entropy of its folded state. Such mutations in the alkalophilic α-amylase of *Bacillus* sp. have been successfully applied to render thermostability. Two positions of BLA, H^{133} and A^{209}, were identified to play a role in stability at high temperature and acidic pH (Joyet *et al*., 1992; Declerck *et al*., 1990). The replacement of H^{133} with isoleucine had the largest stabilising effect – even better than that provided by Y^{133}. The second mutation, $A^{209}V$, was a result of random mutagenesis and the stabilising effect of this mutation was additive in the presence of the Y^{133} mutation. The additive nature of stabilising mutations suggests that they affect the same reversible unfolding step. It was explained that the stabilising effect of $A^{209}V$ was a result of a hydrophobic crevice being filled at the surface of the protein (Declerck *et al*., 1997). Since domain B constitutes the aforementioned major stabilising mutations identified thus far, it was speculated that the initial reversible step is a partial unfolding of the B domain. As thermostability is not a feature that could be attributed by one amino acid change, repeated random mutagenesis and screening have been used as an elegant tool to isolate isoenzymes with greater stability.

Apart from the direct evidence from mutagenesis studies, there are also many hypotheses on the thermostability of enzymes, includes a substantial number of very small stability-enhancing changes throughout the protein. Other suggestions include enhanced internal hydrophobicity, amino acid substitutions that favour the stability of the α-helix regions of the molecules, increased sheet-forming abilities, the presence of aliphatic side chains of amino acids buried within the structures of proteins, leading to increased bulkiness, and an increasing number of prolines, especially the prolines at the second sites of the β-turns and in the first turn of α-helices. In most α-amylases, the raw starch-binding domain may govern thermostability. This might be due to the fact that the β-sheet compact structure of the starch-binding domain provides the molecule with more conformational rigidity. It was also reported that a correlation exists between the thermostability and hydrophobicity of introduced residues in hydrophobic pockets (Shih *et al*., 1995), which might explain why a highly hydrophobic SBD contributes to thermostable behaviour. Accordingly, some enzymes exhibit thermostability properties owing to the presence of starch-binding domains. The *Cryptococcus* α-amylase starch-binding domain located at the C terminus of the molecule confers thermostability properties (Iefuji *et al*., 1996).

It is a known fact that the thermostability of an enzyme depends on the flexibility of each molecule that comprises it (Vihinen, 1990) and on improved conformational rigidity, which is indispensable for enhancing the stability of an enzyme against heat denaturation (Zavodsky *et al*., 1998). However, due to the abundance of structural features with which thermostability can be correlated, these attempts are bound to remain empirical and probably unsuccessful, unless the mechanisms of thermo-inactivation are understood beforehand, thus greatly reducing the number of possible hypotheses. On the other hand, domain engineering, which involves the fusion of the starch-binding motif to an enzyme that lacks it, has been of use. This approach allows us to create chimeras, which, in addition to conferring wild-type properties, can also adsorb to raw starch and, owing to the starch-binding domain, exhibits higher thermostability than the original molecule (Ohdan *et al*., 2000).

In recent years, the number of reports on cold-adapted enzymes has been of special interest since the investigations on their catalytic efficiency and capability to work at low temperatures will help the scientific community to understand structure-function relationships based on temperature. Thus, thermostability is not the only criterion that can be useful to industry. Although an enzyme is believed to be thermally adapted to a temperature at which it has the highest activity, this does not mean that the enzyme is adapted to have the highest turnover rate possible at this temperature (Sheridan *et al*., 2000). Moreover, the optimum temperature of activity of an enzyme is not usually the same as the growth temperature of the organism that is the source of the enzyme. Process-based protein engineering often involves the design of a protein/enzyme that can work optimally at the temperature at which the fermenting organism thrives. It is therefore important that, in addition to studies on enzymes active at higher temperatures, ground research is done on enzymes from psychrophiles that are active at lower temperatures. It is also important to differentiate the structural changes that affect stability (how an enzyme holds together) from those that affect catalysis. Although many cold-active enzymes are thermolabile, it is not essential that structural changes producing cold activity require thermolability (Miyazaki *et al*., 2000).

Alteromonas haloplanctis α-amylase (AHA), for which a three-dimensional structure has been elucidated, is the best studied psychrophilic α-amylase (Aghajari *et al*., 1998) known. The AHA structure has three domains characteristic of mesophilic α-amylases, with the central domain A forming a $(\beta/\alpha)_8$ barrel, a domain B protruding between the β_3 strand and the α_3 helix and a C-terminal globular domain consisting of a Greek key motif. However, the superimposition of AHA on its mesophilic counterpart shows that 24 residues forming the five subsites are strictly conserved and no changes were found in the catalytic centre, indicating that the mechanism of cold adaptation lies elsewhere in the molecule. An additional disulfide bridge, which was absent in AHA, was observed between domain A and domain B in mesophilic enzymes. The cold-active amylase has a lower affinity for calcium binding at the

site situated between domains A and B which in turn increases the flexibility of the enzymes. The loss of this domain tethering may provide the enzyme with a greater degree of flexibility and freedom, which may be the key to its adaptation to low temperatures. No striking differences between the cold-active enzyme and its mesophilic counterparts were observed elsewhere. It is relevant in this context that organisms growing at low temperatures do not encounter high temperatures, which will eliminate the selective pressure to maintain thermostability. Thus, the two characteristics are physically unlinked, but are often found together because the selection of one property (cold activity) is not often co-selected with the need for activity at higher temperatures (thermostability).

7.2 Engineering pH activity profiles

Stability at a wide range of pH values is an important industrial property and the successful use of amylases in industrial processes is dependent on having an enzyme that is both stable and active at the pH of the process. At extreme pH conditions, destabilising phenomena occur, such as the spontaneous formation of peptide succinamides (Asp-Gly and Asn-Gly sequences), the burial of ionised groups and the repulsive electrostatic forces associated with the large net charges of the protein found at both acidic and basic pH, and these contribute to decreased stability in the enzymes. Protein stability engineering techniques, such as helix capping, removal of deamidating residues and cavity filling (Joyet *et al*., 1992; Declerck *et al*., 1995, 1997; Bisgaard-Frantzen *et al*., 1999; Wind et al., 1998), can be applied for improving the stability of enzymes.

Although stability is essential for activity at extreme pH, it is not sufficient; the active site residues must be in a protonated state in order for the enzyme to be active. Thus, the proton donor in α-amylases (Glu) must be protonated, while the nucleophile must be negatively charged. α-Amylases exhibit bell-shaped activity profiles (Qian *et al*., 1994; Strokopytov *et al.,* 1995), which suggest that the acidic limb of the activity profile is determined by the titration of the catalytic nucleophile and the basic limb reflects the titration of the catalytic proton donor. Therefore, if an α-amylase is stable over the entire pH range, it is conceivable that the pH activity profile can be changed if the Pka value of either the nucleophile or the proton donor is changed. Pka values can be changed by altering the electrostatic field experienced by the titrable group (Linderstrøm-Lang and Nielsen, 1959). Both the solvent accessibility of the titrable group, as well as of other charged groups, and the local hydrogen-bonding network can influence the electrostatic field. To exploit this hypothesis, charged residues are typically inserted near the titrable group. Mutagenic experiments that remove and insert charged residues near the active site (Nielsen *et al.,* 1999; Sogaard *et al*., 1993) have been attempted in α-amylases and CGTases, but no defined or predictable way of changing the pH activity profile has been found. Mutations done in BLA and BstA to perturb

the electrostatic field in the active site by inserting or deleting charged amino acids or changing the solvent accessibility of the catalytic nucleophile resulted in shifts opposite to the expected theoretical results (Jackson and Fersht, 1993; Nielsen *et al*., 1999; Fang and Ford, 1998; Bakir *et al*., 1993). Thus mutations in BLA suggest that there could be factors other than the electrostatic repulsions that contribute to the pH activity profile of α-amylases. It is difficult to rationalise the effect produced by several of the stabilising mutations, and the processes leading to the irreversible denaturation of α-amylases are thus still unknown.

Protein engineering techniques have already been used successfully to answer some seminal questions about enzyme mechanisms. They have also yielded several engineered enzymes of great practical utility and commercial value. Nevertheless, the full potential for generating novel (or the improvement of existent) enzymes is still largely untapped. Recent advances in using *in vitro* evolution techniques and the opening of metagenomics approaches offer a larger sequence space for manipulation.

Effective catalysis includes a combination of proficient enzymatic systems and the expression of these in a suitable host for the optimal processing of fermentation-*proficient enzymes in novel cell factories*. While protein engineering methods improve the substrate range for starch hydrolases and deliver isoenzymes with properties best suited for the starch industry, the preparation of whole-cell biocatalysts in which these enzymes can be expressed opens a new trend in emerging biotechnological applications. An engineered multi-enzyme system and microorganism and optimal process methods in conjunction would route the vision of molecular bioengineering towards an attractive, cost-effective and environmentally friendly fuel ethanol initiative.

8. REFERENCES

Abe A, Tonozuka T, Sakano Y, Kamitori S. 2004. Complex structures of *Thermoactinomyces vulgaris* R-47 alpha-amylase 1 with malto-oligosaccharides demonstrate the role of domain N acting as a starch-binding domain. *J Mol Biol* **335(3)**:811-822.

Aghajari N, Feller G, Gerday C, Haser R. 1998. Crystal structures of the psychrophilic alpha-amylase from *Alteromonas haloplanctis* in its native form and complexed with an inhibitor. *Protein Sci* **7**:564-572.

Allen JD, Thoma JA. 1976. Subsite mapping of enzymes. Application of the depolymerase computer model to two alpha-amylases. *Biochem J* **159(1)**:121-132.

Ashikari T, Nakamura N, Tanaka Y, Kiuchi N, Shibono Y, Tanaka T, Amachi T, Yoshizumi H.

1986. *Rhizopus* raw starch degrading glucoamylase: its cloning and expression in yeast. *Agric Biol Chem* **50**:957-964.

Atia KS, Ismail SA, El-Arnaout MB, Dessouki AM. 2003. Use of co-immobilized beta-amylase and pullulanase in reduction of saccharification time of starch and increase in maltose yield. *Biotechnol Prog* **19**:853-857.

Bakir U, Coutinho PM, Sullivan PA, Ford C, Reilly PJ. 1993. Cassette mutagenesis of *Aspergillus awamori* glucoamylase near its general acid residue to probe its catalytic and pH properties. *Protein Eng* **6**:939-946.

Bisgaard-Frantzen H, Svendsen A, Norman B, Pedersen S, Kjærulff S, Outtrup H, Borchert TV. 1999. Development of industrially important α-amylases. *J Appl Glycosci* **46**:199-206.

Brayer GD, Sidhu G, Maurus R, Rydberg EH, Braun C, Wang Y, Nguyen NT. Overall CM, Withers SG. 2000. Subsite mapping of the human pancreatic alpha-amylase active site through structural, kinetic, and mutagenesis techniques. *Biochemistry* **39(16)**:4778-4791.

Brzozowski AM, Davies GJ. 1997. Structure of *Aspergillus oryzae* α-amylase complexed with inhibitor acarbose at 2.0 Å resolution. *Biochemistry* **36**:10837-10845.

Brzozowski AM, Lawson DM, Turkenburg JP, Bisgaard-Frantzen H, Svendsen A, Borchert TV, Dauter Z, Wilson KS, Davies GJ. 2000. Structural analysis of a chimeric bacterial alpha-amylase. High-resolution analysis of native and ligand complexes. *Biochemistry* **39(31)**:9099-9107.

Conrad B, Hoang V, Polley A, Hofemeister J. 1995. Hybrid *Bacillus amyloliquefaciens* × *Bacillus licheniformis* α-amylases. Construction, properties and sequence determinants. *Eur J Biochem* **230**:481-490.

Cornett CA, Fang TY, Reilly PJ, Ford C. 2003. Starch-binding domain shuffling in *Aspergillus niger* glucoamylase. *Protein Eng* **16(7)**:521-529.

Coutinho PM, Henrissat B. 1999. Carbohydrate-active enzymes: an integrated database approach. In *Recent Advances in Carbohydrate Bioengineering*, Gilbert HJ, Davies G, Henrissat B, Svensson B (eds). The Royal Society of Chemistry: Cambridge; 3-12.

Davies GJ, Wilson KS, Henrissat B. 1997. Nomenclature for sugar-binding subsites in

glycosyl hydrolases. *Biochem J* **321**:557-559.

De Moraes LM, Astolfi-Filho S, Oliver SG. 1995. Development of yeast strains for the efficient utilisation of starch: evaluation of constructs that express alpha-amylase and glucoamylase separately or as bifunctional fusion proteins. *Appl Microbiol Biotechnol* **43(6)**:1067-1076.

Declerck N, Joyet P, Gaillardin C, Masson JM. 1990. Use of amber suppressors to investigate the thermo-stability of *Bacillus licheniformis* α-amylase. Amino acid replacements at 6 histidine residues reveal a critical position at His-133. *J Biol Chem* **265**:15481-15488.

Declerck N, Joyet P, Trosset JY, Garnier J, Gaillardin C. 1995. Hyperthermostable mutants of *Bacillus licheniformis* α-amylase: multiple amino acid replacements and molecular modelling. *Protein Eng* **8**:1029-1037.

Declerck N, Machius M, Chambert R, Wiegand G, Huber R, Gaillardin C. 1997. Hyperthermostable mutants of *Bacillus licheniformis* α-amylase: thermodynamic studies and structural interpretation. *Protein Eng* **10**:541-549.

Eksteen JM, Van Rensburg P, Cordero Otero RR, Pretorius IS. 2003. Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the alpha-amylase and glucoamylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*. *Biotechnol Bioeng* **84(6)**:639-646.

Fang TY, Ford C. 1998. Protein engineering of *Aspergillus awamori* glucoamylase to increase its pH optimum. *Protein Eng* **11**:383-388.

Farber GK, Petsko GA. 1990. The evolution of α/β barrel enzymes. *Trends Biochem Sci* **15**: 228-234.

Farber GK. 1993. An α/β barrel full of evolutionary trouble. *Curr Opin Struct Biol* **3:**409-412.

Fujimoto Z, Takase K, Doui N, Momma M, Matsumoto T, Mizuno H. 1998. Crystal structure of a catalytic-site mutant α-amylase from *Bacillus subtilis* complexed with maltopentaose. *J Mol Biol* **277**:393-407.

Gilles C, Astier JP, Marchis-Mouren G, Cambillan C, Payan F. 1996. Crystal structure of pig pancreatic α-amylase isoenzyme II, in complex with the carbohydrate inhibitor acarbose. *Eur*

J Biochem **238**:561-569.

Griffiths AD, Tawfik DS. 2000. Man-made enzymes – from design to in vitro compartmentalisation. *Curr Opin Biotechnol* **11(4)**:338-353.

Hahn-Hagerdal B, Jeppson H, Skoog K, Prior BA. 1994. Biochemistry and physiology of xylose fermentation by yeasts. *Enzyme Microbiol Tech* **16**:933-943.

Hasegawa K, Kubota M, Matsuura Y. 1999. Roles of catalytic residues in alpha-amylases as evidenced by the structures of the product-complexed mutants of a maltotetraose-forming amylase. *Protein Eng* **10**:819-824.

Hiromi K. 1970. Interpretation of dependency of rate parameters on the degree of polymerization of substrate in enzyme-catalyzed reactions. Evaluation of subsite affinities of exo-enzyme. *Biochem Biophys Res Commun* **40**:1-6.

Holm L, Koivula AK, Lehtovaara PM, Hemminki A, Knowles JKC. 1990. Random mutagenesis used to probe the structure and function of *Bacillus stearothermophilus* alpha-amylase. *Protein Engng* **3**:181-191.

Iefuji H, Chino M, Kato M, Limura Y. 1996. Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus* sp. S-2: purification, characterization, cloning and sequencing. *Biochem J* **318**:989-996.

Igarashi K, Hatada Y, Ikawa K, Araki H, Ozawa T, Kobayashi T, Ozaki K, Ito S. 1998. Improved thermo-stability of a *Bacillus* α-amylase by deletion of an arginine-glycine residue is caused by enhanced calcium binding. *Biochem Biophys Res Commun* **248**:372-377.

Jackson SE, Fersht AR. 1993. Contribution of long-range electrostatic interactions to the stabilization of the catalytic transition state of the serine protease subtilisin BPN′. *Biochemistry* **32**:13909-13916.

Janecek S, Svensson B and Henrissat B. 1997. Domain evolution in the α-amylase family. *J Mol Evo.* **45**:322-331.

Janecek S, Svensson B, Macgregor EA. 2003. Relation between domain evolution, specificity and taxonomy of the α-amylase family members containing a C-terminal starch-binding domain. *Eur J Biochem* **270**:635-645.

Janecek S. 1995. Close evolutionary relatedness among functionally distantly related members of the (alpha/beta)8-barrel glycosyl hydrolases suggested by the similarity of their fifth conserved sequence region. *FEBS Lett* **377(1)**:6-8.

Janecek S. 1997. α-Amylase family: molecular biology and evolution. *Prog Biophys Mol Biol* **67**:67-97.

Janecek S. 2002. How many conserved sequence regions are there in the α-amylase family? *Biologia, Bratislava* **57(11)**:29-41.

Janse BJ, Pretorius IS. 1993. Expression of the *Klebsiella pneumoniae* pullulanase-encoding gene in *Saccharomyces cerevisiae*. *Curr Genet* **24(1-2)**:32-37.

Janse BJ, Pretorius IS. 1995. One-step enzymatic hydrolysis of starch using a recombinant strain of *Saccharomyces cerevisiae* producing alpha-amylase, glucoamylase and pullulanase. *Appl Microbiol Biotechnol* **42(6)**:878-883.

Jespersen HM, MacGregor EA, Sierks MR, Svensson B. 1991. Comparison of the domain-level organization of starch hydrolases and related enzymes. *Biochem J* **280**:51-55.

Juge N, Le Gal-Coeffet M, Caroline SM, Furniss A, Gunning P, Kramhoeft B, Morris VJ, Williamson G, Svensson B. 2002. The starch binding domain of glucoamylase from *Aspergillus niger*: overview of its structure, function, and role in raw-starch hydrolysis. *Biologia, Bratislava* **57**: 239-245

Joyet P, Declerck N, Gaillardin C. 1992. Hyperthermostable variants of a highly thermostable α-amylase. *Biotechnology* **10**:1579-1583.

Kadziola A, Søgaard M, Svensson, B, Haser R. 1998. Molecular structure of a barley α-amylase-inhibitor complex: implications for starch binding and catalysis. *J Mol Biol* **278**:205-217.

Kandra L, Gyemant G. 2000. Examination of the active sites of human salivary alpha-amylase (HSA). *Carbohydr Res* **329(3)**:579-585.

Kandra L, Gyemant G, Farkas E, Liptak A. 1997. Action pattern of porcine pancreatic alpha-amylase on three different series of beta-maltooligosaccharide glycosides. *Carbohydr Res*

298(3):237-242.

Kjeldsen T, Brandt J, Andersen AS, Egel-Mitani M, Hach M, Pettersson AF, Vad K. 1996. A removable spacer peptide in an alpha-factor-leader/insulin precursor fusion protein improves processing and concomitant yield of the insulin precursor in *Saccharomyces cerevisiae. Gene* **170(1)**:107-112.

Kondo A, Shigechi H, Abe M, Uyama K, Matsumoto T, Takahashi S, Ueda M, Tanaka A, Kishimoto M, Fukuda H. 2002. High-level ethanol production from starch by a flocculent *Saccharomyces cerevisiae* strain displaying cell-surface glucoamylase. *Appl Microbiol Biotechnol* **58(3)**:291-296.

Kuriki T, Takata H, Okada S, Imanaka T. 1991. Analysis of the active center of *Bacillus stearothermophilus* neopullulanase. *J Bact* **173**:6147-6152.

Lagunas R. 1979. Energetic irrelevance of aerobiosis for *S. cerevisiae* growing on sugars. *Mol Cell Biochem* **27(3)**:139-146.

Lin LL, Ma YJ, Chien HR, Hsu WH. 1998. Construction of an amylolytic yeast by multiple integration of the *Aspergillus awamori* glucoamylase gene into a *Saccharomyces cerevisiae* chromosome *Enzyme Microb Technol* **23**:360-365.

Linderstrøm-Lang K, Nielsen SO. 1959. In *Acid-Base Equilibria of Proteins in Electrophoresis*, Bier M (ed). Academic: New York; 35-89.

Lopes TS, De Wijs IJ, Steenhauer SI, Verbakel J, Planta RJ. 1996. Factors affecting the mitotic stability of high-copy-number integration into the ribosomal DNA of *Saccharomyces cerevisiae*. *Yeast* **12(5)**:467-477.

Ma YJ, Lin LL, Chien HR, Hsu WH. 2000. Efficient utilization of starch by a recombinant strain of *Saccharomyces cerevisiae* producing glucoamylase and isoamylase. *Biotechnol Appl Biochem* 31: 55-59

MacGregor E, Janecek S, Svensson B. 2001. Relationship of sequence and structure to specificity in the α-amylase family of enzymes. *Biochim Biophys Acta* **1546**:1-20.

MacGregor EA, Svensson B. 1989. A super-secondary structure predicted to be common to several alpha-1,4-D-glucan-cleaving enzymes. *Biochem J* **259(1)**:145-152.

Machius M, Wiegand G, Huber R. 1995. Crystal structure of calcium-depleted *Bacillus licheniformis* alpha-amylase at 2.2 Å resolution. *J Mol Biol* **246(4)**:545-559.

Masy CL, Henquinet A, and Mestdagh MM. 1992. Flocculation of *Saccharomyces cerevisiae*: inhibition by sugars. *Can. J. Microbiol*. **38:**1298-306.

Matsui I, Ishikawa K, Miyuri S, Fukui S, Honda K. 1992. A mutant α-amylase with enhanced activity specific for short substrates. *FEMS Lett* **310**:216-218.

Matsuura Y, Kusunoki M, Harada W, Kakudo M. 1984. Structure and possible catalytic residues of Taka-amylase A. *J Biochem (Tokyo)* **95(3)**:697-702.

Matsuura Y, Kusunoki M, Harada W, Tanaka N, Iga Y, Yasuoka N, Toda H, Narita K, Kakudo M. 1980. Molecular structure of taka-amylase A. I. Backbone chain folding at 3 A resolution. *J Biochem (Tokyo)* **87(5)**:1555-1558.

Matsuura Y. 2002. A possible mechanism of catalysis involving three essential residues in the enzymes of alpha-amylase family. *Biologia, Bratislava* **57**:21-27.

Matthews BW, Nicholson H, Becktel WJ. 1987. Enhanced protein thermo-stability from site-directed mutations that decrease the entropy of unfolding. *Proc Natl Acad Sci USA* **84**: 6663-6667.

Miyazaki K, Wintrode PL, Grayling RA, Rubingh DN, Arnold FH. 2000. Directed evolution study of temperature adaptation in a psychrophilic enzyme. *J Mol Biol* **297**:1015-1026.

Montesinos T, Navarro J. 2000. Production of alcohol from raw wheat flour by amyloglucosidase and *Saccharomyces cerevisiae*. *Enzyme Microb Technol* **27(6)**:362-370.

Mori H, Bak-Jensen KS, Svensson B. 2002. Barley alpha-amylase Met53 situated at the high-affinity subsite -2 belongs to a substrate binding motif in the beta-->alpha loop 2 of the catalytic (beta/alpha)8-barrel and is critical for activity and substrate specificity. *Eur J Biochem* **269(22)**:5377-5390.

Nagashima T, Toda S, Kitamoto K, Gomi K, Kumagai C, Toda H. 1992. Site directed mutagenesis of the catalytic active site residues of Taka-amylase A. *Biosci Biotech Biochem* **56**:207-210.

Nakajima R, Imanaka T, Aiba S. 1986. Comparison of amino acid sequences of eleven different α-amylases. *Appl Microbiol Biotechnol* **23**:355-360.

Nakamura A, Haga K, Yamane K. 1993. Three histidine residues in the active center of cyclodextrin glucanotransferase from alkalophilic *Bacillus* sp.1011: effects of replacement on pH dependence and transition state stabilization. *Biochemistry* **32**:6624-6631.

Nakamura A, Haga K, Ogawa S, Kuwano K, Kimura K, Yamane K. 1992. Functional relationships between cyclodextrin glucanotransferase from an alkalophilic *Bacillus* and alpha-amylases. Site-directed mutagenesis of the conserved two Asp and one Glu residues. *FEBS Lett* **296(1)**:37-40.

Nielsen JE, Borchert TV. 2000. Protein engineering of bacterial α-amylases. *Biochim Biophys Acta* **1543**:253-274.

Nielsen JE, Beier L, Otzen D, Borchert TV, Frantzen HB, Andersen KV, Svendsen A. 1999. Electrostatics in the active site of an α-amylase. *Eur J Biochem* **264**:816-824.

Nitta Y, Mizushima M, Hiromi K, Ono S. 1971. Influence of molecular structures of substrates and analogues on Taka-amylase A catalyzed hydrolyses. I. Effect of chain length of linear substrates. *J Biochem (Tokyo)* **69(3)**:567-576.

Ohdan K, Kuriki T, Takata H, Kaneko H, Okada S. 2000. Introduction of raw starch-binding domains into *Bacillus subtilis* alpha-amylase by fusion with the starch-binding domain of *Bacillus* cyclomaltodextrin glucanotransferase. *Appl Environ Microbiol* **66(7)**:3058-3064.

Penninga D, Vanderveen BA, Knegtel RMA, Vanhijuin SAFT, Rozeboom HJ, Kalk KH, Dijkstra BW. 1996. The raw starch-binding domain of cyclodextrin glucosyltransferase from *Bacillus circulans* strain 251. *J Biol Chem* **271**:32777-32784.

Pretorius IS, Marmur J. 1988. Localization of yeast glucoamylase genes by PFGE and OFAGE. *Curr Genet* **14(1)**:9-13.

Pretorius IS, Modena D, Vanoni M, Englard S, Marmur J. 1986. Transcriptional control of glucoamylase synthesis in vegetatively growing and sporulating *Saccharomyces* species. *Mol Cell Biol* **6(9)**:3034-3041.

Przylas I, Tomoo K, Terada Y, Takaha T, Fujii K, Saenger W, Strater N. 2000. Crystal structure of amylomaltase from *Thermus aquaticus*, a glycosyltransferase catalysing the production of large cyclic glucans. *J Mol Biol* **296(3)**:873-886.

Qian M, Haser R, Buisson G, Duee E, Payan F. 1994. The active center of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2-Å resolution. *Biochemistry* **33**:6284-6294.

Sato N, Matsumoto T, Ueda M, Tanaka A, Fukuda H, Kondo A. 2002. Long anchor using Flo1 protein enhances reactivity of cell surface-displayed glucoamylase to polymer substrates. *Appl Microbiol Biotechnol* **60(4)**:469-474.

Sheridan PP, Panasik N, Coombs JM, Brenchley JE. 2000. Approaches for deciphering the structural basis of low temperature enzyme activity. *Biochim Biophys Acta* **1543(2)**:417-433.

Shibuya I I, Tamura G, Ishikawa T, Hara S. 1992a. Cloning of the alpha-amylase cDNA of *Aspergillus shirousamii* and its expression in *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* **56(2)**:174-179.

Shibuya I, Tamura G, Shima H, Ishikawa T, Hara S. 1992b. Construction of an alpha-amylase/glucoamylase fusion gene and its expression in *Saccharomyces cerevisiae*. *Biosci Biotechnol Biochem* **56(6)**:884-889.

Shigechi H, Koh J, Fujita Y, Matsumoto T, Bito Y, Ueda M, Satoh E, Fukuda H, Kondo A. 2004, Direct production of ethanol from raw corn starch via fermentation by use of a novel surface-engineered yeast strain codisplaying glucoamylase and alpha-amylase. *Appl Environ Microbiol* **70(8)**:5037-5040.

Shih P, Holland DR, Kirsch JF. 1995. Thermal stability determinants of chicken egg-white lysozyme core mutants: hydrophobicity, packaging volume, and conserved buried water molecules. *Protein Sci* **4**:2050-2062.

Sogaard M, Kadziola A, Haser R, Svensson B. 1993. Site-directed mutagenesis of histidine 93, aspartic acid 180, glutamic acid 205, histidine 290, and aspartic acid 291 at the active site and tryptophan 279 at the raw starch binding site in barley α-amylase 1. *J Biol Chem* **268**:22480-22484.

Sorimachi K, Le Gal-Coeffet MF, Williamson G, Archer DB, Williamson MP. 1997. Solution structure of the granular starch binding domain of *Aspergillus niger* glucoamylase bound to beta-cyclodextrin. *Structure* **5(5)**:647-661.

Steyn AJ, Pretorius IS. 1991. Co-expression of a *Saccharomyces diastaticus* glucoamylase-encoding gene and a *Bacillus amyloliquefaciens* alpha-amylase-encoding gene in *Saccharomyces cerevisiae*. *Gene* **100**:85-93.

Strokopytov B, Knegtel RMA, Penninga D, Rozeboom HJ, Kalk KH, Dijkhuizen L, Dijkstra BW. 1996. Structure of cyclodextrin glycosyltransferase complexed with a maltononaose inhibitor at 2.6 Å resolution. Implications for product specificity. *Biochemistry* **35**:4241-4249.

Strokopytov B, Penninga D, Rozeboom HJ, Kalk KH, Dijkhuizen L, Dijkstra BW. 1995. X-ray structure of cyclodextrin glycosyltransferase complexed with acarbose. Implications for the catalytic mechanism of glycosidases. *Biochemistry* **34**:2234-2240.

Suganama T, Matsuno R, Ohnishi M, Hiromi K. 1978. A study of the mechanism of action of Taka-amylase A^1 on linear oligosaccharides by product analysis and computer simulation. *J Biochem* **84**:293-316.

Suzuki Y, Ito N, Yuuki T, Yamagata H, Udaka S. 1989. Amino acid residues stabilizing a *Bacillus* alpha-amylase against irreversible thermoinactivation. *J Biol Chem* **264(32)**:18933-18938.

Svensson B. 1994. Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity and stability. *Plant molecular biology* **25**:141-157.

Swift HJ, Brady L, Derewenda ZS, Dodson EJ, Dodson GG, Turkenburg JP, Wilkinson AJ. 1991. Structure and molecular model refinement of *Aspergillus oryzae* (TAKA) alpha-amylase: an application of the simulated-annealing method. *Acta Crystallogr B* **47**:535-544.

Takase K, Matsumoto T, Mizuno H, Yamane K. 1993. Site directed mutagenesis of active site residues near the catalytic site on the activity of *Bacillus stearothermophilus* α-amylase. *Eur J Biochem* **211**:899-902.

Takase K. 1992 Interaction of catalytic-site mutants of *Bacillus subtilis* α-amylase with substrates and acarbose. *Biochim Biophys Acta* **1122**:278-282.

Takata H, Kuriki T, Okada S, Takesada Y, Iizuka M, Minamiura N, Imanaka T. 1992. Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at alpha-(1----4)- and alpha-(1----6)-glucosidic linkages. *J Biol Chem* **267(26)**:18447-18452.

Tomazic SJ, Klibanov AM. 1988. Why is one *Bacillus* α-amylase more resistant against irreversible thermoinactivation than another? *J Biol Chem* **263**:3092-3096.

Uitdehaag JC, van Alebeek GJ, van Der Veen BA, Dijkhuizen L, Dijksta BW. 2000. Structures of maltohexaose and maltopentaose bound at the donor sites of cyclodextrin glycosyltransferase give insight into mechanisms of transglycosylation activity and cyclodextrin size specificity. *Biochemistry* **39**:7772-7780.

Van der Maarel MJ, Van der Veen B, Uitdehaag JC, Leemhuis H, Dijkhuizen L. 2002. Properties and applications of starch-converting enzymes of the alpha-amylase family. *J Biotechnol* **94(2)**:137-155.

Vihinen M, Ollikka P, Niskanen J, Meyer P, Suominen I, Karp M, Holm L, Knowles J, Mantsala P. 1990, Site-directed mutagenesis of a thermostable alpha-amylase from *Bacillus stearothermophilus*: putative role of three conserved residues. *J Biochem (Tokyo)* **107(2)**:267-272.

Visser W. 1995. Oxygen requirements of fermentative yeasts. PhD thesis. Delft University of Technology.

Williamson MP, Le Gal-Coeffet MF, Sorimachi K, Furniss CS, Archer DB, Williamson G. 1997. Function of conserved tryptophans in the *Aspergillus niger* glucoamylase 1 starch binding domain. *Biochemistry* **36(24)**:7535-7539.

Wind RD, Uitdehaag JC, Buitelaar RM, Dijkstra BW, Dijkhuizen L. 1998. Engineering of cyclodextrin product specificity and pH optima of the thermostable cyclodextrin glycosyltransferase from *Thermoanaerobacterium thermosulfurigenes* EM1. *J Biol Chem* **273**:5771-5779.

Yamashita M, Matsumoto D, Murooka Y. 1997. Amino acid residues specific for the catalytic action towards α-1,6 glucosidic linkages in *Klebsiella pneumoniae*. *Journal of Fermentation and Bioengineering* **84**:283-290.

Yoshioka Y, Hasegawa K, Matsuura Y, Katsube Y, Kubota M. 1997. Crystal structures of a mutant maltotetraose-forming exo-amylase cocrystallized with maltopentaose. *J Mol Biol* **271**:619-628.

Zavodsky P, Kardos J, Svingor A, Petsko GA. 1998. Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc Natl Acad Sci USA* **95**:7406-7411.

Thoma JA , Spradlin JE, and Dygert, S; Plant and animal amylases, in *The Enzymes, 3rd Ed., Vol. V: Hydrolysis (Sulfate Esters, Carboxyl Esters, Glycosides) and Hydration*, ed. by P.D. Boyer. 1971: Academic Press, New York, pp. 115-189. A classic review of the a-amylases of mammals and higher plants.

Takagi T, Toda H, Isemura I ; Bacterial and mold amylases, in *The Enzymes, 3rd Ed., Vol. V: Hydrolysis (Sulfate Esters, Carboxyl Esters, Glycosides) and Hydration*, ed. by P.D. Boyer. 1971: Academic Press, New York, pp. 235-271. The classic review of the microbiological a-amylases.

CHAPTER 3

RESEARCH RESULTS I

The role of the N-terminal region of the *Lipomyces kononenkoae LKA1*- and *LKA2*-encoded α-amylases in substrate binding, thermal stability and substrate specificity

This manuscript will be submitted for publication in *European Journal of Biochemistry.*

Role of the N-terminal region of the *Lipomyces kononenkoae LKA1* and *LKA2* encoded α-amylases in substrate binding, thermal stability and substrate specificity

N. Ramachandran[1], I. S. Pretorius[1, 2], R. R. Cordero Otero[1]

[1]Institute for Wine Biotechnology, Department of Viticulture & Oenology, Stellenbosch University, Victoria Street, Stellenbosch, ZA 7602, South Africa; [2]The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, Australia.

1. ABSTRACT

Lipomyces kononenkoae is considered as one of the most efficient raw-starch-degrading yeasts known. It. Previously, we have cloned and characterised two of its α-amylase-encoding genes, *LKA1* and *LKA2*. In addition to these α-amylases' activity on the α-1,4-glucosidic linkages of starch, Lka1p has a side activity on the α-1,6-glucosidic bonds of pullulan, while Lka2p has considerable dextrinase activity. Primary sequence analyses of Lka1p and Lka2p revealed the high homology of these proteins to TAKA amylase from *Aspergillus oryzae*. Using TAKA amylase as the reference protein, homology modelling of Lka1p and Lka2p was performed to identify the domain organisation of these two *L. kononenkoae* α-amylases. Taking the computer coupling alignments and comparative secondary structure data into account, mutant Lka1p and Lka2p proteins lacking the N-terminal sequences preceding their catalytic domains were created. The N-terminal deletions of Lka1p (Lka1pΔN) and Lka2p (Lka2pΔN) retained their starch-degrading capability. The wild-type (Lka1p and Lka2p) and truncated (Lka1pΔN and Lka2pΔN) α-amylases were analysed for their raw starch-binding ability, thermostability, calcium ion dependence and substrate specificity. Here we report the role of the N-terminal sequences of Lka1p and Lka2p in raw starch adsorption, specificity and thermostability. Lka1pΔN showed a 95% reduction in raw starch adsorption, 20% lower hydrolytic activity on complex raw starch and less thermostability than the wild-type protein. On the other hand, Lka2pΔN did not show significant variations in raw starch adsorption or thermostability. Furthermore, the catalytic efficiency of Lka2pΔN on dextrin was 44% lower than that of Lka2p. The results discussed in this paper suggest that the N-terminal sequence in Lka1p is important for adsorption to, and hydrolysis of raw starch substrate.

2. INTRODUCTION

To date, no naturally occurring yeast has been described that can cost-effectively convert inexpensive starch-rich materials into potable or biofuel ethanol in a single step fermentation process. Therefore, it would be highly desirable to develop yeast strains that possess a combination of efficient starch utilisation (e.g. high-level production of raw-starch degrading amylases, and efficient uptake and utilisation of resulting sugars) and high ethanol yields [1]. The development of such yeasts can proceed via either of two strategies, namely (i) the improvement of desired production properties of a naturally amylolytic yeast species having excellent starch utilisation capabilities or, alternatively, (ii) the desired amylolytic properties could be conferred to a yeast having optimal ethanol production properties [1]. Our laboratory has opted for the latter strategy, by attempting to extend the limited range of carbohydrates from which the industrially versatile yeast, *Saccharomyces cerevisiae*, is capable of efficiently producing ethanol [2, 3].

Lipomyces kononenkoae is one of the most efficient starch-degrading yeasts known [4]. This lipid-forming yeast secretes a consortium of highly efficient α-amylase, glucoamylase [5], isoamylase [6] and cyclodextrinase [7] activities but, unfortunately, it cannot be used in existing alcohol fermentation processes because of its low ethanol tolerance, slow growth rate and catabolite repression. On the other hand, *S. cerevisiae,* the prime choice for industrial production of potable and biofuel ethanol [8], lacks efficient amylolytic enzymes to convert starchy biomass to ethanol. In an attempt to overcome this deficiency, our laboratory has cloned and expressed two α-amylase genes (*LKA1* and *LKA2*) from *L. kononenkoae* in *S. cerevisiae* [2, 3, 9, 10]. The *L. kononenkoae LKA1*- and *LKA2*-encoded α-amylases, Lka1p and Lka2p, not only have high specificity for α-1,4 glucosidic linkages, but also show activity on starchy substrates containing α-1,6 glucosidic bonds. In particular, Lka1p has a side activity on pullulan, while Lka2p has significant dextrinase activity [9, 10]. This dual bond specificity, which is usually observed in α-amylase/pullulanase enzymes, makes these enzymes excellent candidates for the development of amylolytic *S. cerevisae* strains.

The α-amylase family (EC 3.2.1.1) to which Lka1p and Lka2p belong, consists of members with different specificities, including glycosyl transferases, hydrolases and isomerases. These enzymes specifically act on the α-1,4 linkages of starch; however, some of them have been reported to contain functional, in addition to their catalytic domains [11]. There is a high degree of amino acid conservation in the sequences of these amylolytic enzymes, although many different domain organisations have been reported [12]. They share a similar catalytic TIM barrel and most α-amylases contain domain B, which protrudes between the third α-helix and the third β-strand of the catalytic domain. Domain C follows the

catalytic domain, consists of β-strands and is usually involved in stability of the catalytic domain. Some members of the family possess an additional N-terminal domain, the role of which has been implicated in specificity, stability and starch binding. The structure of α-amylase family members with such an N-terminal extra domain has not yet been evaluated in detail.

In the present paper, we report on the potential domain architecture of the two *L. kononenkoae* α-amylases, Lka1p and Lka2p. Lka1p possesses an additional N-terminal domain and we have investigated the possible role of this region in raw starch digestion and adsorption. The contribution of the N-terminal regions in both Lka1p and Lka2p to specificity towards various substrates, and the contribution to thermal stability are also discussed.

3. MATERIALS AND METHODS

3.1 PCR mutagenesis, sequencing and plasmid construction

The polymerase chain reaction (PCR) technique was used to amplify copies of *LKA1* and *LKA2* in which the N-terminal-encoded sequences were deleted. The full-length *LKA1* and the truncated version (*LKA1ΔN*) were synthesised by using the 5' PCR primers, FLKA1Δ28 5'-ATGGATTGCACTACAGTTAC-3' and FLKA1Δ160 5'-ATGGTCGTAACTGATAGATT-3', coupled in separate reactions to the 3' primer RLKA1 5'-CTACATGGAGCAGATTCC-3'; both DNA fragments were then purified using High Pure™ PCR product (Boehringer Mannheim) and cloned into the *E. coli* PGEM®-T expression vector (Promega) to obtain the plasmids pLKA1 and pLKA1ΔN (deleted for the first 132 residues), respectively. Using the *LKA2* gene as template DNA, two PCR products were developed using the 5' primers FLKA2Δ23 5'-ATGAAGACCGCGGCAGAAT-3' and FLKA2Δ39 5'-ATGGTCATAACCGATCGTTTC-3' in combination with the 3' primer 5'-TTAAGAACAAAATTTCCCAGCA-3', and the cloned PCR products generated the plasmids pLKA2 and pLKA2ΔN (deleted for the first 16 residues), respectively. None of these amplified sequences contained the fragment encoding the native secretion signal. The PCR reactions contained PCR buffer with 2 mM $MgSO_4$ (BIOTAQ™), 0.8 mM dNTPs, 0.3 μM of each primer, 0.1 μg template, and 5 units of Ex-*Taq* DNA polymerase (Takara) in a total volume of 50 μL. A thermocycler PCR Express gradient (Hybaid Lt. Ashford) was used for DNA amplification under the following conditions: 30 s at 94ºC, 30 s at 48ºC and 45 s at 68ºC. Twenty cycles were run, with a subsequent polymerisation period of 5 min at 68ºC. All of the cloned DNA fragments were verified by sequencing.

Standard recombinant DNA methods were used for the construction of plasmids. The analysis of the sequence of the PCR-generated fragments was carried out using an ABI PRISM® Big Dye™ 377 DNA sequencer (PE/Applied Biosystems). Both the coding and the

non-coding strands were sequenced to ensure the reliable identification of deletions and the absence of unwanted mutations.

3.2 Homology modelling

Sequence-based homology predictions were done using the Clustal W program (European Bioinformatics Institute). Sequences were retrieved from the protein databases SWISSPROT and EMBL. Predict protein software was used to determine the secondary structures of Lka1p and Lka2p. α-Amylase from *Aspergillus oryzae* (TAKA amylase) was used as the 3D structure reference for tertiary coupling structure predictions using the Cn3D 4.0 program (National Center For Biotechnology Information).

3.3 Enzyme purification

E. coli DH5α expressing different α-amylases Lka1p, Lka1pΔN, Lka2p and Lka2pΔN were grown separately at 37°C in 50 mL Luria-Bertani broth medium containing 100 μg mL^{-1} ampicillin [36]. *E. coli* cells were harvested by centrifugation at 2 300 g for 15 min at 4°C, and the supernatants were independently applied to a Spectra/Gel Fast Flow DEAE Sepharose™ column, 2.5 by 20 cm (Pharmacia Amersham, Uppsala, Sweden), which had previously been equilibrated with 50 mM MES (2-[N-Morpholino] ethanesulfonic acid, Sigma Catalogue # M 5287) buffer at pH 6.0. The column containing both the supernatant and 100 mM MES in a ratio of 1:1 was incubated in an end-over-end rotary mixer at 40 rpm at 4°C for 2 h. After being equilibrated with five column-volumes of 30 mM Tris-HCl at pH 8, the bound amylase was eluted using a linear 0.2-1 M NaCl gradient in 30 mM of Tris-HCl (pH 8) at a flow rate of 10 mL h^{-1}, and desalted with a Microcon® ultra-filtration device YM-3. The fractions containing amylolytic activity were pooled and concentrated by ultra-filtration with a Microcon® device YM-50. The purified proteins were runon SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) and the corresponding western blot of the proteins was performed.

Protein concentrations were determined using a Bio-Rad protein assay kit (Bio-Rad Laboratories, GmbH) with BSA (Promega, Madison WI, USA) as the standard.

3.4 SDS-PAGE and Western blot analysis

The molecular weight of the purified enzymes was determined by SDS-PAGE according to the manufacturer's recommendations (Hoeffer Scientific Instruments, San Francisco, USA) and visualised by silver staining [37]. The molecular masses of both the full-length and truncated α-amylases were estimated according to the method of Laemmli [38], with a high molecular-mass standard mixture purchased from Sigma. For the Western blotting analyses, the resolved proteins were transferred with a SEMI-PHOR system (Hoeffer Scientific

Instruments) onto an immobilon^{TM}–P transfer membrane (Millipore Corporation, MA, USA). The blotted proteins were identified immunochemically by sequential addition of rabbit polyclonal anti-α-amylase (Abcam Ltd, Cambridgeshire, UK), followed by donkey anti-rabbit Ig-horse radish peroxidase-linked whole antibody (Amersham Pharmacia Ltd, UK). The secondary antibody was detected using the ECL Western blotting detection system (Amersham Pharmacia Ltd, UK).

3.5 Immunochemical quantification of proteins

The protein samples to be quantified were adsorbed on sterilin (Bibby-sterilin Ltd, Staffordshire, UK) immunoplates overnight at 4°C. ELISA was performed as described by the suppliers of primary antibodies (Abcam Ltd, Cambridgeshire, UK). To prevent non-specific binding, the wells were blocked with 300 µL of 1% skim milk powder at 37°C for 2 h. Phosphate buffered saline (pH 7.5) was used as a washing buffer. After the plate had been washed twice, rabbit polyclonal anti-α-amylase (Abcam Ltd, Cambridgeshire, UK) was added and the plate was incubated at room temperature for 2 h, followed by conjugation with anti-rabbit secondary antibody linked to alkaline phosphatase. The binding of α-amylases was detected using alkaline phosphatase and *p*-nitro phenol phosphate substrate. The release of *p*-nitro phenol was measured by absorbance at 405 nm using a microtitre plate reader (Powerwave X Bio-Tek Instruments). α-Amylase from *Bacillus amyloliquefaciens* (Sigma, St. Luis, Mo.) at different concentrations was used to derive the ELISA standard for quantification.

3.6 Enzyme activity assay

Raw corn starch, soluble potato starch, amylopectin, amylose, pullulan, cyclodextrin, dextrin, maltotriose and panose were purchased from Sigma, St. Luis, MO. to study the hydrolytic pattern of the enzymes.

The activities of Lka1p, Lka2p and the N-terminal truncated variants, Lka1pΔN and Lka2pΔN, lacking the first 159 and 39 amino acids, respectively, were assayed at 50°C for 30 min in a 20 mM sodium acetate buffer at pH 6.8. The activity was assayed by monitoring the release of reducing sugars using the dinitrosalicylic acid method [39]. One unit of activity of standard α-amylase was defined as the amount of enzyme that liberated 1 mg of reducing sugar per min. The amount of reducing sugar released was assessed from a maltose standard curve. The kinetic parameters, V_{max} (µmol min^{-1} mg^{-1}) and the Michaelis-Menten constant (*K*m, mM) were determined from Lineweaver- Burk plots of specific activities at various substrate concentrations. The concentrations of wild-type (Lka1p and Lka2p) and mutant (Lka1pΔN and Lka2pΔN) enzymes were determined using immunochemical quantification and through standard protein assays. The concentration of enzymes in µM was

deduced from the molecular masses of the proteins to allow calculation of the catalytic rate constant (k_{cat}) from the relationship $k_{cat} = V_{max}/[E_0]$, where $[E_0]$ is equal to the total enzyme concentration [40].

3.7 Starch-binding assay

Adsorption standard assay was modified from Gal *et al*. [41]. Wild-type (Lka1p and Lka2p) and mutant (Lka1pΔN and Lka2pΔN) proteins of varied concentrations (1, 2, 3 and 4 μmol) were mixed with 10 mg raw cornstarch suspended in 50 mM citrate buffer at pH 5.5 in a final volume of 0.5 mL in individual reactions. The reaction mixtures were incubated for 3 h at 4°C in an end-over-end rotary mixer at 40 rpm, followed by centrifugation at 15 000 x *g* for 5 min at 4°C. The concentration of α-amylase in the supernatant assayed, and the amount of absorbed enzyme, were calculated by subtraction [42]. The percentage adsorption was calculated using the formula $\%Ads = [(E-F)\cdot E^{-1}]\cdot 100$, where E is the total of enzyme units mixed with the substrate, and F is the amount of free eluted units.

3.8 Temperature profile and stability of enzymes

The effect of temperature on enzyme activity was measured at temperatures ranging from 30 to 90°C after 30 min of incubation in 20 mM sodium acetate (pH 6.8). The thermostability of the wild-type enzymes and their truncated variants was monitored by measuring the residual enzymatic activity after heat treatment at 50°C in 20 mM sodium acetate (pH 6.8). At different time points, 200 μL samples were withdrawn and chilled on ice for 30 min before the residual activity was evaluated.

3.9 Effect of Ca^{2+} on amylase activity

The activity of the wild-type (Lka1p and Lka2p) and the mutant (Lka1pΔN and Lka2pΔN) enzymes was measured on insoluble starch under a varying calcium concentration (0-22 mM) and an EDTA concentration of 1 mM and 5 mM in 20 mM sodium acetate (pH 6.8) and 0.5 mg mL^{-1} BSA at 50°C. Standard deviations were calculated from the data obtained from triplicate experiments performed independently.

4. RESULTS

4.1 Production and purification of α-amylases

The PGEM®-T α-amylase constructs (pLKA1, pLKA1-ΔN, pLKA2 and pLKA2-ΔN) were transformed independently into *E. coli* DH5α and tested for the production of α-amylase. The recombinant transformants were cultured in Luria Bertrani medium until the late logarithmic

phase and the amylolytic activity on soluble starch was measured in the supernatant of the cultures. The theoretical molecular weight of Lka1p was deduced to be 65 kDa, while the truncation of the N-terminal sequence resulted in a 51 kDa protein. The size of Lka2p is approximately 55 kDa and its truncated form 51 kDa, which is consistent with the sizes observed on SDS-PAGE gel (Fig. 1A) and by Western blot (Fig. 1B). For the activity assays on different substrates, 0.5 μM of enzyme was used, while 1 μM was used for the kinetic characterisation.

4.2 Substrate specificity of wild-type and mutant α-amylases

The *LKA1* and *LKA2* encoded α-amylases have the capacity to act on a range of starch-related substrates with different bonding patterns. Consequently, the relative activity of the purified wild-type (Lka1p and Lka2p) and the mutant (Lka1pΔN and Lka2pΔN) α-amylases was determined for various starchy substrates, containing either α-1,4 glucosidic linkages or a mixture of α-1,4 and α-1,6 glucosidic linkages (Table 1). Lka1p effectively hydrolysed the α-1,4 and α-1,6 linkages in corn starch, amylopectin and α-cyclodextrins. It showed a lower degree of specificity to pullulan, which has many α-1,6 linkages [9]. The full-length Lka1p showed the best hydrolytic activity, producing 4.9 μmol of maltose from the amylose fraction of starch and 3.8 μmol maltose from amylopectin. It had a broad specificity towards pullulan, cyclodextrin and dextrin, releasing 4.1, 2.8 and 3 μmol of reducing sugar, respectively. However, Lka1p had a very low activity on maltotriose, while there was no detectable activity on panose. The mutant Lka1p-ΔN showed an overall decrease in activity on all of the aforementioned substrates, while the decrease in activity on raw starch was pronounced. The enzyme lost 30% of its original hydrolytic activity on raw starch, 22% on α-cyclodextrin, 17% on amylose and amylopectin and 10% on pullulan. The loss of hydrolytic activity towards raw starch was more pronounced than towards soluble starch substrates, suggesting that the N-terminus of Lka1p might play an important role in raw starch hydrolysis. The hydrolytic activity of the native Lka2p and its N-terminal deletion mutant were very similar on soluble starch, raw starch, amylose, amylopectin, α-cyclodextrin and pullulan. However, Lka2p-ΔN hydrolysed dextrin at a rate of 45% lower than the wild-type Lka2p. Similarly to Lka1p, the Lka2p and Lka2p-ΔN enzymes were unable to hydrolyse panose and maltotriose. The N-terminal sequence deleted in Lka2p is very short compared to the N-terminal sequence truncated in Lka1p. The 45% decrease in dextrin hydrolysis by Lka2pΔN could possibly be due to amino acids present in this short sequence that promote specificity towards this substrate.

4.3 Kinetic parameters of wild-type and mutant α-amylases

The kinetic constants of the α-amylases, which are summarised in Table 2, were determined for 1 μm of purified enzyme at seven different substrate concentrations (4, 8, 16, 20, 32, 64 and 128 mg mL^{-1}) at 50°C and pH 6.8. Lka1p-ΔN showed a nearly two-fold higher K_m value on raw starch, and approximately 1.0-fold and 1.5-fold higher on pullulan and dextrin, respectively. The V_{max} of raw starch hydrolysis by Lka1p-ΔN was 3.2% lower than that of the wild-type enzyme and 4.2% lower on dextrin, while pullulan hydrolysis increased by 4.9% in comparison with that of the wild-type enzyme. However, the overall catalytic efficiency (K_{cat}/K_m) for Lka1p-ΔN was lower on all the forementioned substrates compared to that of the wild-type protein. Furthermore, the catalytic efficiency of Lka1p-ΔN on raw starch reduced to 48% of the activity of the wild-type protein, whilst on the other substrates the decrease was 30%. On the other hand, the K_m values for both Lka2p and Lka2p-ΔN on raw starch and dextrin were similar, although the K_m value for the mutant was 0.5-fold higher on pullulan. Noticeably, the K_{cat} value for the full-length Lka2p was 1.2 and 1.6-fold higher than that of the mutant enzyme on pullulan and dextrin, respectively. Deletion of the N-terminal sequence of Lka2p did not affect the K_{cat} value of this enzyme on raw starch. As for Lka2p, the truncated mutant had a similar K_{cat}/K_m value for raw starch and pullulan, while the wild-type enzyme exhibited a 58% higher catalytic efficiency for dextrin than the truncated Lka2p.

4.4 Temperature profiles of wild-type and mutant α-amylases

Lka1p has a temperature optimum of about 50°C [9], while that of Lka2p is from 50 to 70°C [12]. To investigate whether the truncation of these amylases caused any change in the temperature optimum, their temperature profiles were investigated over the range 30-90°C (Fig. 2A). The optimum temperature for Lka1p and Lka1p-ΔN was 50°C, and they had similar activity profiles across the range of temperatures tested. Nevertheless, Lka1p exhibited 1.7-fold higher hydrolysis than Lka1p-ΔN at 30°C and at temperatures above 70°C, Lka1p activity reduced less rapidly than that of Lka1p-ΔN. A four-fold decrease in the residual percentage hydrolysis of starch was observed in Lka1p-ΔN at temperatures above 70°C. These results imply that Lka1p-ΔN exhibits lower activity than Lka1p at temperatures other than 50°C, especially above 70°C, suggesting these proteins have different thermal profiles. Lka2p and Lka2p-ΔN have a temperature optimum of around 50°C. There was a marked two-fold decrease in activity at temperatures above 70°C for these enzymes, suggesting a mesophilic thermal character. Physicochemical and kinetic characterisation were subsequently carried out at 50°C.

4.5 Thermostability of wild-type and mutant α-amylases

To investigate the effect of the N-terminal deletions on thermal stability of the enzymes, residual activity was measured after incubating the proteins at 50°C for various lengths of time (Fig. 2B). Lka1p retained its maximal activity after 68 h and an outstanding 90% of its original activity maximum after 96 h of incubation. In contrast, the truncated Lka1p showed a residual activity of 20% and 35% after 30 h and 80 h, respectively. Lka2p and Lka2p-ΔN amylases showed a similar pattern of gradual decrease in percentage residual activity after 24 h. After 96 h of incubation, Lka2p and Lka2p-ΔN retained 65 and 75% of their maximum activity, respectively. It is evident that Lka1p-ΔN is less stable at 50°C over time when compared to the wild-type Lka1p. It was also noted that Lka2p-ΔN was slightly more stable than the wild-type Lka2p.

4.6 Adsorption of wild-type and mutant α-amylases on raw starch

The adsorption of the Lka1p, Lka2p and N-terminal deleted enzymes (Lka1p-ΔN and Lka2p-ΔN) on raw starch was assayed at various time intervals and enzyme dilutions. Approximately 65% of the native Lka1p adsorbed to raw starch after an incubation period of 3 h at 4°C. Adsorption was near saturation when 1 μmol of enzyme was incubated with 10 mg mL^{-1} of raw starch under standard assay conditions (as described in materials and methods), limiting both the availability and accessibility of the substrate to the enzyme. A 3.5-fold lower magnitude of adsorption was recorded for Lka1p-ΔN under similar conditions, although less than 5% residual adsorption was noticed (Fig. 3), suggesting that the N-terminal sequence of Lka1p plays an essential role in raw starch adsorption. The Lka2p and Lka2p-ΔN enzymes had less than 20% raw starch adsorption, indicating inefficient binding of these enzymes.

4.7 Effect of calcium ions on the activity of wild-type and mutant α-amylases

Titration with 1 to 10 mM of $CaCl_2$ did not give rise any distinct variations in enzyme activity on insoluble starch (Fig. 4). A loss of 10% of maximum activity with an increasing concentration of Ca^{2+}, from 12 to 22 mM, was evident for all the enzymes, and a loss of activity was more pronounced for the truncated enzymes than the full-length α-amylases after 15 min. The activity of for Lka1p-ΔN and Lka2p-ΔN decreased with increasing Ca^{2+}. EDTA did not have any influence on the enzyme activity at concentrations of 1 and 10 mM (data not shown). The neutral effect of the chelating agent EDTA suggests that the active site of this enzyme is a not dependent on a metallic cofactor.

5. DISCUSSION

The primary structure of the *L. kononenkoae* α-amylases, Lka1p and Lka2p, are similar with a predicted common, super secondary fold (β/α)8 barrel structure shared with other members of the α-amylase family. The α-amylase from *Aspergillus oryzae* (TAKA amylase) [13] was chosen as a reference for structural interpretations, due to the abundant structural information available on this protein.

The sequence alignment of Lka1p and Lka2p with TAKA amylase shows similarities in the $(\beta/\alpha)_8$ catalytic domain and the C-terminal- stabilising domain C, which forms a Greek key motif in α-amylases. Sequence comparison also led to the mapping of four highly conserved regions in the catalytic domain, which comprise the active site. These amino acids are located in loops connecting the α-helices and β-strands of the catalytic barrel. They include histidine, aspartic acid and glutamic acid residues corresponding to the His122, Asp206, Glu230 and Asp297 of TAKA amylase [14, 15, 16]. However, variations within the four conserved sequences have been recorded for many α-amylases, perhaps being indicative of enzyme specificity [16].

In addition to the catalytic and C-terminal domains, Lka1p has an N-terminal region of 132 amino acids. The N-terminal sequence preceding the catalytic region in Lka2p or TAKA amylase is much shorter than the N-terminal extension in Lka1p. In addition, the *LKA1* and *LKA2* encoded α-amylases share sequence identity with enzymes denoted as cyclomaltodextrinases, maltogenic amylases or neopullulanases. Some members of the α-amylase family are distinguished by the presence of a novel N-terminal domain and by preferential specificities for starch-related substrates over starch [17]. In this context, the neopullulanase from *Bacillus stearothermophilus* also contains an additional N-terminal domain; it has been proposed that this enzyme has a wider substrate-binding cleft due to the specific spatial arrangement of amino acids facilitating the binding and cleavage of a complex pullulan substrate. It is well known that several glycosyl hydrolases have one or more N-terminal domains preceding the TIM barrel; such domains have occasionally been named domain N, although they are not all structurally related [17].

C-terminal starch-binding domains (SBD) that consist of several β-strands and form an open-sided distorted barrel structure have been found in at least 43 members of the α-amylase family. Detailed analyses of all SBDs has revealed that they evolved indepenently of catalytic domains [18]. The proposed primary function of the SBD is to bind tightly to starch in a perpendicular orientation, implying that the SBD does more than merely locate the catalytic

domain onto the substrate surface leading to physical distortion of the substrate[19]; presumably, SBDs facilitate starch hydrolases to bind to and digest raw granular starch [20].

The sequence comparison with known SBD suggests that the N-terminal sequence of Lka1p has a 36% identity with the N-terminal SBD of *Rhizopus oryzae* glucoamylase (Fig. 5) [21]. The latter is the only N-terminal SBD that has been reported thus far and there is no significant sequence similarity with the C-terminal SBDs. In addition to their evolutionary solitude, N-terminal SBDs might present a new type of SBD [22, 23]. The present investigation using deletion mutants reveals that this motif of Lka1p confers binding to raw starch. In contrast, Lka2p lacks this long N-terminal extension and did not adsorb to raw starch. Previous reports on the structure of α-amylases, such as the cyclomaltodextrinases and maltogenic amylases of *Thermus* strain TVAII (ThMA and *T. vulgaris*, respectively), revealed that their unique N-terminal domains modify the active site, conferring distinct specificity [24]. While some researchers have reported that the N-terminal region forms an integral part of the active site [25, 26], it has also been speculated that it is too far from the active site to participate in enzymatic activity. In view of our results on substrate specificity, it is evident that the presence of the N-terminal sequence in Lka1p confers higher affinity for a broad range of substrates. The truncated Lka1p did not show any altered specificity towards the substrates, but there was a noticeable decline in its hydrolytic capacity. Furthermore, the deletion of Lka1p's SBD led to reduced affinity towards raw starch, pullulan and dextrin. Thus, the interaction of specific amino acids of the N-terminal region with the catalytic belt might affect the affinity of this enzyme for its substrates, enabling broad range specificity.

Lka2p possesses similar affinity towards various substrates, although it has an overall lower catalytic efficiency on raw starch. Despite the lack of an SBD, the *LKA2*-encoded enzyme degraded raw starch and, in this case, the hydrolytic capability of the enzyme was not directly correlated with the binding of the enzyme to the substrate, since this enzyme did not shown any adsorption or binding properties. Interestingly, Lka2p possesses higher affinity for dextrin and the deletion of the short N-terminal sequence preceding the catalytic barrel led to a loss of 50% of the original activity on dextrin. This suggests that this short region offers the molecule affinity for, and greater accessibility to, dextrin, leading to increased hydrolysis. Further, this deletion drastically reduced the K_{cat}/K_m for dextrin, clearly suggesting a possible role for amino acids in the N-terminal region of Lka2p in forming the active conformation and imparting stability of the enzyme-substrate complex. Improved conformational rigidity is indispensable for enhancing the stability of an enzyme against heat treatment [27].

Our data also suggest that the presence of the raw starch-binding motif leads to thermostability of the *LKA1*–encoded enzyme (Fig. 2). Our results are in agreement with previous reports that the SBDs in α-amylases might be associated with their thermal stability

[28, 29, 30, 31]. For an enzyme to be thermostable, it is necessary that there is a structural flexibility in each motif that comprises it [32]. The N-terminal domain in Lka1p might offer higher stability and structural flexibility between the domains of this molecule. On the other hand, Lka2pΔN did not show any difference in its thermal stability compared to the full-length enzyme. Consequently, the absence of an SBD-like sequence could explain the lower thermal stability of Lka2p compared to Lka1p.

Calcium contributes to the stability of some α-amylases by binding at specific sites on the enzyme, thereby reducing the flexibility of the protein and its susceptibility to partial unfolding [33]. The calcium-binding site in most amylases is located at the interface between the catalytic domain and the region preceding it [34]. The amino acids Asn121, Asp175, His210 and Glu162, which are implicated in the first calcium binding of TAKA amylase, are fully conserved in both Lka1p and Lka2p [35]. It was noted that the *L. kononenkoae* α-amylases did not need calcium for their activity and stability, as was reported for a *Cryptococcus* α-amylase [31]. The secondary calcium-binding site described for TAKA amylase, located at the bottom of the substrate-binding cleft involving residues Asp206 and Glu230 and presumed to play a catalytic role, is equally well preserved in both the *LKA1*- and *LKA2*-encoded α-amylases. This could explain the partial inhibitory effect of calcium observed at concentrations higher than 10 mM [35]. However, the reason for the apparent major inhibitory effect of calcium at high concentrations on the truncated enzymes, in contrast to their full-length wild types, is still unclear from our results (Fig. 4). It is possible that the lack of N terminus improved the access of Ca^{2+} to the active site and this might interfere/interact with the acidic groups in the amino acids.

The findings presented here will form the basis for mutational analyses to further our understanding of structure-function relationships, the nature and mechanism of the hydrolysis of substrates, and the evolutionary significance of α-amylases with unique specificity properties, such as Lka1p and Lka2p.

6. ACKNOWLEDGMENTS

This project was financially supported by South Africa's grapegrowers and winemakers through their investment body Winetech and by South Africa’s National Research Foundation.

7. REFERENCES

1. Lynd, L.R., Weimer, P.J., Van Zyl, W.H & Pretorius, I.S (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506-577.

2. Eksteen, J.M., Van Rensburg, P., Cordero Otero, R.R & Pretorius, I.S (2003) Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the α-amylase and glucoamylase genes from *Lipomyces kononenkoae* and *Saccharomycopsis fibuligera*. *Biotechnol. Bioen.* **84**, 639-646.

3. Steyn, A.J.C., Marmur, J. & Pretorius, I.S. (1995) Cloning, sequence analysis and expression in yeasts of a cDNA containing a *Lipomyces kononenkoae* α-amylase-encoding gene. *Gene* **166**, 65-71.

4. Horn, C.H., De Kock, A., Du Preez, J.C. & Lategan, P.M. (1988) A comparative study of the amylolytic ability of *Lipomyces* and *Schwanniomyces* yeast species. *System. Appl. Microbiol*. **10**, 106-110.

5. Spencer-Martins, I. & Van Uden, N. (1979) Extracellular amylolytic systems of the yeast *Lipomyces kononenkoae*. *Eur. J. Appl. Microbiol*. **6**, 241-250.

6. Spencer-Martins, I. (1982) Extracellular isoamylase produced by the yeast *Lipomyces kononenkoae*. *Appl. Environ. Microbiol*. **44**, 1253-1257.

7. Spencer-Martins, I. (1984) Production and properties of an extracellular cyclodextrin hydrolase from the yeast *Lipomyces kononenkoae*. *Int. J. Microbiol.* **2**, 31-38.

8. Hahn-Hägerdal, B., Wahlbom, C.F., Gárdonyi, M., Van Zyl, W.H., Cordero Otero, R.R. & Jönsson, L.J. (2001) Metabolic engineering of *Saccharomyces cerevisiae* for xylose utilisation. *Adv. Biochem. Eng. Biotechnol.* **73**, 53-84.

9. Steyn, A.J.C. & Pretorius, I.S. (1995) Characterization of a novel α-amylase from *Lipomyces kononenkoae* and expression of its gene (*LKA1*) in *Saccharomyces cerevisiae*. *Curr. Genet*. **28**, 526-533.

10. Eksteen, J.M., Steyn, A.J.C., Van Rensburg, P., Cordero Otero, R.R. & Pretorius, I.S. (2003) Cloning and characterization of a second α-amylase gene (*LKA2*) from *Lipomyces kononenkoae* IGC4052B and its expression in *Saccharomyces cerevisiae*. *Yeast* **20**, 69-78.

11. Janeček, S. & Sevcik, J. (1999) The evolution of starch binding domain. *FEBS Lett.* **456**, 119-125.

12. MacGregor, E., Janeček, S. & Svensson, B. (2001) Relationship of sequence and structure to specificity in the α-amylase family of enzymes, *Biochim. Biophys. Acta* **1546**, 1-20.

13. Matsuura, Y., Kunusoki, M., Harada, W. & Kakudo, M. (1984) Structure and possible catalytic residues of Taka amylase. *J. Biochem.* **95**, 695-702.

14. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309-316.

15. Takata, H., Kuriki, T., Okada, S., Takesada, Y., Iizuka, M., Minamiura, N. & Imanaka, T. (1992) Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at alpha-(1-4)- and alpha-(1-6)-glucosidic linkages. *J. Biol. Chem.* **267,** 18447-18452.

16. Svensson, B. (1994) Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity and stability. *Plant Mol. Biol*. **25**, 141-157.

17. Janeček, S. (1997) α-Amylase family: molecular biology and evolution. *Prog. Biophys. Molec. Biol.* **67**, 67-69.

18. Janeček, S., Svensson, B. & Macgregor, E.A. (2003) Relation between domain evolution, specificity and taxonomy of the α-amylase family members containing a C-terminal starch-binding domain. *Eur. J. Biochem.* **270**, 635-645.

19. Southall, S.M., Simpson, P.J., Gilbert, H.J., Williamson, G. & Williamson, M.P. (1999) The starch-binding domain from glucoamylase disrupts the structure of starch. *FEBS Lett.* **449**, 58-60.

20. Penninga, D., Van der Veen, B.A., Knegtel, R.M.A., Van Hijum, S.A.F.T., Rozeboom, H.J., Kalk, K.H., Dijkstra, B.W. & Dijkhuizen, L. (1996) The raw starch domain of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251. *J. Biol. Chem*. **271**, 32777-32784.

21. Ashikari, T., Nakamura, N., Tanaka, Y., Kiuchi, N., Shibono, Y., Tanaka, T., Amachi T. & Yoshizumi, H. (1986) *Rhizopus* raw starch degrading glucoamylase: its cloning and expression in yeast. *Agric. Biol. Chem*. **50**, 957-964.

22. Coutinho, P.M. & Reilly, P.J. (1997) Glucoamylase, structural, functional and evolutionary relationships. *Proteins: Structure, Function and Genetics* **29**, 334-347.

23. Bui, D.M., Kunze, I., Förster, S., Wartmann, T., Horstmann, C., Manteuffel, R. & Kunze, G. (1996) Cloning and expression of *Arxula adeninivorans* glucoamylase gene in *Saccharomyces cerevisiae*. *Appl. Microbiol. Biotechnol.* **44**, 610-619.

24. Park, K.H., Kim, T.J., Cheong, T.K., Kim, J.W., Oh, B.H. & Svensson, B. (2000)

Structure, specificity, and function of cyclomaltodextrinase, a multispecific enzyme of the α-amylase family. *Biochim. Biophys. Acta* **1478**, 165-185.

25. Katsuya, Y., Mezaki, Y., Kubota, M. & Matsura Y. (1998) Three-dimensional structure of *Pseudomonas* isoamylase at 2.2 Å resolution. *J. Mol. Biol.* **281**, 885-897.

26. Feese, M.D., Kato, Y., Tamda, Y., Kato, M., Komeda, T., Miura, Y., Hirose, M., Kondo, H., Kobayashi, E. & Kuroki, R. (2001) Crystal structure of glucosyl trehalohydrolase from the hyper thermophilic archaeum *Sulfolobus solfataricus*. *J. Mol. Biol.* **301**, 451-464.

27. Zavodsky, P., Kardos, J., Svingor, A. & Petsko, G.A. (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Nat. Acad. Sci. USA* **95**, 7406-7411.

28. Bahl, H., Burchhardt, G., Spreinat, A., Haeckel, K., Wienecke, A., Schmidt, B. & Antranikian, G. (1991) Alpha-amylase of *Clostridium thermosulfurogenes* EM1: nucleotide sequence of the gene, processing of the enzyme, and comparison of other alpha-amylases. *Appl. Environ. Microbiol*. **57**, 1554-1559.

29. Bahri, S.M. & Ward, M. (1990) Regulation of a thermostable alpha-amylase of *Streptomyces thermoviolaceus* CUB74: maltotriose is the smallest inducer. *Biochimie* **72**, 893-895.

30. Petricek, M., Tichy, P. & Kuncova, M. (1992) Characterization of the alpha-amylase-encoding gene from *Thermomonospora curvata*. *Gene* **112**, 77-83.

31. Iefuji, H., Chino, M., Kato, M. & Iimura, Y. (1996) Raw-starch-digesting and thermostable alpha-amylase from the yeast *Cryptococcus* sp. S-2: purification, characterization, cloning and sequencing. *Biochem. J*. **318**, 989-996.

32. Vihinen, M. & Mäntsälä, P. (1989) Microbial amylolytic enzymes. *Crit. Rev. Biochem. Mol. Biol.* **24**, 329-418.

33. Pantoliano, M.W., Whitlow, M., Wood, J.F., Rollence, M.L., Finzel, B.C., Gilliland, G.L., Poulos, T.L. & Bryan, P.N. (1988) The engineering of binding affinity at metal ion binding sites for the stabilization of proteins: subtilisin as a test case. *Biochemistry* **27**, 8311-8317.

34. Nielsen, J.E. & Borchert, T.V. (2000) Protein engineering of bacterial α-amylases. *Biochim. Biophys. Acta*. **1543**, 253-274.

35. Boel, E., Brady, L., Brozowski, A.M., Derewenda, Z., Dodson, G.G., Jensen, V.J., Swift, H., Thim, L. & Woldlike, H.F. (1990) Calcium binding in α-amylases: an X-ray diffraction study at 2.1 Å resolution of two enzymes from *Aspergillus, Biochemistry* **29**, 6244-6249.

36. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

37. Bollag, D.M. & Edelstein, S.J. (1991) *Protein Methods*, 1st edn. Wiley-Liss, New York.

38. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

39. Miller, G.L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem*. **31**, 426-428.

40. Fersht, A.R. (1984) *Enzyme structure and mechanism*, 2nd edn. Freeman, San Francisco, CA.

41. Gal, L., Pages, S., Gaudin, C., Belaich, A., Reverbel-Leroy, C., Tardif, C. & Belaich, J.P. (1997) Characterization of the cellulolytic complex (cellulosome) produced by *Clostridium cellulolyticum*. *App. Environ Microbiol*. **63**, 903-909.

42. Williamson, G., Belshaw, N.J. & Williamson, M.P. (1992) O-glycosylation in *Aspergillus* glucoamylase. *Biochem. J*. **282**, 423-428.

Table 1. Specific activity of wild-type (Lka1p and Lka2p) and truncated (Lka1p-ΔN and Lka2p-ΔN) *Lipomyces kononenkoae* α-amylases on different starch substrates.

Substrate	Reducing sugar released[a]			
	Lka1p	Lka1p-ΔN	Lka2p	Lka2p-ΔN
Soluble starch	4.8 ± 0.6	4.7 ± 0.3	4.6 ± 0.4	4.5 ± 0.6
Raw starch	4.1 ± 0.8	2.9 ± 0.6	3.9 ± 0.5	3.8 ± 0.5
Amylose	4.9 ± 0.6	4.1 ± 0.5	4.5 ± 0.6	4.4 ± 0.5
Amylopectin	3.8 ± 0.3	3.1 ± 0.7	3.5 ± 0.7	3.6 ± 0.8
Cyclodextrin	2.8 ± 1.1	2.1 ± 0.9	2.4 ± 0.9	2.4 ± 1.2
Dextrin	3.0 ± 0.4	2.4 ± 0.6	4.5 ± 0.3	2.5 ± 0.4
Pullulan	4.1 ± 0.7	3.7 ± 0.5	3.0 ± 0.6	2.9 ± 0.4
Maltotriose	1.4 ± 0.2	ND	ND	ND
Panose	ND[b]	ND	ND	ND

[a] $\mu mol \cdot min^{-1} \mu mol\ enzyme^{-1}$

[b] Not detected

Table 2. Kinetic properties of wild-type (Lka1p and Lka2p) and truncated (Lka1p-ΔN and Lka2p-ΔN) *Lipomyces kononenkoae* α-amylases.

Enzyme	Raw starch		Pullulan			Dextrin		
	V_{max} [a]	K_m [b]	K_{cat} [c]	K_m	K_{cat}/K_m [d]	K_{cat}	K_m	K_{cat}/K_m
LKA1	768.7 ± 39.0	1.34 ± 0.16	885.9 ± 92.3	2.10 ± 0.45	421.4	654.4 ± 46.7	1.78 ± 0.27	367.6
LKA1ΔN	745.8 ± 43.6	2.67 ± 0.33	930.1 ± 109.3	2.99 ± 0.63	311.0	626.8 ± 51.1	2.51 ± 0.39	249.7
LKA2	774.5 ± 38.4	1.83 ± 0.20	916.1 ± 73.0	3.06 ± 0.43	299.3	1008 ± 97.4	2.33 ± 0.44	432.6
LKA2ΔN	793.5 ± 34.1	1.75 ± 0.16	756.4 ± 46.7	2.45 ± 0.30	308.5	613.5 ± 51.7	2.44 ± 0.39	251.4

Each value is the mean ± standard deviations for three independent determinations.

[a] $mg\ mL^{-1}min^{-1}$

[b] mM

[c] s^{-1}

[d] $mg\ mL^{-1}\ s^{-1}$

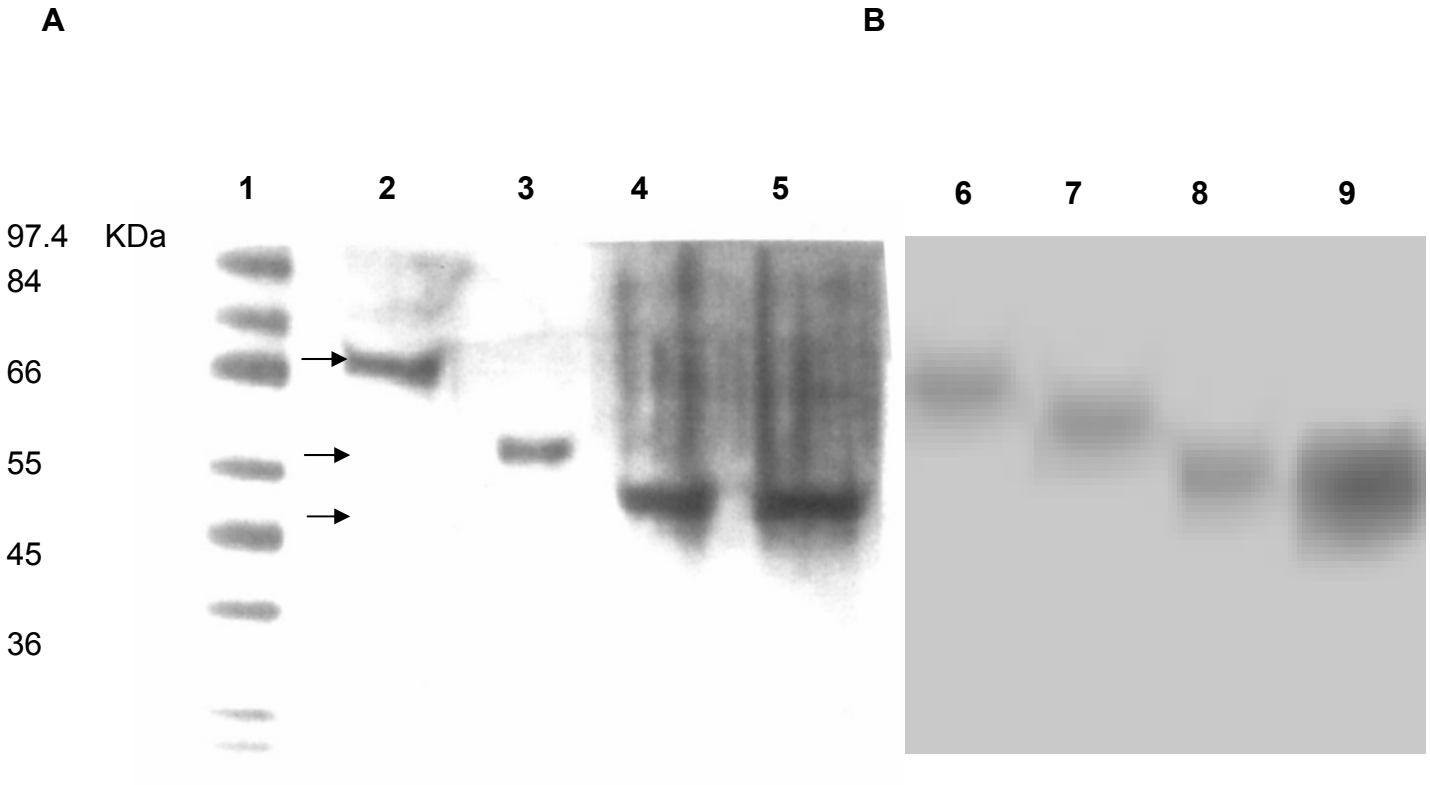
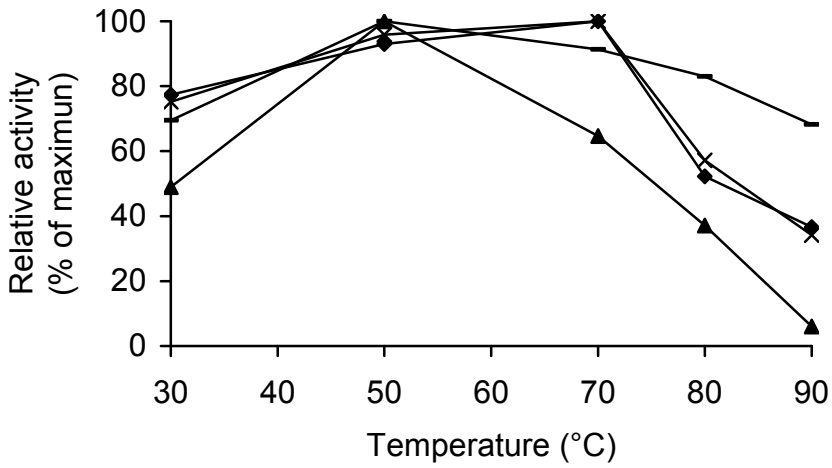


Fig. 1 A SDS-PAGE of purified amylolytic proteins. Lane 1: Molecular mass marker (SigmaMarker^{TM} high molecular weight range); lane 2: Lka1p; lane 3: Lka2p; lane 4: Lka1pΔN; lane 5: Lka2pΔN. The numbers to the right of the figure indicate the position and molecular weight of the marker in kDa. **B Western blot analysis**. lane 6: Lka1p; lane 7: Lka2p; lane 8: Lka1pΔN; lane 9: Lka2pΔN.

A



B

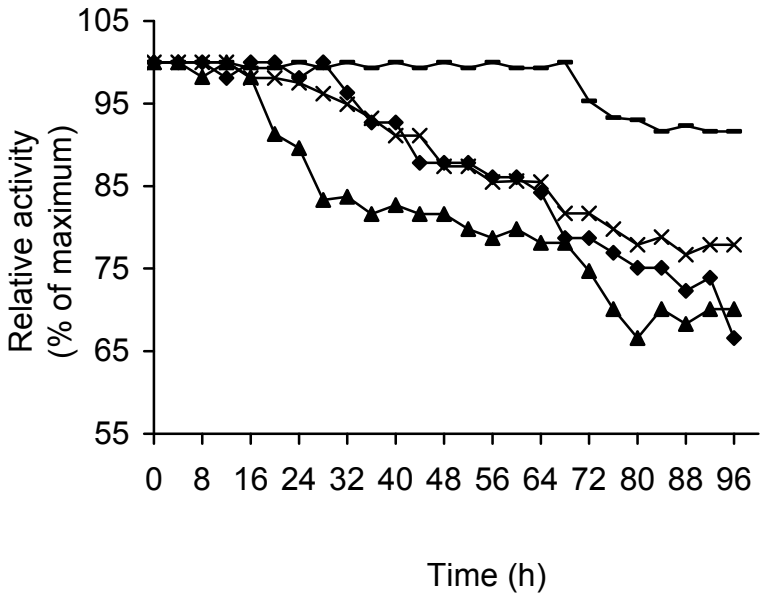


Fig. 2 The effect of temperature on the activity (A) and stability (B) of Lka1p (−), Lka2p (♦), Lka1pΔN (▲), and Lka2pΔN (×). The scale of relative activity (%) indicates the percentage of experimental values at various temperatures relative to the maximum value for each enzyme. The effect of the temperature was determined from 30 to 90°C in 20 mM sodium acetate buffer, pH 6.8, after 30 min reaction. The values shown here are means from assays done in triplicate ± 5% standard deviation.

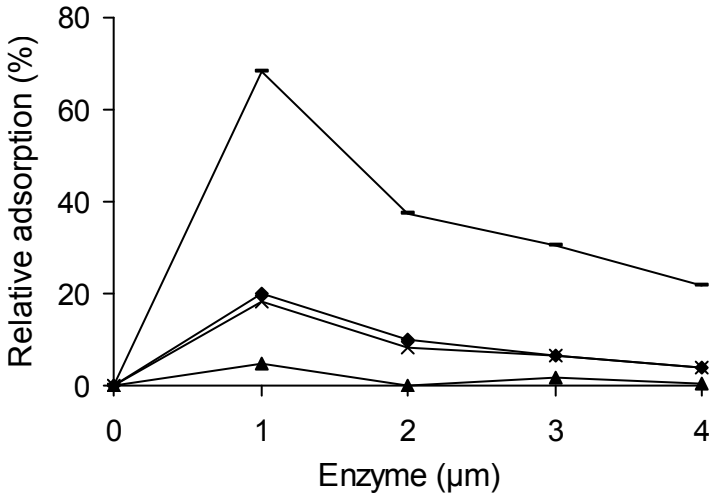


Fig. 3 Relative adsorption of enzymes on raw cornstarch. The values shown here are means from assays done in triplicate ± 5% standard deviation. Lka1p (−), Lka2p (♦), Lka1pΔN (▲), and Lka2pΔN (×).

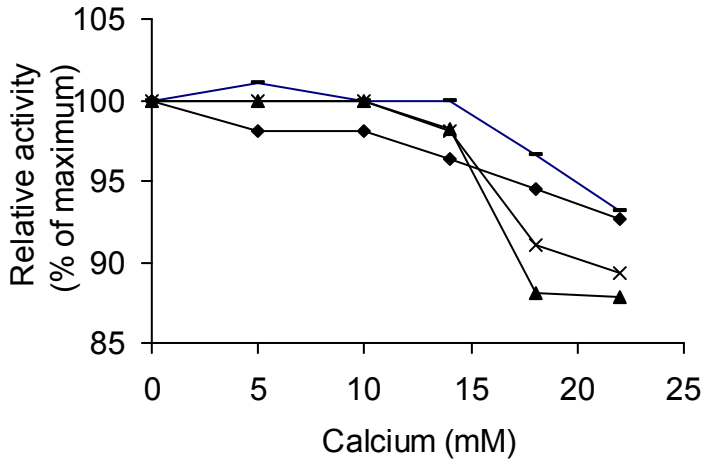


Fig. 4 The effect of Ca^{2+} concentration on the relative activities of amylases. The scale of relative activity (%) indicates the percentage of experimental values at various Ca^{2+} concentrations relative to the maximum value for each enzyme: Lka1p (−), Lka2p (♦), Lka1pΔN (▲), and Lka2pΔN (×). The values shown here are means from assays done in triplicate ± 3% standard deviation

```
Lka1p 41 ESVTGSNHVQ LASYEMCGST LSASLYVYND DYDKIVTLYY LTSSG...TT 86
Rhg   27 ASIPSSASVQ LDSYNYDGST FSGKIYVKNI AYSKKVTVIY ADGSDNWNNN 75

Lka1p 87 GSTLALIL.. PVWSNNWELW TLSAIAAG.. ....AVEITG ASYVDSDTSV 128
Rhg   76 GNTIAASYSA PISGSNYEYW TFSASINGIK EFYIKYEVSG KTYYDNNNSA 125

Lka1p 129 TYTTSLDLPL TTTSASVPTG 147
Rhg   126 NYQVSTSKPT TTTATATTTT 144
```

Fig. 5 Alignment of the amino acid sequences of N-terminal Starch-Binding Domains (SBDs). The sequences of the *Lipomyces kononenkoae* α-amylase Lka1p and *Rhizopus oryzae* glucoamylase Rhg have been aligned by introducing gaps (.) to maximise the similarity. Bold letters represent identical amino acids.

CHAPTER 4

RESEARCH RESULTS II

Site-directed mutagenesis of Gln^{414} at the active site of the *Lipomyces kononenkoae* α-amylase Lka1p

This manuscript will be submitted for publication in *European Journal of Biochemistry.*

Site-directed mutagenesis of Gln^{414} at the active site of the *Lipomyces kononenkoae* α-amylase Lka1p

Nivetha Ramachandran[1], Isak S. Pretorius[1,2], Ricardo R. Cordero Otero[1]

[1]Institute for Wine Biotechnology, Department of Viticulture & Oenology, Stellenbosch University, Victoria Street, Stellenbosch, ZA 7600, South Africa; *[2]The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, SA 5064, Australia*

1. ABSTRACT

The α-amylase family of enzymes is characterised by four highly conserved regions in the catalytic domain and most α-amylases consist of at least six fully conserved amino acids, which form the active site of these molecules. These amino acids include those of the catalytic triad (Asp^{206}, Asp^{297} and Glu^{230}) and three conserved histidines (His^{122}, His^{296} and His^{201}) (the a.a. numbering is according to TAKA amylase). Five of these amino acids are conserved in the *Lipomyces kononenkoae* α-amylase Lka1p (Asp$^{323, 415}$, Glu^{347} and His^{240,} His^{327}), which acts on α-1,4 and α-1,6 linkages in starch and limited activity on pullulan. However, Lka1p has a Gln^{414} at the matching place of His^{296}. In the light of His^{296} being important in transition state stabilisation at the subsite -1 in other α-amylases, Gln^{414} in Lka1p was changed to His, Ala, Asp, Asn or Glu to understand the significance of this amino acid in Lka1p function. Compared to the wild-type Lka1p, the mutant Gln^{414}His resulted in 1.5-fold increased hydrolysis of α-1,4 linkages, while there was a 1.2-fold decrease in activity on raw starch and a pronounced 33% decrease in activity on pullulan. The other mutants, Gln^{414}Ala/Asp/Glu/Asn, exhibited poor activity on all the substrates tested. The effects on specific activity, kinetic characterisation on linear maltooligosaccharide derivatives and differences in bond cleavage pattern of the wild-type Lka1p and the mutant enzymes provide evidence of substrate specificity properties of the α-amylase Lka1p.

2. INTRODUCTION

α-Amylases (α-1,4-D-glucanohydrolases, EC 3.2.1.1) act mainly on the internal α-1,4 linkages of starch and related substrates and belong to glucosyl hydrolase family 13, which constitutes

enzymes with more than 30 different specificities covering hydrolases, transferases and isomerases [1]. Numerous protein engineering studies targeting individual enzymes of this family have been done to explore the possibility of tailoring their properties to meet specific needs by broadening their substrate specificity [2-7]. However, the residues involved in catalysis and substrate binding are located in short stretches of four conserved regions in the catalytic domain and are preserved throughout the family. In the four conserved regions (^{117} DVVANH; ^{202} GLRIDTVKH; ^{230} EVLD; and ^{292} FVENHD, Asp^{206}, which functions as a catalytic nucleophile, Glu^{230}, the proton donor, Asp^{297}, the second conserved aspartate, and the two histidine residues, His^{122} and His^{296}, were postulated to be crucial for catalysis and transition-state stabilisation in TAKA amylase [8]. While the three catalytic residues (Asp^{206}, Glu^{230} and Asp^{297}) are strictly conserved throughout the family, His^{122} is not conserved in the amylomaltases (2.4.1.25), and His^{296} is absent in the maltosyl transferases (2.4.1.-). However, these histidine residues are fully conserved in α-amylases, cyclodextrin glucanotransferases, oligo 1.6 glucosidase, maltotetraohydrolase, isoamylases, neopullulanases and maltogenic α-amylases [9].

Primary structural analysis of other α-amylases reveals that the *Lipomyces kononenkoae* α-amylase Lka1p, which exhibits specificity towards α-1,4 and α-1,6 linkages in starch and specificity towards pullulan, also consists of the four conserved regions in the catalytic domain [10]. However, the second highly conserved histidine, His^{296}, is replaced by a glutamine (Gln^{414}). In this study, the role of the Gln^{414} of Lka1p was investigated, since His^{296} is understood to play a role in transition-state stabilisation during substrate binding and forms part of the sub site -1 [11]. To gain further insight into the role of Gln^{414} in specificity towards substrates, kinetic characterisation was performed and bond cleavage patterns were determined for wild-type Lka1p and mutant enzymes. The results of this study enhance our understanding of the mode of action of the *LKA1*-encoded α-amylase on different substrates and provide insights into the role of Gln^{414} in the hydrolysis of α-1,4 and α-1,6 linkages in polysaccharides.

3. MATERIALS AND METHODS

3.1 Strain and plasmids

Escherichia coli JM109DE3 strain (*end*A1, *rec*A1, *gyr*A96, *thi*, hsdR17 (r_k -, m_k+), *rel*A1, *supE44*, λ-, Δ(lac-proAB), [F'.*tra*D36, *pro*A+B+, lacI ^{q} ZΔM15]+) was used for the propagation of pALTER-*Ex*1 vector carrying the wild-type *LKA1* gene and the desired point mutations in the *LKA1* gene.

3.2 Site-directed mutagenesis

The wild-type *LKA1* gene encoding *Lipomyces kononenkoae* α-amylase without its native secretion signal was amplified using a forward primer (5'-CGGAATTCCGGAAGCTTA TGGATTGCACTACAGTTA-3') and a reverse primer (5'-GGGCGGCCGCCTACA TGGAGCAGATTC-3') to generate the PCR product of the *LKA1* gene with the flanking *Eco*RI and *Not*I sites (underlined). The PCR product was digested with *Eco*RI-*Not*I and cloned into pALTER-*Ex1* to generate the plasmid pALTER-*Ex*1Gln^{414}, which was used as a template for PCR mutagenesis. For the site-directed mutagenesis of *LKA1*, the mega-primer method [23] was followed, using the oligonucleotides in which a *Xho*I site (underlined) was engineered by introducing a point mutation: 5'- C TTTCTCGAGAATXXXGAT AATCCGCGATTC-3', where 'XXX' was CAT, CAG, GAT, GAG or AAC encoding His, Ala, Asp, Glu and Asn, respectively, instead of being Gln^{414}. The PCR reactions contained PCR buffer with 2 mM $MgSO_4$, 0.8 mM dNTPs, 0.3 μM of forward and reverse primers, 0.5 μM of mutagenic primer, 0.1 μg of template, and five units of *Taqara* DNA polymerase in a total volume of 50 μL. A thermocycler PCR Express gradient (Hybaid) was used for DNA amplification under the following conditions: 30 s at 94ºC, 30 s at 48ºC and 45 s at 68ºC. Twenty cycles were run, with a subsequent polymerisation period for 5 min at 68ºC. All of the cloned DNA fragments were verified by sequencing. PCR products were restricted with *Eco*RI-*Not*I and cloned into pALTER-*Ex*1. *E. coli* transformants harboring the recombinant plasmids were selected on Luria Bertani agar plates containing 125 μg mL^{-1} tetracycline.

3.3 Expression and purification of mutant enzymes

Wild-type Lka1p and the mutant enzymes were expressed by culturing *E. coli* JM109DE3 carrying the relevant plasmids at 37°C overnight in Luria Bertani medium with tetracycline (12.5 μg mL^{-1}). The culture supernatant was harvested by centrifuging the culture at 10 000 rpm for 2 min. A protocol based on anion exchange chromatography was devised using a BIO-RAD Model EP-1 Econo Pump (Bio-Rad, Hercules, CA, USA). DEAE Sepharose ™ fast flow matrix (Pharmacia Amersham, Uppsala, Sweden) was equilibrated with 50 mM MES buffer (pH 6.0). An equal proportion of supernatant and 50 mM MES buffer (1:1 ratio) was added to the column and mixed slowly for 2 h at 4°C, using an end-to-end Labinco rotary mixer. After being equilibrated with five column-volumes of 30 mM Tris-HCl at pH 8, the bound amylase was eluted using a linear 0.2-1 M NaCl gradient in 30 mM of Tris-HCl (pH 8) at a flow rate of 10 mL h^{-1}, and desalted with a Microcon® ultra-filtration device YM-3. Further purification and collection of proteins above 30 Kda was done by treating the eluent through YM-30 Microcon® centrifugal filter devices and the purified proteins were further analysed by means of SDS/PAGE.

3.4 SDS-PAGE and western blot

The molecular weight of the purified enzymes was determined by SDS-PAGE according to the manufacturer's recommendations (Hoeffer Scientific Instruments, San Francisco, USA) and visualised by silver staining [24]. Molecular masses of wild-type and mutant α-amylases were estimated according to Laemmli [25]. The resolved peptides were electrophoretically transferred onto an immobilon^{TM} –P transfer membrane (Millipore Corporation, Bedford, MA, USA) by semidry blotting (Towbin et al., 1979). The blotted proteins were identified immunochemically by the sequential addition of rabbit polyclonal anti-α-amylase (Abcam Ltd, Cambridgeshire, UK), followed by donkey anti-rabbit Ig-horseradish peroxidase-linked whole antibody (Amersham Pharmacia Ltd, UK). The secondary antibody was detected using the ECL Western blotting detection system (Amersham Pharmacia Ltd, UK).

3.5 Activity on high molecular weight substrates

Raw corn starch, soluble potato starch, amylopectin, pullulan and dextrin were purchased from Sigma, St. Luis, MO.

The activities of the wild-type Lka1p and the variant enzymes were assayed at 50°C for 30 min in a 20 mM sodium acetate buffer at pH 5.5. The activity was assayed by monitoring the release of reducing sugars using the dinitrosalicylic acid method [26]. One unit of activity of standard α-amylase was defined as the amount of enzyme that liberated 1 mg of reducing sugar per min. The amount of reducing sugar released was assessed from the maltose standard curve using 0.2% W/V maltose.

3.6 Kinetic characterization

Kinetic measurements were performed using 4-paranitrophenyl maltoheptaoside (G_7PNP; Boehringer Mannheim), 4-paranitrophenyl maltohexaoside (G_6PNP; Fluka), 4-paranitrophenyl maltopentaoside (G_5PNP; Sigma) and amylose (DP_{17}) in a coupled assay with 20 U mL^{-1} of yeast α-glucosidase (Boehringer Mannheim). The kinetic parameters were calculated using the initial velocities obtained from seven substrate concentrations of 0.05 to 5mM. A 3.0 nM concentration of wild-type Lka1p and 3.5 nM, 220 nM, and 0.5-1.2 μM of Gln^{414}His, Gln^{414}Ala and Gln^{414}Asp/Asn/Glu mutant enzymes were used, respectively. The assay was performed at 50°C in phosphate buffer at pH 5.5. All the experiments were carried out in triplicate and the average values were recorded.

3.7 pH and temperature profile and enzyme stability

The effect of pH on the activity of the wild-type Lka1p and the mutant enzymes was investigated over a pH range of 3 to 8 on 1% raw starch. Sodium phosphate buffer (20 mM) was used and the

assay was performed at 50°C. The activity profiles for the wild-type Lkalp and the mutant enzymes were measured over a temperature range of 25 to 80°C. For the assessment of stability, the enzymes were incubated at 50°C and aliquots were removed at different time intervals between 15 min and 12 h for activity assays. The release of reducing sugar was measured using dinitrosalicylic acid at A_{540} to calculate the activity of the enzymes.

3.8 Bond cleavage pattern on 4-nitrophenyl α-D maltooligosaccharides

The bond cleavage pattern was analysed using G_7PNP, G_6PNP and G_5PNP in 20 mM sodium phosphate buffer (pH 5.5) at 50°C. Hydrolysis was initiated by adding 4 nM of enzyme to 1 mM of substrate; 12 μL aliquots were removed at different time intervals and 3 μL of 10% acetic acid was added to stop the reaction. The products and residual substrates were separated using HPLC on a hypersil APS2 column at 40°C with 30:70 (v/v) acetonitrile and water as the solvent system at a flow rate of 1.0 mL min^{-1}. G_{1-7}PNP and 4-nitrophenol were detected at 275 nm and quantified against standard mixtures. The bond cleavage frequencies and relative activities were calculated as described by Matsui *et al*. (1997) [27].

4. RESULTS

4.1 Mutagenesis of Gln^{414} in LKA1

α-Amylases and other glycosyl hydrolases have four major conserved segments that are located in their catalytic region, and carry the amino acids that form the active site. On the basis of primary sequence analyses (Table 1) and secondary structure prediction, Gln^{414} of Lka1p clearly corresponds to the conserved residue His^{296}, which has been reported earlier in transition-state stabilisation in α-amylases and enzymes of the glucoside hydrolase family 13 [12, 13]. Using site-specific mutagenesis, Gln^{414} was mutated to His, Ala, Asp, Asn and Glu. The mutation of Gln^{414} to His was done to explore if His could functionally replace Gln^{414} in Lka1p. However, it was also targeted to understand the importance of Gln/His for substrate specificity of Lka1p towards α-1,4 and α-1,6 linkages. Gln^{414} was mutated to Asn to evaluate if it could act as a hydrogen bond donor, to Asp/Glu to understand the effect of substitution with negatively charged amino acids, and to the Ala mutant, which would be useful to assess the importance of hydrophilicity.

4.2 Purification and physical properties of the mutant enzymes

The modified proteins were purified through a combination of ion exchange and size-based column purification. The wild-type Lka1p and Gln^{414}Ala/Asp/His/Glu/Asn were recovered from single culture supernatants, although the mutant Gln^{414}Asn was produced in very low quantities.

A fine band of 68 kD was observed for the wild-type and mutant proteins (Fig.1) and all proteins had similar immunoreactivity towards the *Bacillus stearothermophilus* anti-α-amylase antibody (data not shown). The pH profiles of the wild-type and mutant enzymes are shown in Fig. 2A. The optimal pH of wild-type Lka1p was around 6.0, while the Gln^{414}Asn/Glu/His/Asp mutants also maintained their optimal activity at 6.0, retaining 90% of their original activity at pH 5.0 and 7.0. It was evident, however, that both the wild-type Lka1p and Gln^{414}His had narrow optimum pH ranges, exhibiting approximately 40% and 30% of activity at pH 5.0 and 7.0, respectively. In addition, Gln^{414}Ala had an altered pH optimum, with its highest activity at pH 4.0. It was observed that this enzyme retained only 30% of residual activity at pH 6.0. Wild-type Lka1p has a temperature optimum of around 50°C. To investigate whether the mutations affected the temperature optima, the temperature profiles of the mutant proteins were recorded between 25 and 80°C (Fig. 2B). The data obtained show that all the mutant enzymes exhibit optimal activity at 50ºC. At 70°C, the wild-type Lka1p and Gln^{414}His lost 40% of their original activity, while Gln^{414}Asn and Gln^{414}Asp/Glu/Ala exhibited a decrease of 70% and 90% in activity, respectively. To investigate whether the mutations affected the stability of Lka1p, residual activities were measured after incubation at 50°C for various lengths of time (Fig. 3). It was observed that both the wild type and the Gln^{414}His mutant protein remained stable, retaining 97% of their original activity for over 12 h, while the Gln^{414}Asp/Ala/Glu mutants showed decreases of 25, 25 and 60%, respectively, in residual activity after just 2 h and were clearly unstable after 3 h of incubation at 50°C. The enzyme Gln^{414}Asn, however, lost 20% of its original activity after 2 h, but retained 70% of its original activity for 12 h, showing greater stability than the other mutants at 50°C.

4.3 Enzyme activity assays

The activity of the Gln^{414} mutants and the wild-type Lka1p was compared on high molecular weight substrates including insoluble starch, soluble starch, pullulan, dextrin and amylopectin (Table 2). On insoluble starch, wild-type Lka1p had the highest activity than the mutant enzymes. In the context of His residue being common at this position in other α-amylase members (Table 1), Gln^{414}His showed 17% increased activity on soluble starch compared to the wild-type enzyme, but it also showed a 16% decrease in activity on insoluble starch, and a decrease of 11% and 10% on dextrin and amylopectin, respectively. It was observed that the Gln^{414}His mutant showed a 33% decrease in activity on pullulan, signifying that Gln^{414} might be one of the residues contributing to specificity towards pullulan. The Gln^{414}Ala/Asp/Asn/Glu mutants showed poor activity on all substrates tested, exhibiting 10% or less of the activity of the wild-type enzyme.

4.4 Bond cleavage pattern on 4-nitrophenyl α-D maltooligosaccharide derivatives

The Gln^{414} of Lka1p is located in the place of His^{296}, which, in TAKA amylase, forms interactions with the Tyr^{82} at the subsite -1 [14]. To analyse the role of Gln^{414} in substrate binding and processing, the cleavage propensity of specific bonds in derived maltooligosaccharides was evaluated using the reactions of wild-type Lka1p and mutant enzymes on G_7PNP, G_6PNP and G_5PNP (Table 4). The wild-type Lka1p and the Gln^{414}His/Asn/Glu/Asp mutants acted primarily on the fourth glucosidic bond of G_7PNP, releasing G_3PNP and G_4PNP as major products. It was also noticed that, apart from the major products (G_3PNP and G_4PNP), the Gln^{414}/Ala/Asp/Glu/Asn mutants produced small amounts of PNP, G_1PNP and G_5PNP. However, Gln^{414}Ala showed a different binding mode, cleaving the third glucosidic bond of the substrate and releasing G_2PNP and G_4PNP as major products. This mutant released 12% and 14% of PNP and G_1PNP, respectively, which was relatively higher than that released by its counterparts. On G_6PNP, the dominant binding mode of the wild-type and the Gln^{414}His/Asp/Glu/Asn mutant enzymes resulted in G_2PNP and G_3PNP as the principal products of hydrolysis. Again, the Gln^{414}Ala mutant exhibited a different binding mode, releasing G_3PNP as a major product. The amount of G_3PNP produced by this enzyme was two-fold higher than that of G_2PNP; the other enzymes released slightly higher amounts of G_2PNP and G_3PNP. For the shorter G_5PNP substrate, the action pattern was similar in all enzymes. The major cleavage products formed were G_1PNP and G_2PNP; the Gln^{414}His mutant, however, released 14% of PNP from this substrate. It was evident from the hydrolysis pattern that the Gln^{414}His mutant has higher hydrolytic efficiency than the wild-type enzyme on the linear maltooligosaccharides; it had 88%, 50% and 30% higher activity than the wild type on G_7PNP, G_6PNP and G_5PNP, respectively. The results reflect that, when Gln^{414} is replaced by His, this mutant exhibits increaed activity on linear α-,1,4-linked maltooligosaccharide substrates. However, there was a 1.2-fold decrease in the activity of the Gln^{414}His mutant on insoluble starch compared to the wild-type enzyme.

4.5 Kinetic characterization

The kinetics of the wild-type and mutant enzymes were evaluated on amylose DP_{17}, G_7PNP, G_6PNP and G_5PNP (Table 3). The wild-type enzyme hydrolysed linear amylose DP_{17} less efficiently than the mutant Gln^{414}His; the K_{cat}/K_m for this mutant was 66% higher than the wild type, essentially due to a lower K_m and higher K_{cat} values. For G_7PNP, K_{cat} was the highest for Gln^{414}His, showing a 3.3-fold increase over the wild type, and the lowest for Gln^{414}Ala, with a 57-fold decrease in activity compared to the wild-type enzyme. On this substrate, the mutated Gln^{414}Asp/Glu/Asn peptides showed poor activity, with decreased K_{cat} and higher K_m values. This pattern was similar on G_5PNP, with Gln^{414}His exhibiting a five-fold increase in catalytic efficiency over the wild-type enzyme. Very low activity was detected for other mutants on G_5PNP.

5. DISCUSSION

Like all other TIM barrel enzymes, α-amylases possess short conserved segments in the catalytic domain. They carry catalytic and substrate-binding residues at the C terminus of the β strands of the catalytic domain and in loops that extend from these strands. Sequence alignment of Lka1p with TAKA amylase shows corresponding similarities for the $(\beta/\alpha)_8$ catalytic domain. Four highly conserved regions containing the catalytically important residues, such as Glu^{230}, Asp^{206} and Asp^{297} [15, 16, 17], which were previously identified in the α-amylase family of enzymes [18], remain conserved in Lka1p. However, variations within the four conserved sequences have been reported in many members, which might be characteristic of enzyme specificities [17]. Region IV of Lka1p contains an uncharged, unique glutamine in place of the highly conserved and charged histidine, which is involved in substrate binding in other α-amylases. The importance of this invariant His^{296} has been well demonstrated in structural, kinetic and mutagenesis approaches in studies on other α-amylase members [19, 20, 21, 22]. Inhibitor-binding studies conducted on barley α-amylase also suggest that the of the imidazole from His residue forms a hydrogen bond with the substrate at subsite-1 at the bond cleavage site [12]. It has been stated that His^{296} is involved in transition-state stabilisation, but that it does not play a direct role in catalysis.

In this study, site-specific mutagenesis of Gln^{414} in Lka1p reveals a function of this residue in substrate processing, affinity and pH dependence. However, only the His residue could functionally replace the Gln^{414}, since all the other mutations (Gln^{414}Ala/Asn/Glu/Asp) led to very poor activity. Replacement of the Gln^{414} with amino acids other than His greatly diminished the ability to stabilise the transition state, as indicated by decreases in the K_{cat}/k_m values.

The three-dimensional structure of Lka1p is not available, but, by sequence comparison and this biochemical study, it can be postulated that Gln^{414} retains the geometry of the catalytic site, similar to the corresponding His in other enzymes. The presence of Gln^{414} in the wild type might thus sterically replace His, offering conformational flexibility, and also stabilise the aromatic stacking of Tyr^{82} at sub-site –1; a role played by His in other α-amylases [14]. The Gln^{414}His mutant exhibited higher activity on α-1,4 linkages and a lower activity towards complex substrates that possess α-1,4 and α-1,6 linkages. This was demonstrated on amylopectin, dextrin and pullulan, showing that the active site of the wild-type Lka1p facilitates increased hydrolysis of substrates with α-1,4 and α-1,6 linkages.

Furthermore, the higher activity of the Gln^{414}His mutant on linear maltooligosaccharides and substrates with only α-1,4 linkages might be due to the strengthening of the interaction with the bond to be cleaved at the non-reducing end of the substrate. An evaluation of the kinetics on

maltooligosaccharides such as G_7PNP, G_6PNP and G_5PNP further proved that Gln^{414}His has better interaction with the linear substrates and has the best productive mode of hydrolysis on α-1,4 glucosidic bonds, as seen in the lower K_m and higher K_{cat} values. The activities of the Gln^{414}His mutant and the wild-type enzyme was dependent on substrate length, which might be due to the difference in interaction towards different substrates. Noticeably, the relative activity of the Gln^{414}His mutant increased with increasing substrate length. Compared to the shorter substrates, the longer G_7PNP might possibly flank all sub-sites present in Lka1p, leading to increased hydrolysis. It can, however, also be suggested that the strong affinity at subsite -1, contributed by His, leads to the efficient hydrolysis of short substrates such as G_5PNP. Hence, the smaller the substrate, the larger the decrease in K_{cat}/k_m attributed by faster dissociation of the substrate from the active site, as the Gln^{414}Ala/Asn/Glu/Asp mutations did not facilitate substrate binding as efficiently as the Gln^{414}His mutant.

None of the mutations led to a complete loss of activity, suggesting that Gln^{414} does not play a direct role in catalysis, although it appears to be essential for the mechanism of action of the enzyme itself. However, the mutation with Ala, Glu and Asn led to reduced activity and it is thus clear that these hydrogen bond donors do not fully compensate for Gln in this position. The poor activity of these mutants is possibly due to the lack of sufficient binding or even a non-productive mode of binding, which is clearly observed in Gln^{414}Ala. This mutant binds G_6PNP and G_7PNP differently from the wild type, and the relative activities on these substrates were less than that of the wild-type enzyme and the Gln^{414}His mutant. It might be suggested that there was a loss of stable complex formation during substrate-enzyme interactions. In addition, the Gln^{414}Ala mutant showed a shift in its pH maximum to pH 4.0 compared to the wild-type enzyme (pH 5.0), which indicates that this mutation might have affected the activity of the enzyme at optimal pH, as reflected by its overall poor activity. The relative activities for the Gln^{414}Asp/Glu/Asn mutants were also lower than those of the wild-type enzyme, but there was no shift in bond cleavage pattern and, in turn, the Gln^{414}Asp/Glu mutants were generally unstable.

Thus, the mutations reveal that His can sterically complement Gln^{414} in Lka1p, but that it leads to reduced activity on both raw starch and on pullulan. Since the Gln^{414}His mutation also results in increased activity on amylose and other linear-chain polymers, it is suggested that the presence of Gln^{414} in Lka1p and His in other α-amylases at this equivalent position plays a major role in substrate affinity and processing.

Since substrate preference engineering could be used to extend the substrate range of Lka1p, this data makes a significant contribution to understanding which amino acids play a role in the dual bond specificity of the enzyme. The outcome of the current study is thus promising to further our understanding of the substrate specificity and mode of action of the α-amylase Lka1p.

6. ACKNOWLEDGEMENTS

This project was supported financially by South Africa's grape growers and winemakers through their investment body, Winetech, and by South Africa's National Research Foundation.

7. REFERENCES

1. Svensson, B., Jensen, T.M., Mori, H., Bak-Jensen, S.K., Bonsager, B., Nielsen, P.K., Kramhoft, B., Pretorius-Ibba, M., Nohr, J., Juge, N., Greffe, L., Williamson, G & Driguez, H. (2002). Fascinating facets of function and structure of amylolytic enzymes of glycoside hydrolase family 13. *Biologia, Bratislava*. **57**, 5-19.

2. Kim, J.S., Cha, S.S., Kim, H.J., Kim, T.J., Ha, N.C., Oh, S.T., Cho, H.S., Cho, M.J., Kim, M.J., Lee, H.S., Kim, J.W., Choi, K.Y., Park, K.H. & Oh, B.H. (1999) Crystal structure of a maltogenic amylase provides insights into a catalytic versatility. *J. Biol. Chem* **274**, 26279-26286.

3. Yamamoto, T., Fujiwara, S., Tachibaha, Y., Takagi, M., Fukui, K., Imanaka, T. (2000) Alteration of product specificity of cyclodextrin glucanotransferase from *Thermococcus* sp. B1001 by site-directed mutagenesis. *J. Biosci. Bioeng.* **89**, 206-209.

4. Kuriki, T., Kanenko, H., Yanase, M., Takata, H., Shimada, J., Handa, S., Tajada, T., Umeyama, H & Okada, S. (1996) Controlling substrate preference and transglycosylation activity of neopullulanase by manipulating steric constraint and hydrophobicity in active center. *J. Biol. Chem.* **271**, 17321-17329.

5. Mori, H., Bak-Jensen, K.S., Gottschalk, T.E., Motawia, M.S., Damager, I., Moller, B.L., Svensson B. (2001) Modulation of activity and substrate binding modes by mutation of single and double subsites +1/+2 and -5/-6 of barley alpha-amylase1. *Eur. J. Biochem*. **268**, 6545-6558.

6. Gottschalk, T.E., Tull, D., Aghajari, N., Haser, R. & Svensson, B. (2001) Specificity modulation of barley α-amylase through biased random mutagenesis involving a conserved tripeptide in β-α loop 7 of the catalytic $(\beta/\alpha)_8$-barrel domain. *Biochemistry* **40**, 12844-12854.

7. Leemhuis, H., Dijkstra, B,W., Dijkhuizen, L. (2002) Mutations converting cyclodextrin glycosyltransferase from a transglycosylase into a starch hydrolase. *FEBS Lett.* **514**, 189-192.

8. Matsuura, Y., Kusunoki, W., Harada, W., Kakudo, M. (1984) Structure and possible catalytic residues of Taka amylase A. *J . Biochem*. **95**, 697-702.

9. Janecek, S (2002) How many conserved sequence regions are there in the α-amylase family? *Biologia, Bratislava*. **57**, 29-41.

10. Steyn, A.J.C. & Pretorius, I.S. (1995) Characterization of a novel α-amylase from *Lipomyces kononenkoae* and expression of its gene (*LKA1*) in *Saccharomyces cerevisiae*. *Curr. Genet*. **28**, 526-533.

11. Brzozowski, A.M., Davies, G.J. (1997). Structure of *Aspergillus oryzae* α-amylase complexed with inhibitor acarbose at 2.0 Å resolution. *Biochemistry.* **36**, 10837-10845.

12. Søgaard, M., Kadziola, A., Haser, R. & Svensson, B. (1993) Site-directed mutagenesis of histidine 93, aspartic acid 180, glutamic acid 205, histidine 290, and aspartic acid 291 at the active site and tryptophan 279 at the raw starch binding site in barley α-amylase 1. *J. Biol. Chem.* **268**, 22480-22484.

13. Qian, M., Haser, R., Buisson, G., Duée, E. & Payan, F. (1994) The active center of a mammalian α-amylase. Structure of the complex of a pancreatic α-amylase with a carbohydrate inhibitor refined to 2.2-Å resolution. *Biochemistry* **33**, 6284-6294.

14. Strokopytov, B., Penninga, D., Rozeboom, H.J., Kalk, K.H., Dijkhuizen, L., Dijkstra, B.W. (1995) X-ray structure of cyclodextrin glycosyltransferase complexed with acarbose. Implications for the catalytic mechanism of glycosidases. *Biochemistry* **34**, 2234-2240.

15. Henrissat, B. (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* **280**, 309-316.

16. Takata, H., Kuriki, T., Okada, S., Takesada, Y., Iizuka, M., Minamiura, N. & Imanaka T. (1992) Action of neopullulanase. Neopullulanase catalyzes both hydrolysis and transglycosylation at α-(1-4)- and α-(1-6)-glucosidic linkages. *J. Biol. Chem.* **267,** 18447-

18452.

17. Svensson, B. (1994) Protein engineering in the α-amylase family: catalytic mechanism, substrate specificity and stability. *Plant Mol. Biol*. **25**, 141-157

18. Janeček, S. (1997) α-Amylase family: molecular biology and evolution. *Prog. Biophys. Molec. Biol.* **67**, 67-69.

19. Kuriki, T. Takata, H., Okada, S. & Imanaka, T. (1991) Analysis of the active center of *Bacillus stearothermophilus* neopullulanase. *J. Bacteriol.* **173**, 6147–6152.

20. Ishikawa, K., Matsui, I., Kobayashi, S., Nakatani, H. & Honda, K. (1993) Substrate recognition at the binding site in mammalian pancreaticalph α-amylases. *Biochemistry*. **32**, 6259-6265.

21. Matsui, I., Ishikawa, K., Miyairi, S., Fukui, S. & Honda, K. (1992) A mutant alpha-amylase with enhanced activity specific for short substrates. *FEBS Lett*. **310**, 216-218.

22. Takase, K. (1993) Effect of mutation of an amino acid residue near the catalytic site on the activity of *Bacillus stearothermophilus* alpha-amylase. *Eur. J. Biochem*. **211**, 899-902.

23. Datta, A.K. (1995) Efficient amplification using 'megaprimer' by asymmetric polymerase chain reaction. *Nucl. Acids Res.* **23**, 4530-4531.

24. Bollag, D.M. & Edelstein, S.J. (1991) *Protein Methods*, 1st edn. Wiley-Liss, New York.

25. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.

26. Miller, G.L. (1959) Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem*. **31**, 426-428.

27. Matsui, I. & Svensson, B. (1997) Improved activity and modulated action pattern obtained by random mutagenesis at the fourth loop involved in substrate binding to the catalytic $(\beta/\alpha)_8$-barrel domain of barley α-amylase 1. *J. Biol. Chem.* **272**, 22456-22463.

28. Towbin, H., Staehelin, T. & Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA 76: 4350-4354

Table 1. Sequence comparison of different enzymes of α-amylase family: The presence of Gln^{414} in Lka1p in the place of the highly conserved His^{296} of the fourth conserved region

Enzymes	Region 1	Region II	Region III	Region IV
α-Amylase	117 DVVAN**H**	202 GLRIDTVKH	230 EVLD	292 FVEN**H**D
Pullulanase	600 DVVYN**H**	671 GFRFDLMGY	704 EGWD	827 YVSK**H**D
Neopullulanase	242 DAVFN**H**	324 GWRLDVANE	357 EIWH	419 LLGS**H**D
α-Amylase pullulanase	488 DGVFN**H**	594 GWRLDVANE	627 ENWN	699 LLGS**H**D
CGTase	135 DFAPN**H**	225 GIRFDAVKH	258 EWFL	324 FIDN**H**D
Isoamylase	292 DVVYN**H**	371 GFRFDLASV	435 EPWA	505 FIDV**H**D
LKA1 α-amylase	235 DIVVN**H**	319 GLRIDTVKH	347 EVFD	410 FLEN**Q**D

Table 2. Activity of wild type Lka1p and mutant enzymes on high molecular weight substrates

Enzyme	Activity (%)				
	Raw starch	Soluble starch	Dextrin	Pullulan	Amylopectin
Gln^{414}	100	100	100	100	100
Gln^{414}His	84	107	98	77	90
Gln^{414}Ala	10	7	10	ND[a]	7
Gln^{414}Asp	11	7	10	ND	7
Gln^{414}Glu	10	7	11	ND	8
Gln^{414}Asn	10	7	11	ND	7

[a] Not detectable

Table 3. Activity and kinetic parameters of Gln^{414} Lka1p mutants and wild type towards raw corn starch, amylose and linear maltooligosaccharide derivatives

Concentration of soluble and raw starch used was 10 mg mL^{-1}; Standard deviation was less than 6%

Enzyme	Activity on raw corn starch (U mg^{-1})	Activity on soluble starch (U mg^{-1})	Amylose DP_{17}			G_7PNP			G_5PNP		
			K_{cat} [a]	K_m [b]	K_{eff} [c]	K_{cat}	K_m	K_{eff}	K_{cat}	K_m	K_{eff}
Gln^{414}	4575	7029	841.0	1.2	700.8	188.8	0.6	309.5	157.2	0.8	201.5
Gln^{414}His	3860	7523	1046.0	1.0	1046.0	628.4	0.4	1571.0	597.0	0.5	1148.1
Gln^{414}Ala	459	488	211.0	2.4	87.9	11.7	1.0	11.6	ND[d]	ND	ND
Gln^{414}Asp	488	459	71.0	2.5	28.4	21.2	1.0	20.8	12.0	1.2	10.0
Gln^{414}Glu	459	488	36.0	2.1	17.1	22.7	0.9	25.2	ND	ND	ND
Gln^{414}Asn	459	517	41.0	1.9	21.6	17.8	0.8	22.3	ND	ND	ND

[a] s^{-1}

[b] mM

[c] $s^{-1}.mM^{-1}$

[d] Not detectable

4. Hydrolytic pattern of Gln^{414} Lka1p mutant and wild type enzymes on G_7PNP, G_6PNP and G_5PNP

Lka1p	Cleavage frequency (%)	[E] (nm)	Time (min)	Degree of cleavage (%)	Relative activity (%)
G_7PNP	G - G - G - G - G - G - G - PNP				
Wild type	1 22 65 12	3	2	9.0	100.0
Gln^{414}His	2 20 67 11	3.5	1	10.0	188.0
Gln^{414}Ala	1 22 10 43 14 12	55	10	12.0	1.4
Gln^{414}Asp	1 25 62 9 2 1	57	10	15.0	1.8
Gln^{414}Glu	3 24 61 9 1 2	57	10	16.0	1.9
Gln^{414}Asn	2 21 63 10 2 2	57	10	15.0	1.8
G_6PNP	G - G - G - G - G - G - PNP				
Wild type	3 5 41 49 2	65	4	11.5	100.0
Gln^{414}His	4 6 40 50	60	3	12.0	150.0
Gln^{414}Ala	2 65 32 1	125	30	18.6	9.0
Gln^{414}Asp	5 7 39 47 2	125	30	19.2	11.6
Gln^{414}Glu	3 5 42 48 2	125	30	18.7	11.2
Gln^{414}Asn	2 6 40 49 3	125	30	19.2	11.5
G_5PNP	G - G - G - G - G - PNP				
Wild type	5 14 30 49	120	5	9.3	100.0
Gln^{414}His	4 7 22 53 14	120	4	9.7	130.0
Gln^{414}Ala	15 12 34 39	575	60	21.2	4.0
Gln^{414}Asp	16 14 33 37	575	70	23.0	3.7
Gln^{414}Glu	17 16 30 38	575	90	25.0	3.1
Gln^{414}Asn	17 13 32 38	575	60	22.1	4.1

The relative activity was estimated for individual enzymes from the degree of substrates cleaved, enzyme concentration [E] and time (min) in relation to the wild type enzyme.

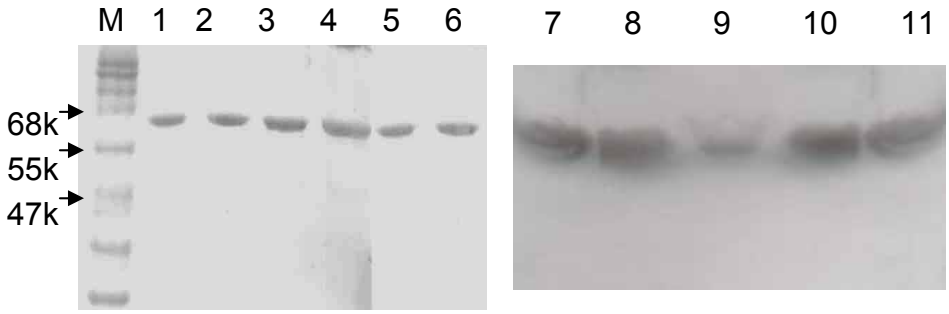
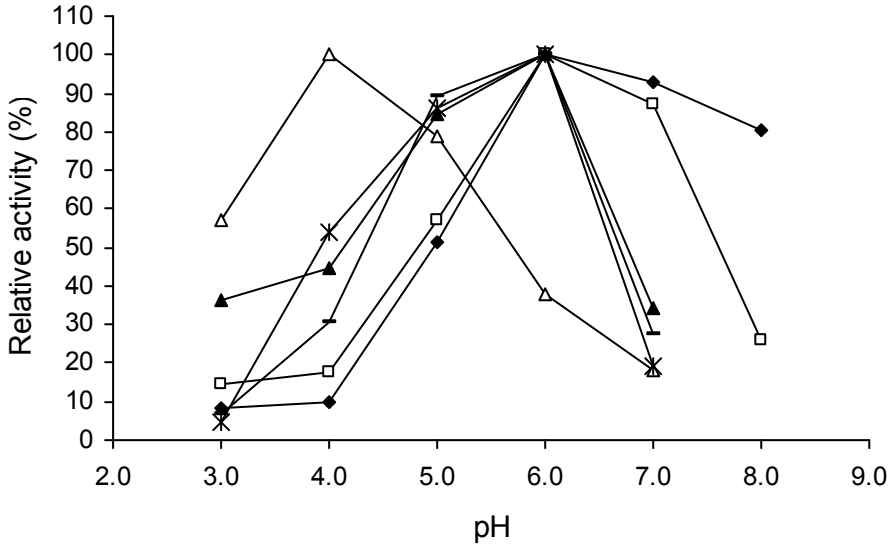


Fig 1. Poly-acrylamide gel electrophoresis (PAGE) of purified wild-type and mutant enzymes. Lane M, sigma wide range molecular weight marker; lane 1, wild type; lane 2, $Gln^{414}His$; lane 3, $Gln^{414}Ala$; lane 4, $Gln^{414}Asp$; lane 5, $Gln^{414}Glu$; lane 6, $Gln^{414}Asn$ **and Western blot analysis.** Lane 7, wild type; lane 8, $Gln^{414}His$; lane 9, $Gln^{414}Ala$; lane 10, $Gln^{414}Asp$; lane 11, $Gln^{414}Glu$

A



B

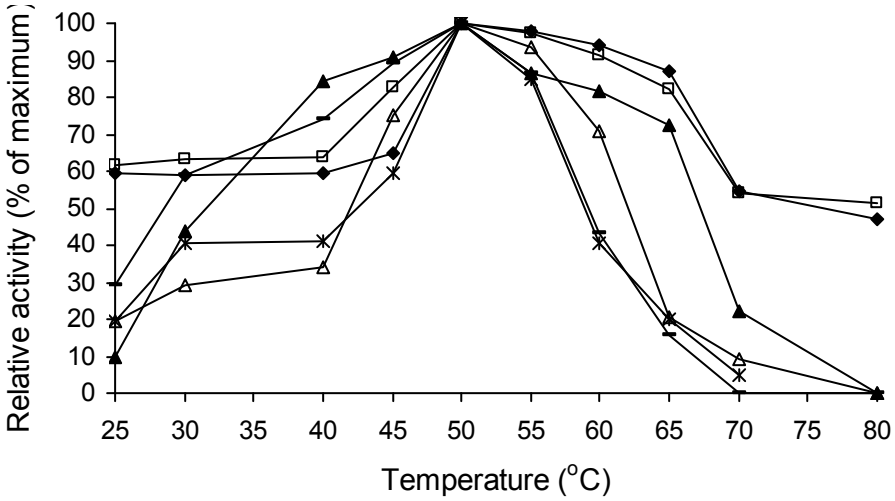


Fig 2. (A) pH profile of wild type Lka1p and mutant enzymes. **(B) Temperature profile of wild type Lka1p and mutant proteins.** The scale of relative activity (%) indicates the percentage of experimental value at various pH and temperatures relative to the maximum value of each enzyme: (◆) Gln^{414}; (□) Gln^{414}His; (△) Gln^{414}Ala; (–) Gln^{414}Asp; (∗) Gln^{414}Glu; (▲) Gln^{414}Asn.

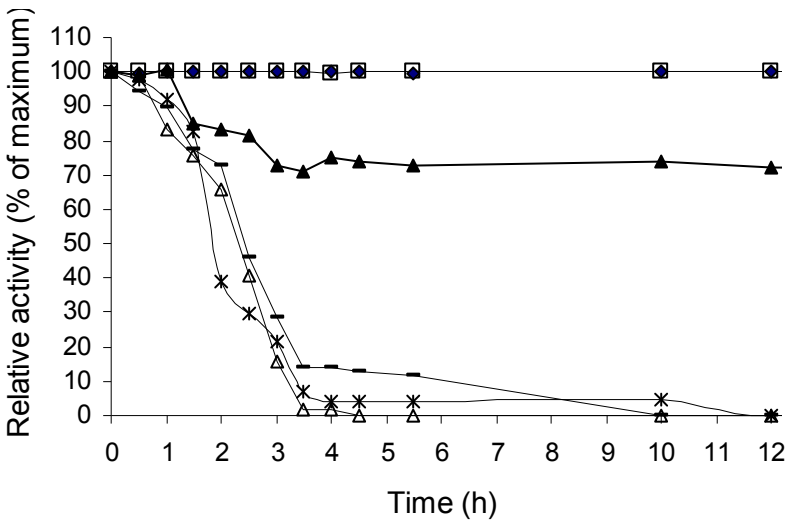


Fig 3. Stability of wild type Lka1p and mutant proteins at 50° C at different time periods; relative specific activity (%) is plotted as a function of time. (◆) Gln^{414}; (□) $Gln^{414}His$; (△) $Gln^{414}Ala$; (–) $Gln^{414}Asp$; (∗) $Gln^{414}Glu$; (▲) $Gln^{414}Asn$.

CHAPTER 5

RESEARCH RESULTS III

Effect of flocculation on *Saccharomyces cerevisiae* expressing *Lipomyces kononenkoae* α-amylase Lka1p for raw starch fermentations

This manuscript will be submitted for publication in *Applied and Environmental Microbiology.*

Effect of flocculation on *Saccharomyces cerevisiae* expressing *Lipomyces kononenkoae* α-amylase Lka1p for raw starch fermentations

N. Ramachandran[1], L. Joubert [3], I. S. Pretorius[1, 2], R. R. Cordero Otero[1*]

[1]Institute for Wine Biotechnology, Department of Viticulture & Oenology, Stellenbosch University, Victoria Street, Stellenbosch, ZA 7602, South Africa; [2]The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, Australia.
[3] Department of Microbiology, Stellenbosch University, Stellenbosch-7602

1. ABSTRACT

To develop novel yeasts capable of bioconversion of starch to ethanol, laboratory *Saccharomyces cerevisiae* strains expressing the *LKA1*-encoded raw starch-degrading α-amylase (Lka1p) from *Lipomyces kononenkoae* in flocculent and non-flocculent genetic backgrounds were constructed. The *LKAI* gene was maintained both episomally and also by integration into the chromosome. Expression of *LKA1* in *S. cerevisiae* was directed by the *PGK1* promoter and terminator sequences, and secretion of Lka1p was directed by the *MFα1*$_S$-signal sequences. Both flocculent and non-flocculent recombinant strains were able to grow on raw starch, although the flocculent strains degraded starch at an earlier time than the non-flocculent strains, despite similar levels of α-amylase in the extracellular culture media. The interaction between flocculent cells and starch granules were investigated visually using a flow-cell system and light microscopy, as well as scanning and atomic force microscopy. The fermentation performance of the strains in raw corn starch, the effect of permeation on whole cell catalysts and the effect of flocculation are discussed. Small-scale batch fermentations were performed. The non-flocculent strain (BELK1) consumed 76% of the starch supplied in the media and produced 4.61 gL^{-1} of ethanol after 90 hrs of fermentation, while the flocculent strain (BELK1F) utilised 82% of the raw starch, resulting in 5.1gL^{-1} of ethanol.

2. INTRODUCTION

The bioprocessing of starch-rich materials for the production of alternative energy sources has received considerable attention in recent years. Impediments to the effective

bioconversion of starchy materials to valuable products include the identification of potential organisms and biosynthetic pathways needed to produce valuable intermediates and byproducts, and the metabolic optimisation of these organisms for high yield of the desired products. This has stimulated research into amylolytic microorganisms and their enzymes and the broadening of substrate range for *Saccharomyces cerevisiae* (24). However, not all naturally-occurring biocatalysts in the form of organisms and enzymes are optimally suited for industrial applications.

Saccharomyces cerevisiae has been the most exploited microorganism for industrial fermentations, but it lacks the amylolytic system to degrade starch. There are several reports on the development of recombinant *Saccharomyces* strains capable of secreting heterologous amylolytic enzymes (2, 3, 7, 9, 17, 18, 21). Although a wide variety of amylases have been successfully secreted in *S. cerevisiae*, their starch decomposition abilities are poor due to limited availability of the enzyme in the extracellular medium. In an attempt to exploit the secreted enzymes efficiently for the production of ethanol, various strategies have been applied in starch fermentations, e.g. cell surface display of amylolytic enzymes on flocculent yeast cells, mixed culture fermentations, and the use of immobilized whole cell biocatalysts. Among various approaches used, flocculation is an interesting phenomenon for ethanol production because the time of occurrence and intensity of flocculation can affect the degree of fermentation, and therefore, the alcohol level (13). Flocculation is also one of the important characteristics of yeast strains used in industrial applications like brewing and production of recombinant proteins, because of the ease of separation of flocculent yeast cells during downstream processing (9). The *LKA1*-encoded α-amylase (Lka1p) from *Lipomyces kononenkoae* acts on the α-1, 4 and α-1, 6 linkages of starch, facilitating the break down of complex raw corn starch. We have previously reported the use of the α-amylase Lka1p in conjunction with a second α-amylase, Lka2p from the yeast

L. kononenkoae α-amylases were expressed under *PGK1* promoter-terminator *vs.* native promoters and directed by their native secretion signals for the fermentation of soluble starch to ethanol (3) The current hypothesis is that a flocculent phenotype might decrease the time interval for interaction between an α-amylase and its substrate, thus facilitating substrate degradation with a shorter time lag. To understand the effect of flocculation in this context, flocculent and non-flocculent *S. cerevisiae* strains expressing the *L. kononenkoae* raw starch-binding α-amylase (Lka1p) were constructed. The effects of flocculent phenotypes on bioconversion of starch to ethanol were subsequently examined in small-scale batch fermentations.

3. MATERIALS AND METHODS

3.1 Strains and plasmids

Escherichia coli strain DH5α (F- φ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *deo*R *rec*A1 *end*A1 *hsd*R17 (rK^-, mK^+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1) was used as an intermediate host for the cloning and propagation of all the plasmids. The yeast expression vector pSTA1 (5.99 kb), a derivative of YIp5, was used as the integration plasmid. The 2μ plasmid, YEp*lac*112 (4.9 kb), with the *TRP5* auxotropic marker, was used as the multicopy expression vector. Both forms of plasmid possess the *PGK1* sequence with the $MF\alpha1_S$ secretion signal, as well as the ampicillin resistance gene and the sequences required for replication in both *E. coli* and *S. cerevisiae*. Plasmid pIP*LKA1* was the source of the *LKA1* gene from *L. kononenkoae* (3). *S. cerevisiae* non-flocculent strain FY834 (*MATα his3 leu2 lys2 trp1 ura3 flo8*) and the flocculent strain FY835 (*MATα his3 leu2 FLO8 lys2 trp1 ura3*) were used.

3.2 Media and growth conditions

E. coli DH5α was grown at 37°C in Luria Bertani medium containing 100 μg mL^{-1} ampicillin. The corn starch medium used for the observation of amylolytic activity consisted of 20 g L^{-1} corn starch (Sigma), 3 g L^{-1} starch azure, 5 g L^{-1} peptone buffered with citrate buffer (pH 5.5) and 20 g L^{-1} agar. The medium used to evaluate invasive growth contained 20 g L^{-1} glucose as sole carbon source with 50 μM ammonium sulphate as sole nitrogen source, and 1.7 g L^{-1} yeast nitrogen base without ammonium sulphate and amino acids (Difco). The growth characteristics of the amylolytic strains were examined using YPRS medium containing 10 g L^{-1} yeast extract, 2 g L^{-1} peptone and 10 g L^{-1} raw corn starch. The growth pattern was monitored by sampling at 4 h intervals. For growth under selection, minimal medium (YNB without amino acids) was used and supplied with the corresponding amino acid requirement. The medium used for fermentation consisted of 20 g L^{-1} raw corn starch, 6.7 g L^{-1} yeast nitrogen base, 0.4 g L^{-1} Tween-80 and 0.01 g L^{-1} ergosterol (the stock solution was made up with ethanol, contributing to 11 mM ethanol in the medium).

3.3 Construction of recombinant amylolytic plasmids and strains

The *LKA1* gene was amplified from plasmid pIP*LKA1* with F-*LKA1* (5'-CGGAATTCCGGAAGCTTATGGATTGCACTACAGTTA-3') and R-*LKA1* (5'-GGCTCGAGCTACATGGAGCAGATTC-3') primers and cloned into the *Hin*dIII-*Xho*I sites of pSTA1 for the construction of integration plasmid pSTA2 (7.9 kb). Subsequently, the $PGK1_P$-$MF\alpha1_S$-*LKA1*-$PGK1_T$ cassette from pSTA2 was amplified using F-*PGK*634 (5'-GCGACGTCCTTTATTTTGGCTTCACCC-3') and R-*PGK* B (5'-GCGGATCCGATAAATAATAGTCTATATATACG-3') and cloned into the 2μ plasmid YEp*lac*112 to construct the multicopy expression vector YEp*LK*112. Standard methods were

used for the isolation, restriction, purification and ligation of DNA, the plasmid transformation into *E. coli*, and agarose gel electrophoresis. The final constructs were verified by sequence analysis using the ABI PRISM® Big Dye™ Terminator cycle sequencing kit with an ABI PRISM™ 377 DNA sequencer (PE/Applied Biosystems). Both the coding and non-coding strands were sequenced to ensure the reliable identification of all constructs. *S. cerevisiae* non-flocculent strain FY834 (*MATα his3 leu2 lys2 trp1 ura3 flo8*) with a non-reverting mutation in *flo8* was transformed with a functional copy of *FLO8* to restore the flocculent phenotype (FY835). *Saccharomyces cerevisiae* strains FY834 (non-flocculent) and FY835 (flocculent) were integrated with the recombinant plasmid pSTA2 to create the amylolytic recombinant strains BELK1 and BELK1F respectively. The integration plasmids were linearised with the unique *Apa*I site within the *URA*3 sequence to ensure targeting of the *URA3* locus in the *Saccharomyces cerevisiae* genome. Transformants were selected on SC medium without uracil. The transformation of FY834 and FY835 with the 2μ plasmid YEpLK112 resulted in the multicopy recombinant strains, BELK100 and BELK100F, respectively. The transformants were selected on SC medium without tryptophan. Transforming the native plasmids pSTA1 and YEplac112 into the flocculent and non-flocculent strains of *S. cerevisiae* generated the corresponding reference strains. High efficiency transformation was performed using Lithium acetate method (5).

3.4 Southern hybridisation

Genomic DNA was isolated from five independent colonies for each transformation, digested with *Ns*iI, subjected to agarose gel electrophoresis and then blotted onto a nylon membrane. The blot was probed with 1.8 kb DIG-labelled *LKA1* gene to confirm integration of the cassette at the *URA3* locus.

3.5 Hydrolytic activity on raw starch

Recombinant strains were evaluated for their amylolytic activity on 20 g L^{-1} of raw corn starch medium consisting of 3 g L^{-1} starch azure, 6.7 g L^{-1} yeast nitrogen base with amino acids and 20 g L^{-1} agar (Difco). The cultures were incubated at 30°C for 40 hrs to observe starch degradation.

3.6 Measurement of flocculation

The flocculation ability of the yeast cells was measured according to the method previously described (22). Flocculent yeast cells were dispersed by washing twice with 100 mM EDTA and twice with sterilised water. The cells were resuspended in 3.0 ml of 50 mM phosphate buffer (pH 6.0), with and without 0.1% $CaCl_2$, to a final concentration indicated by an OD_{600} of 2.0. After being incubated at room temperature for 30 min, the cell suspension was agitated for 20 s using a vortex mixer at maximum speed and then incubated for 5 min. A 1 mL

sample from the upper phase of the cell suspensions was transferred to a cuvette, and the OD_{600} with and without $CaCl_2$ was measured. The flocculation ability (FA) was determined by means of the equation, FA=1-B/A, where A is the OD without $CaCl_2$ and B the OD with $CaCl_2$. Flocculation ability value (FA) close to 1.0 indicates strong flocculation ability.

3.7 Cell surface hydrophobicity

The solvent partition assay for cell surface hydrophobicity (19) was carried out at 4°C, with wide bore pipettes used for flocculating cell cultures. The cells were diluted to a concentration of $5x10^6$ cells mL^{-1} in sterile distilled water, after which 10 mL aliquots of yeast samples in a serum tube were washed once and resuspended in PUM buffer (22.2 g L^{-1} K_2HPO_4, 7.26 g L^{-1} KH2PO4, 1.8 g L^{-1} urea, 0.2 g L^{-1} $MgSO_4.7H_2O$). The absorbance of the cell suspension (I) was measured spectrophotometrically. Aliquots of 2.4 mL yeast solution were transferred to round-bottomed test tubes. Xylene (0.2 mL) was added to each tube, whereafter samples were vortexed for 30 s and then left to stand for 15 min. The xylene layer from each tube was subsequently removed and discarded. Absorbance of the aqueous layer was measured at 660 nm (F). The modified hydrophobicity index (MHI) was defined as I-(F/I), with MHI positively correlated to hydrophobicity.

3.8 α-Amylase activity assays

Amylase activity was measured using the dinitrosalicylic acid (DNS) method. The reaction mixture contained 1 mL of 1% raw starch in sodium monophosphate buffer at pH 6.9 and 500 μL culture supernatant. The mixture was made up to 2 mL using distilled water. It was incubated at 50°C for 30 min, 1 mL of DNS reagent was added, and the mixture was boiled for 15 min. After cooling, the reaction mixture was diluted with 9 mL of distilled water and the absorbance was measured at 540 nm. One unit of amylase activity was defined as the amount of enzyme causing the release of the equivalent of 1 mg of maltose per minute under the assay conditions.

3.9 Protein purification

A protocol based on anion exchange chromatography for the partial purification of proteins was devised using a Bio-Rad Model EP-1 Econo Pump (Bio-Rad). DEAE Sepharose™ fast flow matrix (Pharmacia Amersham) was equilibrated with 50 mM of MES buffer (pH 6.0). Equal proportions of supernatant and 100 mM MES buffer (1:1 ratio) were added to the column and mixed slowly for 2 hrs at 4°C using an end-to-end Labinco rotary mixer. Bound proteins were eluted with an increasing concentration series of NaCl (0.2 to 1M). Different fractions of the sample were collected and run through a YM-3 Microcon® centrifugal filter at 10 000 x g for 50 min to desalt the samples. Further purification and collection of proteins above 50 kDa was carried out by treating the eluent at 14000 x g for 12 min through a YM-50

Microcon® centrifugal filter. The retentate was collected by centrifugation at 1000 x g for 3 min. Partially purified proteins were further analysed by SDS/PAGE. The total protein concentration was determined using the 'protein assay dye reagent' (Bio-rad), with bovine serum albumin (Sigma) as the standard.

3.10 SDS-PAGE and western blot

SDS polyacrylamide gel electrophoresis (PAGE) was run essentially as described by the supplier (Hoeffer Scientific). Proteins were resolved in a discontinuous buffer system using a Mighty Small Vertical electrophoresis unit with a 7.5% stacking gel and resolving gel. After electrophoresis, the gels were subjected to Coomasie blue R250 staining and the resultant bands identified using wide range molecular mass standard (Sigma). The resolved peptides were then electrophoretically transferred onto an immobilon™ – P transfer membrane (Millipore) by semidry blotting (Towbin et al. 1979). The blotted proteins were identified immunochemically through the sequential addition of rabbit polyclonal anti-alpha amylase (Abcam), followed by donkey anti-rabbit I g horseradish peroxidase-linked whole antibody (Amersham Pharmacia). The secondary antibody was detected using the ECL Western blotting detection system (Amersham Pharmacia).

3.11 Immunochemical quantification of proteins

The protein samples to be quantified were adsorbed overnight on sterilin (Bibby-sterilin) immunoplates at 4°C. ELISA was performed as described by the suppliers of the primary antibodies (Abcam). To prevent non-specific binding, the wells were blocked with 300 µl of 1% skim milk powder at 37°C for 2 h. Phosphate-buffered saline at pH 7.5 was used as a washing buffer. After washing the plate twice, rabbit polyclonal anti-alpha amylase (Abcam) was added and the samples were incubated at room temperature for 2 h, followed by conjugation with anti-rabbit secondary antibody linked to alkaline phosphatase. The binding of the α-amylases was detected using alkaline phosphatase and para-nitro phenol phosphate substrate. The absorbance of released para-nitro phenol was measured at 405 nm using a microtitre plate reader (Bio-Tek Instruments). Different concentrations of α-amylase from *Bacillus amyloliquefaciens* (Sigma) were used to derive the ELISA standard for quantification.

3.12 Thin-layer chromatography

Starch hydrolysates were resolved on a silica gel 60 F_{254} thin-layer chromatography plate (Merck) with a solvent system of n-propanol-ethanol-water (7:1:2). After eight hours of separation, the plate was developed with 5% sulphuric acid in ethanol spray, followed by incubation at 110° C for approximately 5 min.

3.13 Flow-cell analyses

A four channel flow-cell system, with individual channel dimensions of 2.5 x 4 x 40 mm, was assembled and sterilized. The flowcell channels were filled with 0.3% starch azure, which was allowed to settle in the channels before inoculation of yeast strains. Samples (400ul) of overnight cultures of BELK1 and BELK1-F strains, each having a cell count of 2 $x10^7$ were inoculated into separate channels of duplicate flow-cell systems. After inoculation, the starch and cells were incubated for 1 h whereafter the interaction of starch and yeast cells was visualized with a Nikon Eclipse E400 epifluorescence microscope at 600x magnification. Images were captured with a Nikon Coolpix 990 digital camera. Channels were irrigated with peptone water at a constant flow rate of 6 ml h^{-1} using a Watson Marlow 205S peristaltic pump. Free-floating cells from different channels were collected and enumerated.

3.14 Atomic force microscopy

Adhesive forces between yeast cells and starch grains in flocculent versus non-flocculent strains were evaluated with atomic force microscopy (AFM). BELK1 and BELK1F strains were grown in YPD media and harvested by centrifugation at 5000 x g for 5 min. Cells were washed once with PBS (pH 8.0), and twice with distilled water. Flocculent and non-flocculent cells at an optical density (OD) of 0.2 were incubated in 1% corn starch for 30 h. A 2 μL aliquot was used for observation under the atomic force microscope. The Nanoscope III (Explorer) AFM (Veeco instruments) was used in contact mode to image cells and measure adhesive forces at different locations on the sample. Nanoprobe silicon nitride cantilevers with a spring constant of 0.06 N/m were used. Force measurements were carried out by engaging the AFM tip to scan the surface. The tip was approached to the surface in 100 nm increments with a specified Z scan size of 300 nm at a frequency of 1 Hz. Surfaces were imaged after every force to confirm the presence of a continuous lawn of sample. For control measurements starch samples without yeast cells, and cells without starch were used.

3.15 Scanning electron microscopy

The time frames of degradation, and pattern of hydrolysis of starch granules by flocculent and non-flocculent strains of *S. cerevisiae,* were monitored using scanning electron microscopy (SEM). BELK1 and BELK1F samples from the fermenter were removed at different time points, gold-coated with an Edwards sputter-coater and observed with a LEO435VP scanning electron microscope operated at 7kV. FY834 and FY835 were used as control strains.

3.16 Permeation of recombinant yeast cells

Recombinant strains expressing amylolytic enzyme activity were permeabilized by treatment with isopropyl alcohol (11). Following centrifugation at 5000 rpm for 5 min., yeast cell pellets

were washed once with 0.85% NaCl, whereafter wet cell samples (appr. 0.1g) were resuspended in 1 mL of 40% iso-propanol, and incubated at 4°C for 10 min with shaking. The permeabilized cells were then separated by centrifugation at 5000 rpm for 5 min, and washed with 0.85% NaCl.

3.17 Small-scale fermentations

Recombinant strains pre-grown in YPD medium at 30° C for 24 hrs were aseptically inoculated into the fermentation medium at a high cell density of 2 x 10^9 cells mL^{-1}. Small-scale bench fermentations were carried out with 100 mL of raw starch medium in 250 mL serum bottles plugged with rubber stoppers, and a gas outlet was secured by inserting a cannula. Fermentation was performed at 30°C for 90 hrs with agitation by magnetic stirring. Samples were taken every 12 hrs for the analyses of the fermentation products.

3.18 HPLC analyses of sugars and fermentation products

For the HPLC analyses, the supernatant samples from the fermentation were filtered using a 0.22μ millipore filter. The samples were centrifuged and the residual starch was subjected to acid hydrolysis using hydrolchloric acid. The glucose liberated was measured and the residual starch concentration was deduced. The concentrations of reducing sugars (maltotriose and maltose), ethanol, glycerol and acetic acid were determined by HPLC using an Aminex HPX87H column (Bio-Rad, Richmond, California). The following conditions were used: mobile phase, H_2S0_4 (6 mmol L^{-1}); flow rate, 0.8 mL min^{-1}; and temperature, 65°C, using a refractometer (23).

4. RESULTS

4.1 Construction of amylolytic *S. cerevisiae* strains with flocculent and non-flocculent phenotype

The list of plasmids and strains constructed for this study is given in Table 1. The integration of *LKA1* into the recombinant strains of *S. cerevisiae* was confirmed by Southern blotting (data not shown). Trp^+ transformants carrying multicopy vectors (containing the *TRP5* marker gene) were confirmed by their growth on selection plates lacking tryptophan (*TRP5*). The transformation of a functional copy of *FLO8* into the strain FY834 resulted in a flocculent strain, FY835. This flocculent strain (FY835) grew invasively into SLAD agar plates, but was unable to use starch as a sole carbon source. The flocculent amylolytic recombinant strain BELK1F showed adhesion and invasive growth on SLAD agar plates, whilst the non-flocculent amylolytic strain BELK1 did not invade agar (Fig. 2B).

4.2 Expression, purification and analysis of amylolytic activity in flocculent and non-flocculent *S. cerevisiae*

Constitutive expression using the *PGK1* promoter and terminator and the use of the *MFα1S* secretion signal sequences led to the expression and secretion of active biological LKA1 protein in *S. cerevisiae*. The use of the *MFα1S* secretion signal resulted in increased secretion of biologically active protein into the culture supernatant while only 10% of the enzyme was intracellular (data not shown). The wild-type non-flocculent strain FY834, and the flocculent FY835, did not exhibit α-amylase activity on raw starch substrate. All recombinant strains expressing *LKA1* showed significant amylolytic activity, which was detected by zones of degradation on corn starch plates. The α-amylase Lka1p from the supernatant of the recombinants was partially purified and analysed using SDS-PAGE (Fig.1) and Western blotting. A clear band near 68 kDa, which represents α-amylase, was observed in the enzyme preparation from recombinant strains while the parent strain consistently did not show this band.

Amylolytic activity of the recombinant strains was analysed using raw corn starch (Table 2). Expression is essentially a linear function of gene copy number; however, in this case the transformation of multicopy episomal plasmid did not cause any significant increase in the enzyme activity in comparison with that of the single copy integrants. Permeation of the flocculent and non-flocculent cells using iso-propanol did not have any significant effect on the secretory efficiency of the cells.

4.3 Effect of flocculation on the time window of starch hydrolysis

The wild-type strains (FY834 and FY835) and the recombinant strains (BELK1, BELK1-F, BELK100 and BELK100-F) were efficient maltose utilizers and, when grown on maltose as a sole carbon source, the performance of both the flocculent and non-flocculent amylolytic strains was similar (data not shown). All the recombinant strains expressing *LKA1* were able to grow on raw corn starch as a sole carbon source (Fig. 2). Despite similar Lka1p levels in the supernatant, the non-flocculent strains BELK1 and BELK100 had a lag phase of 48 h before entering logarithmic growth phase, while the flocculent strains (BELK1-F and BELK100-F) had a shorter lag period and active growth occurred after only 36 h (Fig. 2A). Qualitative analyses of the supernatant from the recombinant strains were performed at time points between 12 to 90 h of growth in starch media, using thin layer chromatography to observe the time frame of hydrolysis. It was noted that detectable amounts of reducing sugars, such as maltose and maltotriose, were present in the media as early as 36 h in the flocculent amylolytic strains, although these were only detected in the non-flocculent strains after 48 h (Fig. 2B and 2C). This suggested the existence of an earlier time window of hydrolysis in the flocculent phenotype, despite similar enzyme concentrations and activity in both the flocculent and non-

flocculent cells in the stipulated time period. It was also noted that maltose was readily utilised by the strains, but there was a build-up of maltotriose in the media.

4.4 Assessment of starch-cell interactions *in situ*

To assess the properties of flocculent yeast cells that might enable them to shift the window of hydrolysis, criteria such as hydrophobicity and cell adhesion were studied. A solvent partition assay showed that flocculent cells had a higher cell surface hydrophobicity index compared to non-flocculent strains (Fig. 3A). This property of flocculent cells was growth-phase dependent, although it was observed that these cells were consistently more hydrophobic than their non-flocculent counterparts throughout late log phase and stationary phases. The invasive phenotype of flocculent strain BELK1 F is shown in Fig. 3B.

To follow whether the higher surface hydrophobicity of the flocculent cells contributed to starch-"floc" interactions, a flow-cell system was used for *in situ* observations (Table 3). Applying light microscopy, it was noticed that the flocculent cells adhered to the surface crevices of the raw starch, in contrast to the non-flocculent cells, which did not exhibit such adherence properties. Representative samples of the flocculent and non-flocculent strains from the flow-cell channels revealed that, of the 10^7 cells inoculated in each flow-cell channel, only 10^3 flocculent cells were recovered in the effluent, while as many as 10^5 cells were present in the non-flocculent effluent.

Multiple force curves recorded at various locations using atomic force microscopy demonstrated single and multiple unbinding forces in about 50% of a total of 100 force measurements in both flocculent and non flocculent strains. A force of 120 nN $\pm$ 12 was recorded for the flocculent cell samples in the starch matrix. A lower value of 32 nN $\pm$ 7 was observed for the non-flocculent cells in the starch matrix, suggesting lower adhesion compared to the flocculent cells. This affirms significant differences in interactive forces between flocculent cells and non-flocculent cells incubated in the starch matrix.

4.5 SEM analysis of the hydrolysis pattern

The shift in the time window of hydrolysis and differences in the hydrolytic pattern of flocculent and non-flocculent cells were observed with scanning electron microscopy (Fig. 4). In flocculent amylolytic strains, the degradation actively began after 36 h and progressed until 72 h, showing characteristic patterns of extreme pitting, shredding and liquefaction. The surface pores on the corn starch granules allow hydrolytic attack by the enzyme and, in the later stages of degradation, a "sieving effect" was observed in which the flocculent cells synergistically sheared the liquefied starch matrix by rapid enzymatic hydrolysis. In non-flocculent cells, degradation was initiated after 48 h and pit formation, gritting and hydrolysis were observed. Unlike the situation in flocculent cells, the rate and the time of sieving of the starch matrix was slower and less efficient, leading to prolonged times of degradation.

4.6 Fermentation characteristics of flocculent *vs.* non-flocculent amylolytic *S. cerevisiae*

The groups of recombinant flocculent and non-flocculent *S. cerevisiae* strains (both the permeated (BELK1P, BELK100P, BELK-FP, BELK100-FP) and the non-permeated (BELK1, BELK100, f BELK-F, BELK100-F) were studied for their fermentative ability on raw starch in small batch-scale fermentations at 72 h. The byproducts formed during the fermentation are listed along with the residual starch (Table 4). The flocculation efficiency of the strains after 72 h of fermentation is also presented. The data indicates that the fermentation performances of multicopy transformants are generally similar to that of the integrants. It was also noticed that there was a modest improvement in the production of all fermentation byproducts by the flocculent strains, with no influence from the permeation factor. However, a significant difference was found in the residual starch concentrations, with the flocculent strains showing less residual starch than the non-flocculent strains at 72 h. Furthermore, the fermentation profile of the integrant strains with both non-flocculent (BELK1) and flocculent (BELK1-F) phenotypes were studied in detail (Fig. 5). With initial concentration of 20g raw starch supplied in the fermenter, the non-flocculent strain consumed 76% of starch in the medium, producing 4.61 g L^{-1} of ethanol after 90 h of fermentation, whereas the flocculent strain utilised 82% of raw starch, releasing 5.1 g L^{-1} of ethanol at 90 h. In the non-flocculent strain BELK1, detectable amounts of ethanol were formed in the media after 48 h while in the flocculent strain BELK1-F ethanol production commenced after only 36 h, with a shorter lag in the fermentation phase. Ethanol yield by flocculent strains increased by at least 0.5 g L^{-1} when compared to non-flocculent strains. A similar trend was also noticed in the production of other byproducts, such as glycerol and acetic acid. Furthermore, an early decrease in the residual starch concentration was noticed for the flocculent strains when compared to their non-flocculent counterparts indicating a shift in the time period of hydrolysis.

5. DISCUSSION

In our previous study, we reported that non-flocculent *Saccharomyces cerevisiae* strains expressing the α-amylases Lka1p and Lka2p from *Lipomyces kononenkoae* degrade soluble starch efficiently by synergistic enzymatic activities (3). In the present study, we describe the effect of a flocculent phenotype on raw starch by *S. cerevisiae* strains expressing the raw starch-degrading α-amylase Lka1p. Although the recombinant flocculent and non-flocculent amylolytic strains exhibited little or no difference in their α-amylase activities, a significant shift in the growth pattern and time period of raw starch hydrolysis was noticed for the flocculent strains. Our data suggest that the physicochemical properties such

as hydrophobicity and cell adhesion, which are associated with the surface of flocculent amylolytic strains, may imply an advantage over the non-flocculent cells in the earlier hydrolysis of starch substrate. In addition, we also suggest that flocculation ability by itself would have contributed towards the earlier production of ethanol, since a high cell density can be maintained within the fermentor without the aid of immobilising agents (20). At the same time, the fact that flocculation and the invasive growth phenotypes in yeast are co-regulated (10) implies that the improved ability of flocculent strains to act and hydrolyse starch was not only due to their increased adhesion to the substrate, but would also have been due to their ability to grow invasively into the growth substrate (4, 8, 12). Moreover, these adhesion properties of the flocculent cells would effectively have reduced the time interval for enzyme-substrate interactions, thus amplifying the rate of hydrolysis at an earlier time period.

Starch consists of amylose and amylopectin, which are inherently incompatible molecules, and the amylose component forms helical structures which are relatively stiff and may present contiguous hydrophobic surfaces (6). During the initial enzymatic attack, the inner hydrophobic core of starch rotates when it is in contact with water, thereby exposing the hydrophilic part of the polymer to the aqueous phase. The hydrophobic surfaces in starch may interact with the hydrophobic flocculent cells, thus imbibing them into the starch matrix. The hydrolysis and degradation of the substrate subsequently take place from the inside of the starch matrix. This mode of hydrolysis seems to be more effective, since the flocculent biocatalysts do not dilute the enzyme (in this case, Lka1p) secreted into the surrounding medium. The encounter of non-flocculent strains with the substrate is merely due to the presence of cells and starch in the same suspension, but no physicochemical interactions occur. Thus the enzyme secreted by the non-flocculent cells disperses into and dilutes in the liquid media, prolonging the start of effective hydrolysis and thus of the fermentation. This might explain the early shift in the hydrolytic window in the flocculent cells. Scanning electron micrographs confirm this unique pattern of degradation by flocculent biocatalysts. The synergistic sieving effect of flocculent cells on a starch matrix enables the cells to enter the fermentation phase with a shorter lag phase.

While there was no difference in performance between transformants carrying an episomal (so-called) multi-copy plasmid, this does not mean that the level of expression of α-maylase does not impact on overall amylase activity. The multicopy status of Yeplac12 transformants was not checked and, more importantly, expression levels of the amylase gene were not determined]

The formation of a “yeast net” does not affect the secretion of enzymes into the culture medium. This was confirmed by data indicating that both the permeated and non-permeated cells had similar extracellular enzyme levels. Permeation by cells did not affect their flocculation efficiency (19) or their enzyme levels, while their bioconversion efficiency

decreased slightly due to lower metabolic rates. It has been reported that the presence of ethanol enhances the permeability of *S. cerevisiae* (15) which might suggest that increasing ethanol concentrations affected cell integrity and thus the metabolic machinery of the cells.

Similar to various other surface display systems (1, 9, 15, 20), the recombinant amylolytic flocculent cells in this study utilized the starch matrix as a surface to adhere to, resulting in early hydrolysis. As has been shown, the flocculent biocatalysts with suitable physicochemical properties could serve as efficient cell factories for industrial fermentations.

In conclusion, we successfully demonstrated that flocculent biocatalysts are efficient in the direct fermentation of raw starch to ethanol with physico- chemical characteristics like hydrophobicity and adhesion playing an important role in their performance. One of the direct effects of flocculation was that it facilitated the economic use of secreted α-amylase Lka1p, reducing the dilution effect on the enzyme and thus the time period of fermentation. Co-expression of Lka1p with other amylolytic enzymes could also be an effective strategy. Since rapid fermentations and high yields of ethanol are desirable, further optimisation of fermentation processes using the flocculent strains from this study might lead to enhanced rates of ethanol production in shorter fermentation times.

6. ACKNOWLEDGMENTS

This project was financially supported by South Africa's grapegrowers and winemakers through their investment body Winetech and by South Africa's National Research Foundation. The authors would like to thank Mr.Jaco Minaar, Insitute for Wine Biotechnology for the HPLC analysis and Ms. Martina M, Institute for Polymer Science for atomic force microscopy.

7. REFERENCES

1. **Atia, K. S, S. A. Ismail., M. B. El-Arnaouty., A. M. Dessouki.** 2003 Use of co-immobilized beta-amylase and pullulanase in reduction of saccharification time of starch and increase in maltose yield. Biotechnol. Prog. **19:**853-857.
2. **Birol, G., I.Onsan., B. Kidar., and S. G. Oliver**. 1998. Ethanol production and fermentation characteristics of recombinant *Saccharomyces cerevisiae* strains grown on starch. Enzyme. Microb. Technol. **22:**672-677.
3. **Eksteen, J.M., P. van Rensberg., R.R. Cordero Otero., and I. S. Pretorius**. 2003. Starch fermentation by recombinant *Saccharomyces cerevisiae* strains expressing the α-amylase and glucoamylase genes from *Lipomyces kononenkaoe* and *Saccharomycopsis fibuligera*. Biotechnol. Bioeng. **84**:639-646.

4. **Gagiano, M., F. F Bauer., and I. S. Pretorius**. 2002. The sensing of nutritional status and the relationship to filamentous growth in *Saccharomyces cerevisiae.* FEMS. Yeast. Res. **2**:433-470.
5. **Gietz, R. D., and R. H. Schiestl**. 1991. Applications of high efficiency lithium acetate transformation of intact cells using single-stranded nucleic acid as carrier. Yeast. **7**:253-263.
6. **Immel, S., and F. W. Lichtenthaler** 2000. The Hydrophobic Topographies of Amylose and its Blue Iodine Complex Starch – Stärke **52:**1-8.
7. **Janse, B. J. H., and I.S. Pretorius**. 1995. One step enzymatic hydrolysis of starch by a recombinant strain of *Saccharomyces cerevisiae* encoding α-amylase, glucoamylase and pullulanase. Appl. Microbiol. Biotechnol. **42**:878-883.
8. **Kobayashi, O., H. Suda., T. Ohtani., and H. Sone**. 1996 Molecular cloning and analysis of the dominant flocculation gene FLO8 from *Saccharomyces cerevisiae*. Mol. Gen. Genet. **251**:707-715.
9. **Kondo, A., H. Shigechi., M. Abe., K. Uyama., T. Matsumoto., S. Takahashi., M. Ueda., A. Tanaka., M. Kishimoto., and H. Fukuda**. 2002. High-level ethanol production from starch by a flocculent *Saccharomyces cerevisiae* strain displaying cell-surface glucoamylase. Appl. Microbiol. Biotechnol. **58:**291-296.
10. **Liu, H., C. A. Styles., and G. R. Fink**. 1996. *Saccharomyces cerevisiae* S288C has a mutation in *FLO8*, a gene required for filamentous growth. Genetics **144**:967-978.
11. **Liu, Y., H. Hama., Y. Fujita., A. Kondo., Y. Inoue., A. Kimura., and H. Fukuda**. 1998. Production of S-Lactoglutathione by high activity whole cell biocatalysts prepared by permeabilization of recombinant *Saccharomyces cerevisiae* with alcohols. Biotechnol. Bioeng. **64:**54-60.
12. **Madhani, H. D., and G.R. Fink**. 1998 The control of filamentous differentiation and virulence in fungi. Trends. Cell. Biol. **8**:348-353.
13. **Masy, C. L., A. Henquinet., and M. M. Mestdagh** 1992. Flocculation of *Saccharomyces cerevisiae*: inhibition by sugars. Can. J. Microbiol. **38:**1298-306.
14. **Struhl, K., D. T. Stincomb., S. Scherer., and R.W. Davis**. 1979. High frequency transformation of yeast: Autonomous replication of hybrid DNA molecules. Proc. Natl. Acad. Sci. USA **76**:1035-1039.
15. **Mizoguchi, H., S. Hara.** 1996. Effect of fatty acid saturation in membrane lipid bilayers on simple diffusion in the presence of ethanol at high concentrations. J. Ferment. Bioeng. **81**: 406-411.
16. **Murai, T., M. Ueda, M. Yamamura., H. Atomi., Y. Shibasaki., N. Kamasawa., M. Osumi., T. Amachi., and A. Tanaka**. 1997. Construction of starch-utilizing yeast by cell surface engineering. Appl. Environ. Microbiol. **63:**1362–1366.

17. **Murai, T., M.Ueda., Y. Shibasaki., N. Kamasawa., M. Osumi., T. Imanaka., and A. Tanaka**. 1999. Development of an arming yeast strain for efficient utilization of starch by co-display of sequential amylolytic enzymes on the cell surface. Appl. Microbiol. Biotechnol. **51:**65-70.
18. **Nakamura, Y., F. Kobayashi., M. Ohnaga., and T.Sawada**. 1997. Alcohol fermentation of starch by genetic recombinant yeast having glucoamylase activity. Biotechnol. Bioeng. **53:**21–25.
19. **Powell, C. D, D. E. Quain., and K. A. Smart**. 2003 The impact of brewing yeast cell age on fermentation performance, attenuation and flocculation. FEMS. Yeast. Res. **3:**149-57.
20. **Rodriguez, E., and D. A. S. Callieri**. 1983 Conversion of sucrose to ethanol by a flocculent *Zymomonas* sp. isolated from sugar cane juice. Biotechnol. Lett. **8**:744-748.
21. **Shigechi, H., K. Uyama., Y. Fujita., T. Matsumoto., M. Ueda., A. Tanaka., H. Fukuda., and A. Kondo**. 2002. Efficient ethanol production from starch through development of novel flocculent yeast strains displaying glucoamylase and co-displaying or secreting α-amylase. Journal of Molecular catalysis B: Enzymatic **17:** 179-187.
22. **Smit. G., M. H. Straver., B. J. J. Lugtenberrg., and J. W. Kijne**. 1992 Flocculence of *Saccharomyces cerevisiae* cells is induced by nutrient limitation, with cell surface hydrophobicity as a major determinant. Appl. Environ. Microbiol. **58**: 3709-3714.
23. **Sreenath, H.K., R.G. Koegel., A.B. Moldes., T.W. Jeffries., and R.J. Straub 2001**. Ethanol production from alfalfa fiber fractions by saccharification and fermentation. Process Biochemistry. **36**:1199-1204.
24. **Steyn, A. J. C., and I. S. Pretorius**. 1990. Expression and secretion of amylolytic enzymes by *Saccharomyces cerevisiae.* Acta Varia **5:**76-126.
25. **Towbin, H., T. Staehelin., and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA **76**: 4350-4354.
26. **Winston, F., C. Dollard., and S. L. Ricupero Hovasse**. 1995. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. Yeast **11:**53-55.

TABLE 1. Strains and plasmids used in this study.

Strains	Genotype	Reference
FY834	*MATα his3 leu2 lys2 trp1 ura3flo8*	25
FY835	*MATα his3 flo8::FLO8 lys2 trp1 ura3*	This study
BELK1	*MATα his3 leu2 lys2 trp1 ura3::LKA1flo8*	This study
BELK1-F	*MATα his3 flo8::FLO8 lys2 trp1 ura3::LKA1*	This study
BELK100	*MATα his3 leu2 lys2 trp1 LKA1flo8*	This study
BELK100-F	*MATα his3 FLO8 lys2 trp1 LKA1*	This study
Plasmids		
YIpLac22-FLO8	Ap^R $FLO8_{PT}$ *FLO8 LEU2*	Dewald van Dyk (unpublished)
YIp5	Ap^R Tc^R *URA3*	This study
pSTA1	Ap^R $Tc^R$$PGK_{PT}$ *URA3*	This study
pSTA2	Ap^R $Tc^R$$PGK1_P$ *LKA1* $PGK1_T$ *URA3*	This study
YEpLac112	2μ origin Ap^R *LacZ* $PGK1_P$ *LKA1* $PGK1_T$*TRP5*	4
YEpLK112	2μ origin Ap^R *LacZ* PGK_P *LKA1* PGK_T*TRP5*	This study

TABLE 2. Enzymatic activity of LKA1 α- amylase from recombinant *S. cerevisiae* strains at 48hrs

Strain	Enzyme activity ($U \cdot mg^{-1} \cdot$ total protein)[a]	α-amylase concentration ($U \cdot L^{-1}$)
Control	11.3 ± 3.2	ND
BELK1	121.0 ± 23.5	87
BELK1F	123.8 ± 27.4	82
BELK1^{P}	124.7 ± 32.1	90
BELK1F^{P}	125.2 ± 31.7	85
BELK100	133.5 ± 27.9	83
BELK100F	132.6 ± 28.9	86
BELK100^{P}	131.2 ± 27.7	85
BELK100F^{P}	132.7 ± 29.2	87

[a] 1 unit of enzyme liberates 1 mg of maltose from 1% starch per minute

[p] permeated

TABLE 3. Flow cell analyses of flocculent (BELK1-F) and non-flocculent (BELK1) strains in starch azure.

Strains	Initial inoculum[a]	Eluted cells[a]	Retained cells[a]
BELK1F	1×10^{7}	1×10^{2}	1×10^{5}
BELK1	1×10^{7}	1×10^{4}	1×10^{3}

[a] cell count$\cdot mL^{-1}$, Standard deviation <10%

TABLE 4. Comparison of starch consumption and byproducts among flocculent and non-flocculent amylolytic *Saccharomyces cerevisiae* strains at 72 h of the fermentation phase.

Strain	Ethanol[a]	Glycerol[a]	Acetic acid[a]	Residual starch[a]	%Flocculation ability
BELK1	4.57	1.23	2.31	4.8	ND
BELK1F	4.92	1.17	2.23	3.2	53 ± 3
BELK1^{P}	4.50	1.12	2.26	5.2	ND
BELK1F^{P}	4.89	1.21	2.34	4.2	57 ± 4.5
BELK100	4.42	1.23	2.25	5.2	ND
BELK100F	4.90	1.24	2.27	4.2	59 ± 2.6
BELK100^{P}	4.62	1.17	2.24	4.8	ND
BELK100F^{P}	4.88	1.15	2.03	3.8	57 ± 2

[a] $g \cdot L^{-1}$

[p] permeated

ND - Not detectable

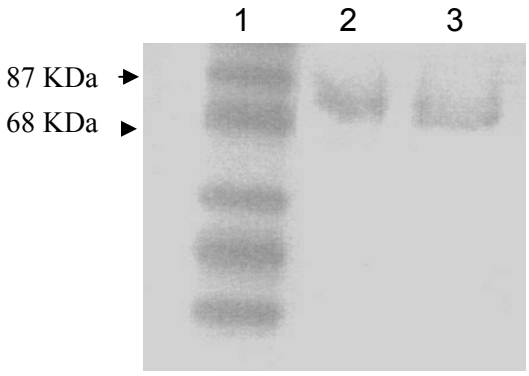


FIG.1. SDS PAGE of purified α-amylase Lka1p from the yeast supernatant: Lane 1. Sigma wide range molecular weight marker, lane 2. BELK1, lane 3. BELK1F

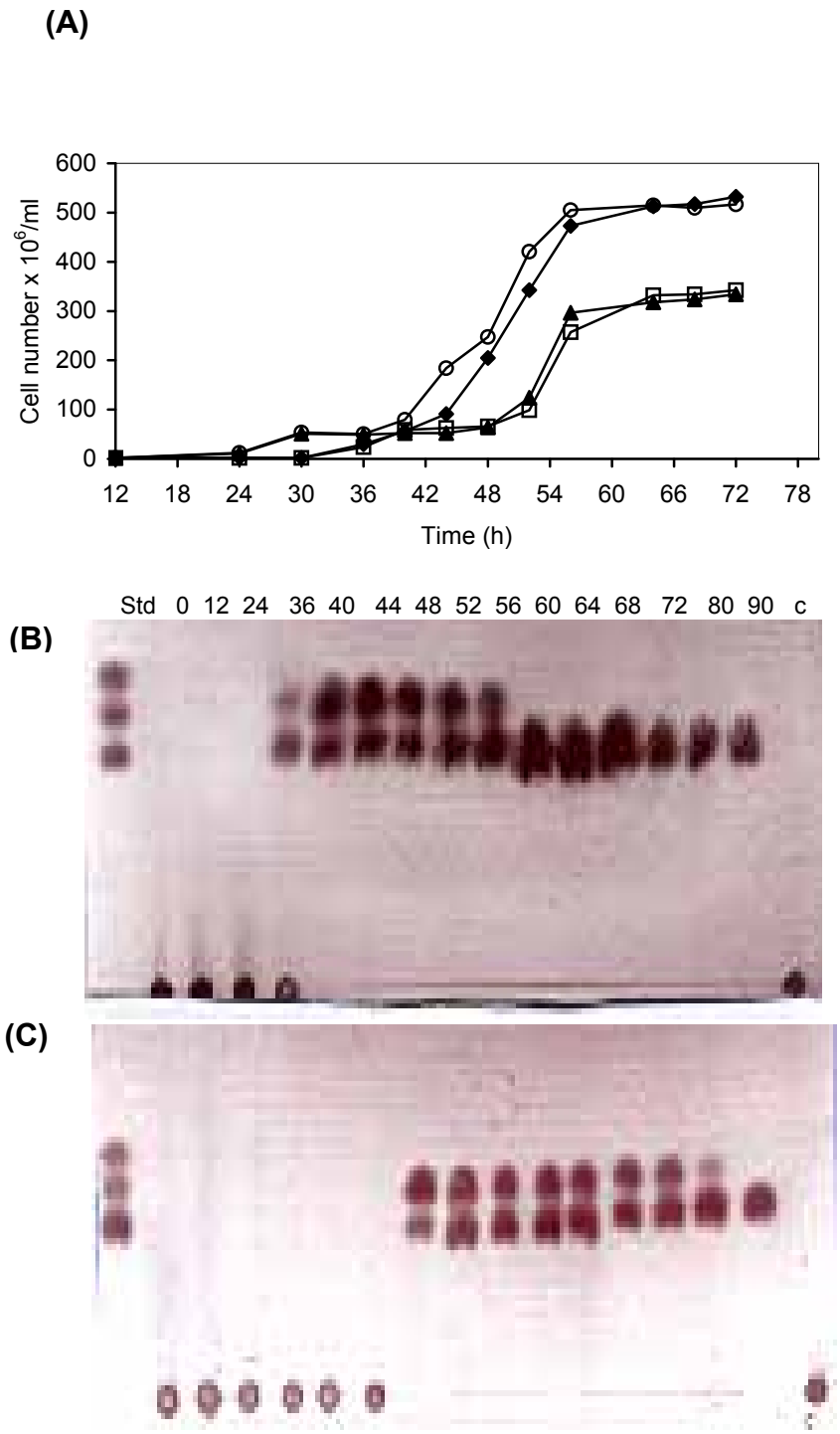


FIG. 2. (A) Growth pattern of recombinant amylolytic strains on 2% raw corn starch: BELK1 (▲), BELK1-F (◇), BELK100 (□), BELK100-F (◆). Each value is a mean of three determinations ± standard deviation, SD<10%, (B) Thin layer chromatography: Comparison of time window of active starch hydrolysis during the growth phase of flocculent and (C) non-flocculent amylolytic recombinants G-glucose; G2-maltose; G3-maltotriose; c - non amylolytic strains (FY835 in (B) and FY834 in (C); sampling time 60 h); Std-standard mix of glucose, maltose and maltotriose.

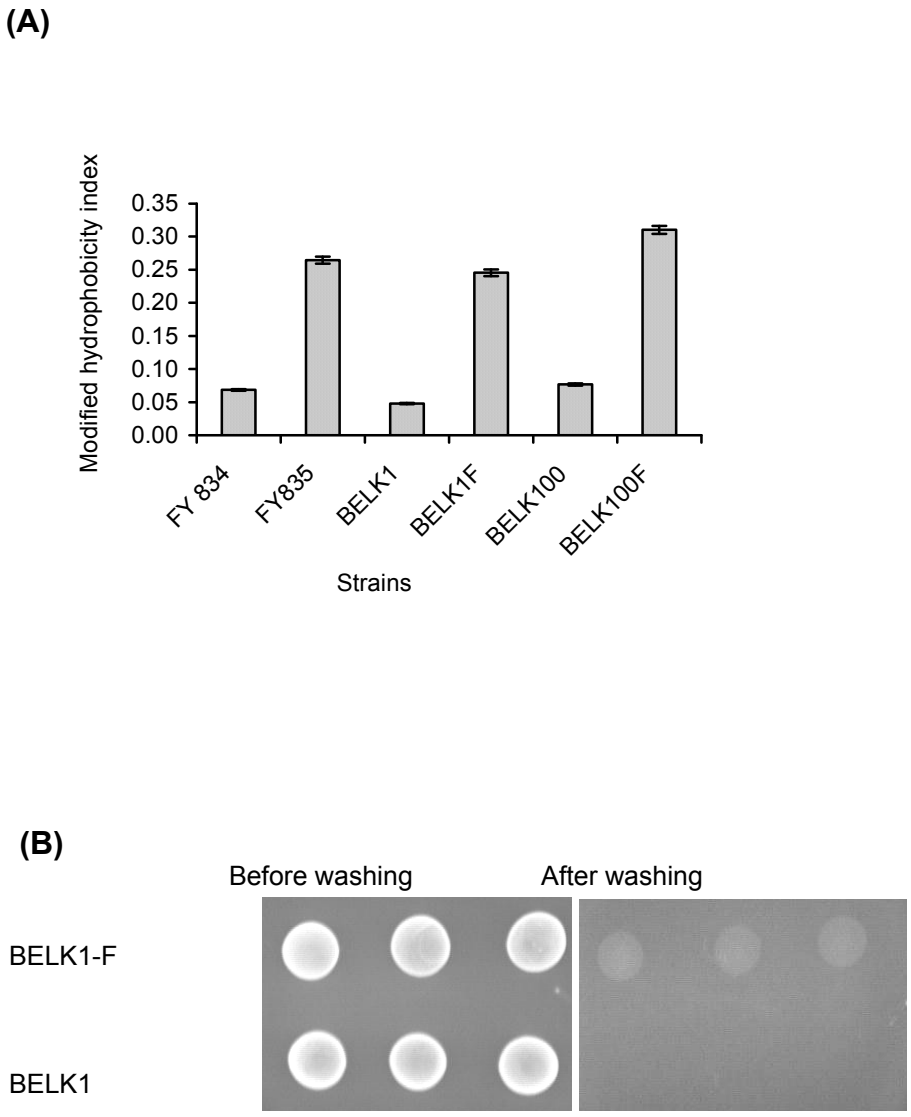


FIG. 3. (A) Assessment of cell surface hydrophobicity of flocculent and non-flocculent *Saccharomyces cerevisiae* cultures grown in YPD for 36 h at 30°C. Data represented are mean of two individual experiments and SD was >10% (B) Assessment of invasive phenotype of flocculent (BELK1F) and non-flocculent (BELK1) amylolytic *S. cerevisiae* strains grown on SLAD agar plates for 3 days at 30°C.

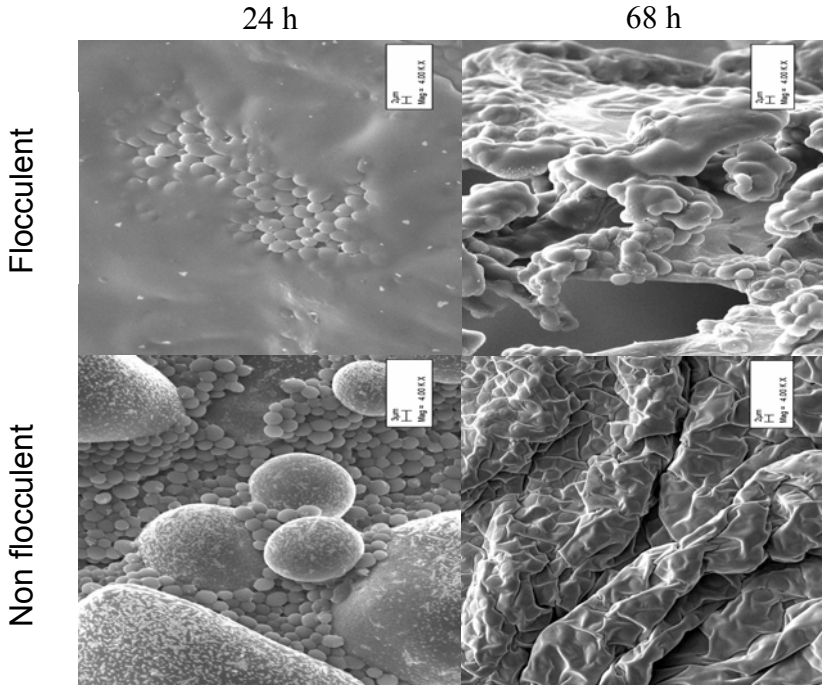
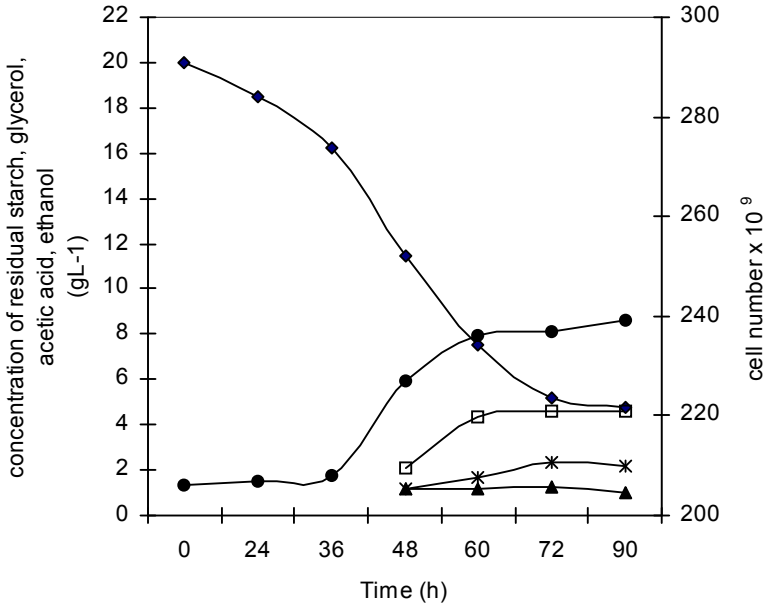


FIG. 4 Ultra structure of starch hydrolyzed by flocculent and non-flocculent amylolytic *S. cerevisiae* at different time periods in the fermentor. All observations recorded at 4000 X; scale bars are 3µm.

(A)



(B)

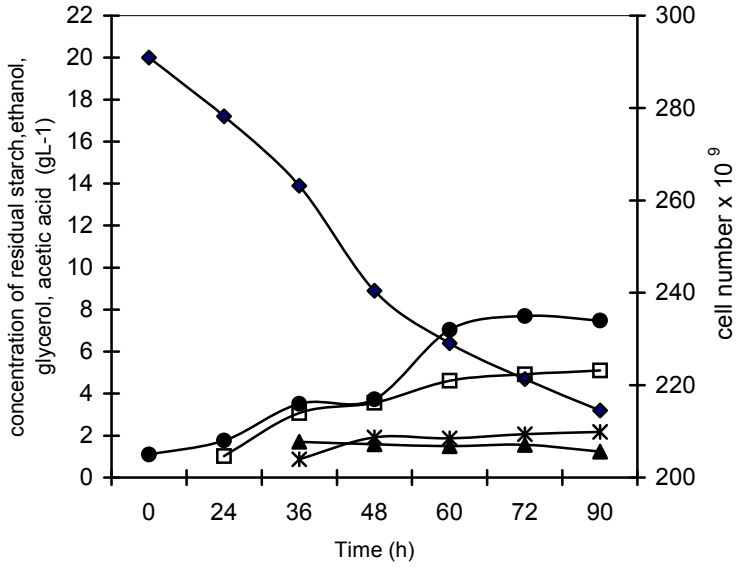


FIG. 5. Comparison of product profiles of flocculent and non-flocculent recombinant *S. cerevisiae* strains during small-scale batch fermentations. A) Non-flocculent amylolytic strain BELK1; B) Flocculent amylolytic strain BELK1F: (□) Ethanol; (♦) starch; (▲) acetic acid; (✳) glycerol; (●) biomass.

CHAPTER 6

GENERAL CONCLUSION

1. GENERAL CONCLUSION

The potential for biofuel production and generating energy from biomass is both very large and underutilized. However, the feasibility of using biomass as a source of fuel for automobiles, factories and electricity generation depends on the availability of appropriate technologies and harvesting of sometimes-remote biomass resources.

Technologies through which biomass can be converted into fuels and energy include enzyme biotechnology and fermentation processing. Fermentation of plant sugar into ethanol is well known since the beginning of recorded history. New technologies have sped up these processes. Today's novel enzymes make it possible to generate alcohol from starch rich plant biomass, several thousand times more efficiently than the earlier brewers and distillers.

Biofuel forms part of major research schemes in both developed and developing countries, but the research and development activities with regard to biofuels were not only based on reducing the oil dependency but also on product improvement and environmental issues.

The promise of bio-catalyst technolology in the quest for renewable energy conversion and sustainable development is that it can one day offer an extensive range of products that can fulfil a wide variety of human, industrial and ecological needs. Thus in a broad sense, biotechnology that we focus on today concerns the application of biosystems-cells of microbial, plant and animal origin, parts thereof and molecular analogues in bioindustries.

Although biotechnology and its applications are rapidly developing, a few methodologies have acquired a strong position and can be expected to influence the future development of applied biotechnology. These in brevity could be referred to as the biotechnological paradigms –a set of methods that experts try first when they seek new solutions or try to understand novel phenomena. The capabilities of the research teams, enterprises and countries depend on their ability to recognise, master and further develop the methodologies of the biotechnological paradigm.

One of the important methodologies used in the applied biotechnology concerns screening, rapid systematic evaluation of a large number of compounds/ enzymes and microbial species for a targeted application. The term "high-thoroughput" screening has been adopted to depict methodologies that include rapid screening of a number of compounds. In direct DNA screening, DNA sequences of living organisms can be screened by locating genes that code proteins having desired properties.The second set of methods concerns the combinatorial chemistry. It is used to produce large number of modifications of compounds, materials and enzymes. As a whole the process resembles natural selection. As it reduces the need for theoretical modelling studies and yields candidate compounds that can be tested rapidly, the methodology is efficient in applied research.

Yet another leap in the field of applied biotechnology concerns the “molecular modification” or changing the structure of an individual molecule or gene in a controlled fashion. In developing modification technologies, the aim is to enable accurate and yet productive operations at molecular level. While modification may involve only minor changes in the structure of the genes or proteins, the resulting changes in the behaviour of the enzyme could be dramatic. The practical success however, depends on the extent to which the properties of the modified molecules can be predicted by computational and other means.

With the advent of recombinant technology, the use of gene transfer or giving an individual or a species a capability to express a certain protein, for example, genetic manipulation of microorganisms to render them applicable in a specific hydrolytic process or bioremediation etc. is also employed widely in bioprocessing.To complement and to further the understanding, “bioinformatics” or the application of computer based technologies has been a useful tool. Computerised methods include those used in building databases of proteins, other molecules and genes. Various techniques of molecular modelling are being developed with one of the goals being prediction of protein folding patterns on the basis of global optimisation of potential energy function of the protein molecule.

The use of starch hydrolases to convert renewable starch biomass into sugars and further to ethanol by fermentative organisms has been a major study topic for years. Increasing environmental pressures and energy prices will make this application a real possibility one day. In the current thesis we have used some of the above-mentioned methodologies to develop molecular and cellular catalysts for effective bioconversion of raw starch. The *LKA1*-encoded α-amylase (Lka1p) from *Lipomyces kononenkoae* belongs to glycosyl hydrolase family 13 and the catalytic properties of these enzymes, such as their broad substrate specificity and raw starch-hydrolysing function, motivated the study of structure-function relationships in these enzymes. In this study, we constructed variants of the Lka1p enzyme so as to understand the native enzyme and its functional properties. The presence of variations in the primary structure in comparison with other α-amylases and bioinformatics-based data led to mutational analyses with regard to its raw starch-degrading function and its substrate specificity towards complex starches and related polymers. The results from mutagenic and kinetic approaches presented in this thesis will form the basis for the further development of fusion proteins or efficient enzymes with a broad substrate range and different industrially important biochemical properties. The two major findings of this study include the characterisation of a raw starch-binding domain at the N terminus of the Lka1p enzyme, and the discovery of the importance of a glutamine residue that replaces the invariant histidine in glycosyl hydrolases present at this position in substrate specificity towards α-1,4 and α-1,6 linkages. These findings are summarised in chapters 3 and 4. Thus, the information available on the organisation observed in Lka1p could be used for the construction of fusion proteins or chimera, either for the manipulation of activity towards

substrates or, for instance, for binding to solid raw substrates. Although the current study does not provide a deep insight into the exact mechanism behind these processes, it provides some understanding of possible sugar-binding sites and chemical and physical properties of the molecule when essential amino acids are mutated or truncations of protein are created.

The project also focussed on the development of an appropriate whole-cell catalyst expressing Lka1p for starch hydrolysis. The construction of a flocculent *S. cerevisiae* strain expressing *LKA1* and the effect of this genetic background on the hydrolysis of raw starch was discussed in Chapter 5. It was demonstrated, however, that the physicochemical properties of flocculent cells led to an earlier time window of hydrolysis, merely as a result of the reduction in the time interval of interaction between the raw starch substrate and the enzyme secreted. This understanding will be a platform for future efforts to develop proficient catalysts for the conversion of the renewable starch resources to ethanol.

In general, knowledge about the glycosyl hydrolase family members is expanding. In particular, specificity and substrate recognition studies represent advances in understanding the biology of these enzymes. Tailoring enzymes for specific applications will be a future trend with continuously improving tools and understanding of structure-function relationships and increased search for enzymes from exotic environments. New technical tools to use enzymes as crystalline catalysts, ability to recycle cofactors, and engineering enzymes to function in various solvents with multiple activities are important technological developments, which will steadily create new applications. Enzymes should, however, not be considered alone but rather as a part of a biocatalyst technology. Whole cell catalysts, increased ability to engineer metabolic pathways and a combination of specific biocatalytic reactions with organic chemistry form a basis to develop new technologies for chemical production.

In a long run, it is thus possible that a new holistic approach emerges in creating means of effective degradation of renewable raw materials for ethanol production. Such an approach could focus on rapid development of new enzymes and enzymatic processes to suit the needs of the growing industry.

APPENDIX

Amylolytic enzymes from the yeast *Lipomyces kononenkoae*

This manuscript has been accepted for publication as a mini-review in the journal

BIOLOGIA

Amylolytic enzymes from the yeast *Lipomyces kononenkoae*

Nivetha RAMACHANDRAN[1], Isak S PRETORIUS[1,2] & Ricardo R CORDERO OTERO[1*]

1 *Institute for Wine Biotechnology, Department of Viticulture & Oenology, Stellenbosch University, Victoria Street, Stellenbosch, ZA 7600, South Africa; Phone: ++27 21 8083770, fax: +27 21 8083771, email: rrco@sun.ac.za*

2 *The Australian Wine Research Institute, Waite Road, Urrbrae, Adelaide, SA 5064, Australia; Phone:* ++61 8 8303 6600, Fax: ++61 8 8303 6601

1. Abstract

The *Lipomyces kononenkoae* α-amylases LKA1 and LKA2 belong to the glucosyl hydrolase Family 13 and exhibit specificity towards α-1,4 and α-1,6 linkages in starch and related substrates. LKA1 exhibits specificity towards α-1,4 and α-1,6 linkages and large amounts of reducing sugars are liberated from highly branched amylopectin and glycogen and linear amylose. LKA2, on the other hand, shows high reactivity towards lintner starch, dextrin and amylase, although only small amounts of reducing sugars are liberated from branched substrates such as amylopectin and glycogen. These enzymes share the four conserved segments of the catalytic domain found in other members of the family, but have some major variant amino acids within these segments. In addition, LKA1 consists of an N-terminal starch-binding domain (SBD). This is the only α-amylase known to possess this N-terminal domain and it exhibits homology to the N-terminal SBD of *Rhizopus oryzae* glucoamylase. It shares no homology with the C-terminal starch-binding domains present in the cycloglucanotransferases, glucoamylases or α-amylases. The evolutionary tree based on the sequence alignment of SBDs reveals that the N-terminal SBDs are separated from the C-terminal SBDs.

Keywords: *Lipomyces kononenkoae* α-amylases, sequence analyses, substrate specificity

2. Introduction

Starch is a major renewable resource consisting of two types of molecules, amylose, which generally makes up 20-30%, and amylopectin, which generally constitutes 70-80%. Both these types of molecules consist of polymers of a-D-glucose units in the $^4C_1$ conformation. In amylose, these are linked α-(1 ,4), with the ring oxygen atoms all on the same side, whereas in amylopectin, about one residue in every twenty or so is also linked α-(1,6), forming branch points (IMMEL *et al.,* 2000). The use of amylolytic enzymes to break down starch to soluble sugars has become increasingly important in the production of various commercial products, such as fuel ethanol, alcoholic beverages and sweeteners. α-Amylases were originally recognised as a group of starch hydrolases and related enzymes that exhibit clear sequence similarities and a predicted common super secondary fold, a parallel $(\beta/\alpha)_8$ barrel (FARBER & PETSKO, 1990; FABER, 1993). The α-amylase family, GH-H clan of glucosyl hydrolases, is the largest family of glucoside hydrolases, transferases and isomerases, comprising over 30 different specificities. Currently, the α-amylase family constitutes the clan of GH-H families 13, 70 and 77, and the enzymes of these different families can operate on α-1,1-, α-1,2-, α-1,3- and α-1,5-linkages, as well as on α-1,4- and α-1,6-glucosidic linkages (VAN DER MAAREL *et al*., 2002). In nature, there are more than 150 starch-assimilating yeasts, but only a few of them secrete a combination of enzymes that can cleave the α-1,4 and α-1,6 linkages of the complex starch molecule. The ability of the yeast species *Lipomyces kononenkoae* to degrade raw starch completely has been well documented (SPENCER MARTINS & VAN UDEN, 1979; WILSON & INGLEDEW, 1982; DE MOT et al., 1984). SPENCER MARTINS & VAN UDEN (1979) evaluated 81 raw starch-assimilating yeasts, representing 59 species, and reported that the highest biomass production on starch media was obtained with strains of *Lipomyces kononenkoae* and *Lipomyces starkeyi.* These two yeast species secrete a set of starch-hydrolysing enzymes of which the genes encoding for some of these have been cloned and expressed in simpler systems, such as *Saccharomyces cerevisiae* and *Escherichia coli* (BIGNELL et al., 2000; KANG et al., 2004; STEYN & PRETORIUS, 1995; EKSTEEN et al., 2003a). The genes encoding two novel α-amylases, LKA1 and LKA2, were isolated from *L. kononenkoae* strain IGC4052B, cloned and expressed in *S. cerevisiae* (STEYN & PRETORIUS, 1995; EKSTEEN et al., 2003b). These amylases exhibit different substrate specificities towards starch and related substrates,

acting on their α-1,4 and α-1,6 linkages. In this current review, we summarise the sequence analyses and properties of *L. kononenkoae* α-amylases LKA1 and LKA2.

3. Amylolytic system of *Lipomyces kononenkoae*

All the starch-assimilating yeasts necessarily produce one or more extracellular amylolytic enzymes. The crude concentrates of culture fluids of *Lipomyces* strains are able to extensively hydrolyse a wide variety of polysaccharides, including starch, and two strains of *L. kononenkoae* were identified to produce a different extracellular amylolytic system capable of total starch hydrolysis (SPENCER MARTINS & VAN UDEN, 1979). However, due to their complex genetics, only a few genes encoding these enzymes have been identified and further characterised. The culture filtrate from *L. kononenkoae* strain CBS5608 consists of three different enzymes: α-amylase, glucoamylase, and a third enzyme with debranching activity on amylopectin and transferase activity hydrolysing α-1,6 chains in panose. *L. kononenkoae* strain IGC4051 was reported to secrete an extracellular isoamylase, which increases the amylolysis of amylopectin and glycogen, completely hydrolysing these substrates into maltose when combined with a β-amylase. This enzyme does not have any action on pullulan or dextrin (SPENCER MARTINS, 1982).

The strain *L. kononenkoae* IGC4052B, which is a catabolite-derepressed mutant of IGC4052, secretes two enzymes with activity on starch, namely a novel raw starch-degrading α-amylase, LKA1, and a second α-amylase, LKA2. LKA1 α-amylase acts on glucose polymers containing α-1,4 and α-1,6 bonds by endohydrolysis, producing maltose, maltotriose and maltotetraose. LKA2 exhibits properties of α-amylase, but also possesses side activity towards dextrin. The genes encoding these enzymes have been identified, cloned and expressed in *Saccharomyces cerevisiae* (STEYN & PRETORIUS, 1995; EKSTEEN et al., 2003a). The expression of LKA1 in *Saccharomyces cerevisiae* in conjunction with LKA2 results in higher ethanol yields as a result of the synergistic action of these enzymes (EKSTEEN et al., 2003b).

3.1 Biochemical properties of *L. kononenkoae* amylases

LKA1 α-amylase has a theoretical molecular mass of approximately 76 kD (STEYN & PRETORIUS, 1995). The enzyme showed properties of a glycoprotein, indicated by DIF-glycan detection. The temperature optimum of LKA1 ranged from 40 to 50°C (STEYN &

PRETORIUS, 1995), and this is within the range of optimum temperatures reported for other yeast α-amylases (DE MOT, 1990; IEFUJI et al., 1996). The stability of the enzyme is maintained between 20 and 50°C, however, and it retains 83% of its activity at 55°C. The optimal pH range of 3-8 was similar to the values reported for *L. starkeyi* α-amylase (KELLY et al., 1985) and other yeast amylases (DE MOT, 1990).

LKA1 α-Amylase activity was partially inhibited by Cu^{2+} and Zn^{2+}, whereas other metal ions, such as Co^{2+} and Mg^{2+}, had only a slight stimulating effect. Calcium, which stabilises most α-amylases, had no effect on the activity of LKA1. Enzymatic activity was not affected by EDTA or EGTA, but was stimulated by SDS and low concentrations of urea, and was sensitive to acetone.

The second α-amylase expressed in *Saccharomyces cerevisiae*, LKA2, had a temperature optimum of 60°C and maintained its maximum activity in the temperature range of 50 to 65°C. The optimum pH for activity was estimated to be 3.5. The molecular weight of this protein was around 55 kD and, similar to the LKA1 α-amylase, neither Ca^{2+} nor EDTA had any effect on enzymatic activity. (STEYN & PRETORIUS, 1995, EKSTEEN *et al*., 2003a)

LKA1 exhibited specificity towards α-1,4 and α-1,6 linkages and large amounts of reducing sugars were liberated from the highly branched amylopectin and glycogen and amylose. The k_m values for lintner starch and amylopectin were 2.31 and 1.97 mg/ml respectively. LKA1 also exhibited a side activity on pullulan. LKA2, however, showed high reactivity towards lintner starch, dextrin and amylose, but only small amounts of reducing sugars were liberated from branched substrates such as amylopectin and glycogen (EKSTEEN et al., 2003a). The relative activity of the LKA1 and LKA2 enzymes on different glucose polymers is summarised in Table 1.

3.2 Sequence alignment of *L. kononenkoae* amylases

Clustal W analyses of LKA1 and LKA2 reveals homology to various yeasts, plant α-amylases, α-amylases of *Aspergillus nidulans*, *Debaryomyces occidentalis*, *Saccharomycopsis fibuligera* and *S. pombe*, cyclodextrin glucanotranferases, pullulanases, α-glucosidase and β-amylases from *Bacillus polymyxa* (KANEKO et al., 1998; KAWAZU et al., 1987; TAKIZAWA & MUROOKA, 1985). The sequence comparison of α-amylases from different yeasts is presented in Fig 2 (EKSTEEN et al., 2003a). LKA1 α-amylase has a high homology of around 78% and LKA2 has a homology of about 36%

homology to the α-amylase from *L. starkeyi* (KANG et al., 2004). The 3D structure of *Aspergillus oryzae* TAKA amylase (MATSURA et al., 1984) was used as a reference model for the structure interpretation of both the α-amylases, LKA1 and LKA2. The primary sequence analyses of LKA1 and LKA2 with TAKA amylase revealed corresponding similarities for the catalytic domain and the C-terminal stabilising domain. The sequence comparison suggested that the amino acid sequence of LKA1 and LKA2 included the four conserved regions comprising the strands α_3, α_4, α_5 and α_7, containing the catalytically important residues His122, Asp206, Glu230, and Asp297 respectively (HENRISSAT, 1991; TAKATA et al., 1992; SVENSSON, 1994), as previously identified in the α-amylase family of enzymes (KURIKI, 1991). Further, six amino acid residues, which are strictly conserved in the four conserved regions of the α-amylase family, are also completely preserved in LKA1 and LKA2. It has been suggested earlier that conservations in the catalytic region of α-amylases might be related to the maintenance of structure and variations within the segments may be characteristic of certain enzyme specificities (JANEČEK, 1997).

3.3 Variations in the conserved segments of the catalytic domain

The sequence analyses showed some amino acid residues that are variant in the conserved segments of *L. kononenkoae* amylases, as shown in Table 2. Region IV of LKA1, LKA2 and LSA contains an uncharged Gln in the place of the highly conserved, charged His, which is involved in substrate binding in other α-amylase members. The importance of this invariant His296 has been well demonstrated through structural, kinetic and mutagenesis approaches in other α-amylase members. The mutation of His296 to Asn led to a 100-fold decrease in the K_{cat} values in barley α-amylase (SOGAARD et al., 1993); a similar mutation led to a loss of 50% of the original activity in cyclodextrin glucanotransferase from *Bacillus*, and similar results have been observed in human pancreatic α-amylase (ISHIKAWA et al., 1993). In *Bacillus stearothermophilus* neopullulanase, the mutation of the corresponding His to Glu led to a 70% decrease in its activity towards α-1,4 linkages, leading to a lower maltose release and higher panose production (KURIKI et al., 1991). Inhibitor-binding studies conducted on barley alpha amylase suggest that the imidazole His residue forms a hydrogen bond with the OH of the glycone residue at the bond cleavage site. Thus, mutation of this His296 greatly diminishes the ability of the enzyme to stabilise the transition state, as indicated by the

decreases in K_{cat}/K_m for the substrates used. These research findings reiterate the importance of the invariant His296 in this family of enzymes.

The uncharged Gln in the place of His296 at he fourth conserved segment is unique to the *Lipomyces* α-amylases cloned and characterised thus far. This is clear from the sequence comparison of these enzymes with other members of the α-amylase family. The significance of Gln at this position in LKA1 α-amylase is currently under investigation (RAMACHANDRAN et al., UNPUBLISHED DATA) ; understanding the role of variant amino acids in such highly conserved regions of these enzymes might broaden our knowledge on the characteristic properties of individual enzymes. The significance of variations in the *LKA2*-encoded α-amylase 'L' in region 2, 'A' and 'N' in region 3, and 'D' and 'Q' in region 4, is yet to be explored. In addition, LKA2 α-amylase consists of an Arg residue in the place of the Lys121 (TAA numbering), which is well conserved in the α-amylase family. Previous reports on mutational analyses of Lys to Arg/Asn in *Saccharomycopsis fibuligera* showed a three-fold increase in activity towards shorter substrates (MATSUI et al., 1992).

3.4 Starch binding and raw starch hydrolysis

In some distinct members, the catalytic domain is preceded by an extra sequence – a domain whose role is uncertain (MAC GREGOR et al., 2001). The N-terminal region of the LKA1 α-amylase preceding the catalytic domain consists of an extension of 132 amino acids that exhibits significant homology to the N-terminal regions of *L. starkeyi* α-amylase, and a 36% similarity to the glucoamylases from *Rhizopus oryzae* and *Arxula adeninivorans* (Fig1). Predictions of a hypothetical secondary structure using the predict protein tool (Expasy predict protein software) suggest the presence of six β strands. This region in *R. oryzae* has been reported to be the starch-binding domain (ASHIKARI et al., 1986). Deletion of the N-terminal region (1-132) in LKA1 resulted in a loss of the binding function of this enzyme to raw starch (RAMACHANDRAN et al., MANUSCRIPT SUBMITTED). It is also known, that the presence of starch binding domain may confer thermostability properties to enzymes (IEFUJI et al., 1996). LKA2 α-amylase, on the other hand, possesses a short extension of 19 amino acids, which does not possess any significant homology to any known α-amylase sequences. Whether this region functions as an extension of the catalytic domain or is involved in any protein function is yet to be understood.

On the basis of the knowledge of different starch-binding domains present in this family, they have been widely classified into the carbohydrate-binding module families 20, 21, 25 and 26 (COUTINHO & HENRISSAT, 1999). The starch-binding domains present at the C-terminal end of the enzymes have been classified under Carbohydrate-Binding Module Family 20, along with the SBDs from cyclomaltodextrin glucanotransferases (EC 2.4.1.19, CGTases, Domain E), α-amylases (EC 3.2.1.1), β-amylases (EC 3.2.1.2), maltotetraose-forming exo-amylases (EC 3.2.1.60), maltogenic α-amylases (EC 3.2.1.133) and other hydrolases. The tertiary structures of nine of the family 20 SBDs are known thus far (CORNETT et al., 2003). The *Rhizopus oryzae* and *Arxula adeninivorans* SBDs are currently classified as Carbohydrate-Binding Module Family 21 . The sequence homology of LKA1 to the other N-terminal SBDs is shown in Fig. 1. However, no three-dimensional structures are available for this family of SBDs, which limits the knowledge of the interactions of the SBD with the other domains on a structural level. It is evident, however, that the N-terminal SBDs from *R. oryzae* glucoamylase and LKA1 α-amylase are very distantly related to all the other known C-terminal SBDs. The differences in the sequences of the C- and N-terminal SBDs belonging to Carbohydrate-Binding Module Families 20 and 21 respectively, may reflect the possibility that, during their evolution, the *Rhizopus* and *Aspergillus* glucoamylases obtained their SBDs independently (TANAKA et al., 1986). While most amylases possess their binding domains at the C terminus, it can be speculated that multiple domain organisations existed in ancestral fungi and that these enzymes did not follow phylogenetic lines. This might explain why *R. oryzae* and *A. adeninivorans* possess a putatively newer type of SBD at the N terminus, even though Zygomycetes like *Rhizopus* diverged from *Ascomycetes* before the appearance of yeasts (COUTINHO & REILLY, 1997). Moreover, the presence of SBD in an amylolytic enzyme is closely connected to the enzyme origin (JANEČEK et al., 2003). The presence of the N-terminal region in the *LKA1*-encoded α-amylase identical to the SBD of *Rhizopus oryzae* glucoamylases might suggest the independent evolutionary behaviour of starch-binding domains in glycosyl hydrolases. BORK et al. (1998) further reported alignment of the starch-binding domains of bacterial and fungal amylases and a bacterial glucanotransferase with some mammalian glycogen-targeting subunits. Strikingly, the amino acids Gly146, Asn152, Lys157, Asp189, Phe193, Tyr213 and Asn224 of these mammalian polysaccharide-binding domains are well conserved only in the N-terminal SBD of *R. oryzae* and LKA1. No similarity was noted in any of the C-terminal starch-binding domains examined.

The differences between the C-terminal SBDs and the N-terminal SBDs are clearly reflected in their evolutionary trees (Fig. 2), which were calculated using the amino acid sequences (Table. 4) according to JANECEK et al., 2003; in the current study we have included the *Lipomyces* amylases and enzymes with N terminal starch binding domains to this evolutionary tree. The C-terminal SBDs of α-amylases and glucoamylases cluster together, while the CGTases are on the branch adjacent to the α-amylase-GA group. The *Lipomyces* enzymes, LKA1 and LSA α-amylases are on the same branch as that of the *R. oryzae* glucoamylase starch-binding domain. The fact that the SBD occurs in GH13, GH14 and GH15 (JANECEK & SEVCIK, 1999) supports the idea that there has been a separate evolution of this domain. The different organisations found in the amylase family (JESPERSEN et al., 2001; JUGE et al., 2002) with regard to the starch-binding domains is represented in Fig. 3.

4. Future prospects

The *L. konoenkoae* α-amylases and their unique specificity properties support the rationale for structure-function relationship studies in the α-amylase family. The unique modular organisation of LKA1 also invites the construction of novel enzymes using *L. kononenkoae* α-amylases with advantageous combinations of properties such as raw starch adsorption and activity towards a wide range of substrates. Further such studies would provide insight into various structure-function implications and would assist in unravelling how the extra domains interact with the catalytic domain or how the variant amino acids influence catalysis and specificity towards different substrates. Among the major questions that need to be addressed in relation to the basic knowledge about these proteins are the positioning of the N-terminal SBD, insight into cooperation between the SBD and the catalytic domain, and the investigation of the reaction mechanism and substrate specificity.

5. Acknowledgements

The authors would like to thank the National Research Foundation, South Africa for funding.

6. References

ASHIKARI, T., NAKAMURA, N., TANAKA, Y., KIUCHI, N., SHIBONO, Y., TANAKA, T., AMACHI, T. & YOSHIZUMI, H. 1986. Agric. Biol. Chem. **50**: 957-964.

BAHL, H., BURCHHARDT, G., SPREINAT, A., HAECKEL, K., WIENECKE, A., SCHMIDT, B. & ANTRANIKIAN, G. 1991. Appl. Environ. Microbiol. **57**:1554–1559.

BIGNELL, G.R., BRUCE, I.J & EVANS, I.H. 2000. Biotechnol. Lett. **22**: 1713-1718.

BOEL, E., HANSEN, M.T., HJORT, I., HOEGH, I & FIIL, N.P.1984.EMBO J. **3**:1581-1585.

BORK, P., DANDEKAR, T., EISENHABER, F. & HUYNEN, M. 1998. J. Mol. Med. **76**: 77-79.

BUI MINH, D., KUNZE, I., FOERSTER, S., WARTMANN, T., HORSTMANN, C., MANTEUFFEL, R & KUNZE, G.1996. Appl. Microbiol. Biotechnol. **44**:610-619.

CORNETT, C.A.G., FANG, T.Y., REILLY, P.J. & FORD, C. 2003. Protein Eng. **16**: 521-529.

COUTINHO, P.M. & HENRISSAT, B. 1999. pp. 15-23. In: OHMIYA, K., HAYASHI, K., SAKKA, K., KARITA, S. & KIMURA, T. (eds) Genetics, Uni Publishers Co, Tokyo.
COUTINHO, P.M. & REILLY, P.J. 1997. Proteins: Struc. Func. Gen. **29**: 334-347.

DE MOT, R., VAN OUDENDIJICK, E. & VERACHTERT, H. 1984. Biotechnol. Lett. **6**: 581-586.

DE MOT, R. & VERACHTERT, H. 1987. Eur. J. Biochem. **164**: 643-654.

DE MOT, R. 1990. pp. 163-222. In: VERACHTERT, H. & DE MOT, R. (eds) Marcel Dekker, New York.

DIDERICHSEN, B. & CHRISTIANSEN, L. 1988. FEMS Microbiol. Lett. **56**: 53–60.

EKSTEEN, J.M., STEYN, A.J.C., VAN RENSBURG, P., CORDERO OTERO, R.R. & PRETORIUS, I.S. 2003a. Yeast **20**: 69-78.

EKSTEEN, J.M., VAN RENSBURG, P., CORDERO OTERO, R.R & PRETORIUS, I.S. 2003b. Biotechnol. Bioeng. **84**: 639-646.

FARBER, G.K. & PETSKO, G.A. 1990. Trends. Biochem. Sci. **15**: 228-234.

FARBER, G.K. 1993. Curr. Opin. Struct. Biol. **3**: 409-412.

FUJIWARA, S., KANEMOTO, M., KIM, B., LEJEUNE, A., SAKAGUCHI, K. & IMANAKA, T. 1992. Appl. Environ. Microbiol. **58**: 4016–4025.

HATA, Y., KITAMOTO, K., GOMI, K., KUMAGAI, C., TAMURA, G & HARA, S. 1991. Agric. Biol. Chem. **55**:941-949.

HAYASHIDA, S., KURODA, K., OHTA, K., KUHARA, S., FUKUDA, K & SAKAKI, Y. 1989.Agric. Biol. Chem. **53**:923-929.

HENRISSAT, B.1991. Biochem. J. **280**: 309-316.

HILL, D.E., ALDAPE, R. & ROZZELL, J.D. 1990. Nucleic Acids Res. **18**: 199

IEFUJI, H., CHINO, M., KATO, M. & LIMURA, Y. 1996. Biochem. J. **318**: 989-996.

IMMEL, S. & LICHTENTHALER, F. W. 2000. Stärke. **52:**1-8.

ISHIKAWA, K., MATSUI, I., KOBAYASHI, S., NAKATANI, H. & HONDA, K. 1993. Biochemistry **32**: 6259-6365.

JANEČEK, S. 1997. Prog. Biophys. Molec. Biol. **67**: 67-69.

JANECEK, S., SVENSSON, B. & MACGREGOR, E.A. 2003. Eur. J. Biochem. **270**:635-645.

JANECEK,S & SEVICEK, J.1999. FEBS Lett.**456**: 119-125

JESPERSEN, H.M., MACGREGOR, E.A., SIERKS, M.R. & SVENSSON, B. 1991. Biochem. J. **280**:51-55.

JOUTSJOKI, V.V & TORKKELI, T.K. 1992. FEMS Microbiol. Lett. **78**:237-243.

JUGE, N., LE GAL-COEFFET, M., CAROLINE, S.M., FURNISS, A., GUNNING, P., KRAMHOEFT, B., MORRIS, V.J., WILLIAMSON, G., SVENSSON, B. 2002. Biologia, Bratislava. **57**: 239-245

KANEKO, A., SUDO, S., SAKAMOTO, Y., TAMURA, G., ISHIKAWA, T. & OHBA, T. 1996. J. Ferment. Bioeng. **81**: 292–298.

KANEKO, T., HAMAMOTO, T. & HORIKOSHI, K. 1988. J. General Microbiol. **134**: 97–105.

KANEKO, T., SONG, K.B., HAMAMOTO, T., KUDO, T. & HORIKOSHI, K. 1989. J. General Microbiol. **135**: 3447–3457.

KANG, H.K., LEE, J.H., KIM, D., DAY, D.F., ROBYT, J.F., PARK, K.H., MOON, T. 2004. FEMS Microbiol. Lett. **233**: 53-64.

KAWAZU, T., NAKANISHI, Y., UOZUMI, N., SASAKI, T., YAMAGATA, H., TSUKAGOSHI, N. & UDAKA, S. 1987. J. Bacteriol. **169**: 1564-1570.

KELLY, C.T., MORIARTY, M.E. & FOGARTY, W.M. 1985. Appl. Microbial. Biotechnol. **22**: 352-358.

KURIKI, T., TAKATA, H., OKADA, S. & IMANAKA, T. 1991. J. Bact. **173**: 6147-6152.

LONG, C.M., VIROLLE, M.J., CHANG, S.Y., CHANG, S. & BIBB, M.J. 1987. J. Bacteriol. **169**: 5745–5754.

MACGREGOR, E.A., JANECEK, S., SVENSSON, B. 2001. Biochim. Biophys. Acta. **1546**:1-20.

MATSUI, I., ISHIKAWA, K., MIYARI, S., FUKUI, S. & HONDA, K. 1992. FEBS Lett. **310**: 216-218.

MATSUURA, Y., KUSUNOKI, W., HARADA, W. & KAKUDO, M. 1984. J Biochem. **95**: 697-702.

NAGASAKA, Y., MURAKI, N., KIMURA, A., SUTO, M., YOKOTA, A & TOMITA F. 1995. Appl. Microbiol. Biotechnol. **44**:451-458.

ONISHI, H. & HIDAKA, O. 1978. Can. J. Microbial. **24**: 1017-1023.

PETRICEK, M., TICHY, P. & KUNCOVA, M. 1992. Gene **112**:77–83.

SCHMID, G., ENGLBRECHT, A. & SCHMID, D. 1988 pp. 71–76 In : HUBER, O. & SZEJTLI, J. (eds),. Kluwer Academic Publishers, Dordrecht, Germany/Boston, USA.

SHIDA, O., TAKANO, T., TAKAGI, H., KADOWAKI, K. & KOBAYASHI, S. 1992 Biosci. Biotechn Biochem. **56**: 76–80

SOGAARD, M., KADIZOLA, A., HASER, R. & SVENSSON, B. 1993. J. Biol. Chem. **268**: 1391-1396.

SPENCER MARTINS, I. & VAN UDEN, N.1979. European J. Appl. Microbiol. **6**: 241-250.

SPENCER MARTINS, I.1982. Appl. Environ. Microbiol. **4**: 29-35.

STEYN, A.J.C. & PRETORIUS, I.S. 1995. Curr. Genet. **28**: 526-533.

STONE, P.J., MAKOFF, A.J., PARISH, J.H & RADFORD, A.1993. Curr. Genet. **24**:205-211.

SVENSSON, B. 1994. Plant Mol. Biol. **25**: 141-157.

TAKATA, H., KURIKI, T., OKADA, S., TAKESADA, Y., LIZUKA, M., MINAMIURA, N. & IMANAKA, T. 1992. J. Biol. Chem. **267**: 18447-18452.

TAKIZAWA, N. & MUROOKA, Y. 1985. **49**: 294-298.

TANAKA, Y., ASHIKARI, T., NAKAMURA, N., KIUCHI, N., SHIBANO, Y., AMACHI, T. & YOSHIZUMI, H. 1986. Agric. Biol. Chem. **50**: 965-969.

VAN DER MAAREL, M.J., VAN DER VEEN, B., UITDEHAAG, J.C., LEEMHUIS, H & DIJKHUIZEN, L. 2002. J Biotechnol. **94(2)**:137-155.

VIGAL, T., GIL, J.A., DAZA, A., GARCIA-GONZALEZ, M.D. & MARTIN, J.F. 1991. Mol. General Genet. **225**:278–288.

WILSON, J.J. & INGLEDEW, W.M.1982. Appl. Environ. Microbiol. **44**: 301-307.

YAMAMOTO, T., SHIRAKI, K., FUJIWARA, S., TAKAGI, M., FUKUI, K. & IMANAKA, T. 1999. Biochem. Biophys. Res. Commun. **265**: 57–6177.
ZHOU, J., BABA, T., TAKANO, T., KOBAYASHI, S. & ARAI, Y. 1989. FEBS Lett. **255**: 37–41.

Table 1. Activity of *Lipomyces kononenkoae* α-amylases on starch and related substrates (STEYN & PRETORIUS, 1995; EKSTEEN et al., 2003a).

Glucose polymer	Relative enzyme activity (%)	
	LKA1	LKA2
Amylopectin (potato)	100	17
Starch (Lintner)	95	100
Glycogen (oyster)	75	9
Dextrin	65	91
Pullulan	2	2

Table 2. Variations between and similarity within the conserved regions of the catalytic domain of *Lipomyces kononenkoae* amylases.

Enzymes	Source	Region 1	Region 2	Region 3	Region 4
α-Amylase	*Aspergillus oryzae*	117 DVVANH	202 GLRIDTVKH	230 EVLD	292 FVEN**H**D
Pullulanase	*Klebsiella aerogenes*	600 DVVYNH	671 GFRFDLMGY	704 EGWD	827 YVSK**H**D
Neopullulanase	*Bacillus stearothermophilus*	242 DAVFNH	324 GWRLDVANE	357 EIWH	419 LLGS**H**D
α-Amylase-pullulanase	*Clostridium thermohydrosulfulcum*	488 DGVFNH	594 GWRLDVANE	627 ENWN	699 LLGS**H**D
CGTase	*Bacillus macerans*	135 DFAPNH	225 **GIRFD**AVKH	258 EWFL	324 FIDN**H**D
Amylomaltase	*Streptococcus pneumoniae*	224 DMWAND	291 IVRIDHFRG	332 EELG	391 YTGT**H**D
Isoamylase	*Pseudomonas amylodermosa*	292 DVVYNH	371 GFRFDLASV	435 EPWA	505 FIDV**H**D
LKA1 α-amylase	*Lipomyces kononenkoae*	264 DIVVNH	348 GLRIDTVKHV	376 EVFD	439 FLEN**Q**D
LKA2 α-amylase	*Lipomyces kononenkoae*	137 DVVINH	223 **GIRLD**TA**R**HV	251 E**ALN**	312 FLDN**Q**D
LSA	*Lipomyces starkeyi*	287 DIVVNH	371 GLRIDTVKH	400 EVFD	462 FLENQD

Table 3. Protein sequences used for comparative analyses and construction of evolutionary tree (JANECEK et al., 2003).

Enzyme	Source	Abbr	Swissprot	Reference
C-terminal SBDs				
α-Amylase	*Aspergillus nidulans*	Aspnd	Q9UV09	Unpublished
	Aspergillus kawachii	Aspka	P13296	KANEKO et al.,1996
	Cryptococcus sp. S2	Crcsp	Q92394	IEFUJI et al., 1996
	Streptomyces albidoflavus	Stral	P09794	LONG et al., 1987
	Streptomyces griseus	Strgr	P30270	VIGAL et al.,1991
	Streptomyces lividans TK21	Strli21	O86876	Unpublished
	Thermomonospora curvata	Thscu	P29750	PETRICEK et al.,1992
Maltotetraohydrolase	*Pseudomonas saccharophila*	Psesa	P22963	ZHOU et al., 1989
Maltopentao-hydrolase	*Pseudomonas* sp.KO-8940	Psesp	Q52516	SHIDA et al.,1992
Maltogenic amylase	*Bacillus stearothermophilus*	Bacst	P19531	DIDERICHSEN et al., 1988
CGTase	*Bacillus* sp. 1-1	Bac11	P31746	SCHMID et al., 1988
	Bacillus sp. 17-1	Bac17	P30921	KANEKO et al., 1989
	Bacillus sp.38-2	Bac38	P09121	KANEKO et al., 1988
	Bacillus sp.6.6.3	Bac663	P31747	Unpublished
	Bacillus licheniformis	Bacli	P14014	HILL et al., 1990
	Bacillus stearothermophilus No 2	Bacst2	P31797	FUJIWARA et al., 1992
	Thermoanaerobacter thermosulfurogenes	Thbth	P26827	BAHL et al., 1991
	Thermococcus sp. B1001	Thcsp	Q9UWN2	YAMOMOTO et al., 1989
Glucoamylases	*Aspergillus* X-100	Ax	Q12537	Unpublished
	Aspergillus kawachi	Ak	P23176	HAYASHIDA et al., 1989
	Aspergillus niger	An	P04064	BOEL et al., 1984
	Aspergillus oryzae	Ao	P36914	HATA et al., 1991
	Neurospora crassa	Nc	P14804	STONE et al., 1993
	Humicola grisea	Hg	Q12623	Unpublished
	Hormoconis resinae	Ho	Q03045	JOUTSJOKI et al.,1992
	Corticium rolfsii	Cr	Q12596	NAGASAKA et al., 1995
N-terminal SBDs				
α-Amylases	*Lipomyces kononenkoae*	LKA1	Q01117	STEYN et al., 1995
	Lipomyces starkeyi	LSA	Q6YF33	KANG et al., 2004
Glucoamylases	*Arxula adeninivorans*	Aa	P42042	BUI et al., 1996
	Rhizopus oryzae	Rg	P07683	ASHIKARI et al., 1986

Fig. 1. Sequence alignment of the N-terminal starch-binding domains of LKA1 (*Lipomyces kononenkoae* α-amylase) (STEYN et al., 1995) and Rg (*Rhizopus oryzae* glucoamylase) (ASHIKARI et al., 1986) with the N-terminal regions of Aa (*Arxula adeninivorans* glucoamylase)(BUI et al., 1996) and LSA (*Lipomyces starkeyi* α-amylase)(KANG et al., 2004).

```
AMYA    21  AEWRSQSIYF LLTDRFARTD NSTTAECDTS .AKYCGGTWQ GIINQLDYIQ   70
ALP1    31  DKWRSQSIYQ IVTDRFARTD GDTSASCNTE DRLYCGGSFQ GIIKKLDYIK   81
SWA2    36  AAWRSESIYQ LVTDRFARTD GSTSATCNTG DRVYCGGTFQ GIIDKLDYIQ   86
LKA1   151  ANWRGRSIYQ VVTDRFARTD GSITYSCDVT DRVYCGGSYR GIINMLDYIQ  200
AMY2    29  HGWRKQSIYS LLTDRFASTN ...PKPCNPE DREYCGGNWR GIIDKLDYIQ   76
LKA2    27  AEWKELSIYQ VITDRFATTN ...LTAPDCW IRAYCGGTWK GLERKLDYIQ   74

AMYA    71  GMGFTAIWIT PVTANLEDGQ .HGEAYHGYW QQDIYALNPH FGTQDDLRAL  120
ALP1    82  DMGFTAIWIS PVVENIPDNT AYGYAYHGYW MKNIYKINEN FGTADDLKSL  132
SWA2    87  GMGFTAIWIS PVVEQIPDDT GYGYAYHGYW MKDIYAINSN FGTADDLKNL  137
LKA1   201  GMGFTAIWIS PIVENIPDDT GYGYAYHGYW MKDIFALNTN FGGADDLIAL  251
AMY2    77  GMGFTAIWIS PIIKNIEGRT KYGEAYHGYW PQDLYTLNPH FGTEQDLIDL  127
LKA2    75  NMGFDAVWIS PVIHNIEVNT TWGFAFHGYW GDDPYRLNEH FGTAADLKSL  125
                                   ** **
                        A
AMYA   121  SDALHDRGMY LMVDVVANHF GYDAPAASVD YSAFNP.FNS ADYFHTPCDI  170
ALP1   133  AQELHDRDML LMVDIVTNHY GSDGSGDSID YSEYTP.FND QKYFHNYCLI  182
SWA2   138  SNELHKRNMK LMVDIVTNHY AWNGAGSSVA YSNYNP.FNQ QSYFHDYCLI  187
LKA1   252  ATELHNRGMY LMVDIVVNHF AFSGNHADVD YSEYFP.YSS QDYFHSFCWI  301
AMY2   128  ADALHDRGMY LMVDTVVNHM GSSDPR.NID YGIYRP.FNQ SSHYHPMCPI  176
LKA2   126  SDSLHARGMS LMVDVVINHL ASYTLPQDVD YSLYPAPFNT SSAFHQPCPI  176
                         °   °*

AMYA   171  TDYDNQTQVE DCWLYTD.AV SLPDVDTTNE EVKEIWYDWV GDLVSDYSID  220
ALP1   183  SNYDDQAQVQ SCWEGDS.SV ALPDLRTEDS DVASVFNSWV KDFVGNYSID  232
SWA2   188  TNYDDQTNVE DCWEGDN.TV SLPDLRTEDS DVSSIFNLWV AELVSNYSID  237
LKA1   302  TDYSNQTNVE ECWLGDD.SV PLVDVNTQLD TVKSEYQSWV KQLIANYSID  351
AMY2   177  E.QDKPLSLE QCWIGTE.DM TLPDIDTENP QIIETLYNFI HDQVKQFKID  225
LKA2   177  D.FSNQSSIE DCWLVTEPAP ALVDLKNEDQ VILDALINSV VDLVETYDID  226
              *             °          * °
            B                        B'
AMYA   221  GLRIDTARHV QKDFWRDYND AAGVYCVGEV FQGDPDYTCG YQEVMDGVLN  271
ALP1   233  GLRIDSAKHV DQGFFPDFVS ASGVYSVGEV FQGDPAYTCP YQNYIPGVSN  283
SWA2   238  GLRIDSAKHV DESFYPSFQS AAGVYLLGEV YDGDPAYTCP YQNYMSGVTN  288
LKA1   352  GLRIDTVKHV QMDFWAPFQE AAGIYTVGEV FDGDPSYTCP YQENLDGVLN  402
AMY2   226  GLRVDATKHV RRTFWPGFCE SAGVYCQGEE WTGQADLFCE WQEYMDGLHN  276
LKA2   227  GIRLDTARHV PKPSLAKFQE KVGVFVTGEA LNQSVPYVAQ YQGPLNSAIN  277
               ▲  *○                    *▲* *
                                                      C
AMYA   272  YPIYYPLLRA FSSTSG.SLS DLANMIETVK YTCSDATLLG NFIENHDNPR  321
ALP1   284  YPLYYPTTRF FKTTDS.SSS ELTQMISSVA SSCSDPTLLT NFVENHDNER  333
SWA2   289  YPLYYPMLRF FQGTSN.SVD ELNAMISSLE SDCKDITLLG NFIENHDQPR  338
LKA1   403  YPVYYPVVSA FQRVGG.SIS SLVDMIDTLK SECIDTTLLG SFLENQDNPR  452
AMY2   277  FPVQGVAAES VIPLNDRALR KTAIAMNLVA HHCKDSTLLG LFLESQDAPR  327
LKA2   278  YPLWYALVDS FMGRTT..FD YLESVVKSEQ ATFSDAHALT NFLDNQDQPR  326
            *                                                * *▲

AMYA   322  FASYTDDIS. ....LAKNVA AFVILSDGIP IIYAGQEQHY SGAGDPANRE  367
ALP1   334  FASMTSDQS. ....LISNAI AFVLLGDGIP VIYYGQEQGL SGKSDPNNRE  379
SWA2   339  LPSYTSDSA. ....LIKNAI AFNLMSDGIP IIYYGQEQGY SGSSDPNNRE  384
LKA1   453  FPSYTSDES. ....LIKNAI AFTILSDGIP IIYYGQEQGL NGGNDPYNRE  498
AMY2   328  LAALNNDYT. ....VLKNAM TLNLMSDGIP IVFYGQEQMF NGSHDPVNRP  373
LKA2   327  FASYLGDGNG DDVLRDENAA TFLFFVSGIP VIYYGFEQRF DGGFDPVNRE  377
                                                        *   *
```

```
AMYA  368  ATWLSGYDST SELYQFISKT NQIRNHAIWQ NET..YLSYK NYAIYNENNV  416
ALP1  380  ALWLSGYNKE SDYYKLIAKA NAARNAAVYQ DSS..YATSQ LSVIFSNDHV  428
SWA2  385  ALWLSGYSTS NGYYKLISSV NQIRNQAIYK DSK..YTTYW SDVLYASGHV  433
LKA1  499  ALWPTGYSTT STFYEYIASL NQIRNHAIYI DDT..YLTYQ NWVIYSDSTT  547
AMY2  374  ALWDQGYNTD GPLYQYTSKV NKIRRDLINS EDGEIYIRSI THAIMIGDHV  424
LKA2  378  PMWTSGYNTS TPLYNYLARL NAIRKYAASI TGTQVFYSDD TVFLGSGVSH  428

AMYA  417  LAMRKGFDGS QIITILTNAG ADAGS.STVS VPNTGFTAGA AVTEIYTCED  466
ALP1  429  IATKRGS... .VVSVFNNLG SSGSS..DVT ISNTGYSSGE DLVEVLTCST  474
SWA2  434  IALQRGADDQ RIVSVFNNLG SSGS...QTV TFSTKYSGGE KVVDVLTCQT  481
LKA1  548  IAMRKGFTGN QIITVLSNLG SSGSS.YTLT LSNTGYTASS VVYEILTCTA  597
AMY2  425  MVMYKGP... .VITFITNYG AVDK...EYL IK...MPGSE TMIDLLTCTL  465
LKA2  429  MAMQRGP... .LVIVLTNVG QHIIDNTGYT VTGSQFSAGD SLTDLVSCTK  475

AMYA  467  ITVSGSG.EV SVPMESGLPR VLYPKAKLEG SGICGL                 502
ALP1  475  VSGSSD...L QVSIQGGQPQ IFVPAKYASD ICS                    505
SWA2  482  SYANSDS.TL TVSISGGAPR IYAPASLIAN SGICNF                 517
LKA1  598  VTVDLSG.NL AVPMSGGLPR VFYPESQLVG SGICSM                 633
AMY2  466  IEVEGEV..M RTSIKKGEPK ILYPYQLAFR DGFCQEQITL QEIDDVFMGR  514
LKA2  476  VKVVGANGTF TSPSNGGKAR IWIKSKYAGK FCS                    509
```

Fig. 2. Alignment of the amino acid sequences of several amylases. The sequences of the mature proteins from *Aspergillus nidulans* (AMYA), *Saccharomycopsis fibuligera* (ALP1), *Debaromyces occidentalis* (SWA2), *Lipomyces kononenkoae* (LKA1), *Schizosaccharomyces pombe* (AMY2) and *Lipomyces kononenkoae* (LKA2) have been aligned by introducing gaps (.) to maximize the similarity. Bolds letters represent identical amino acids or conservative replacements. ▲, residues implicated in catalysis; *, residues for substrate binding; ○, residues for Ca++ binding (Eksteen et al., 2003a).

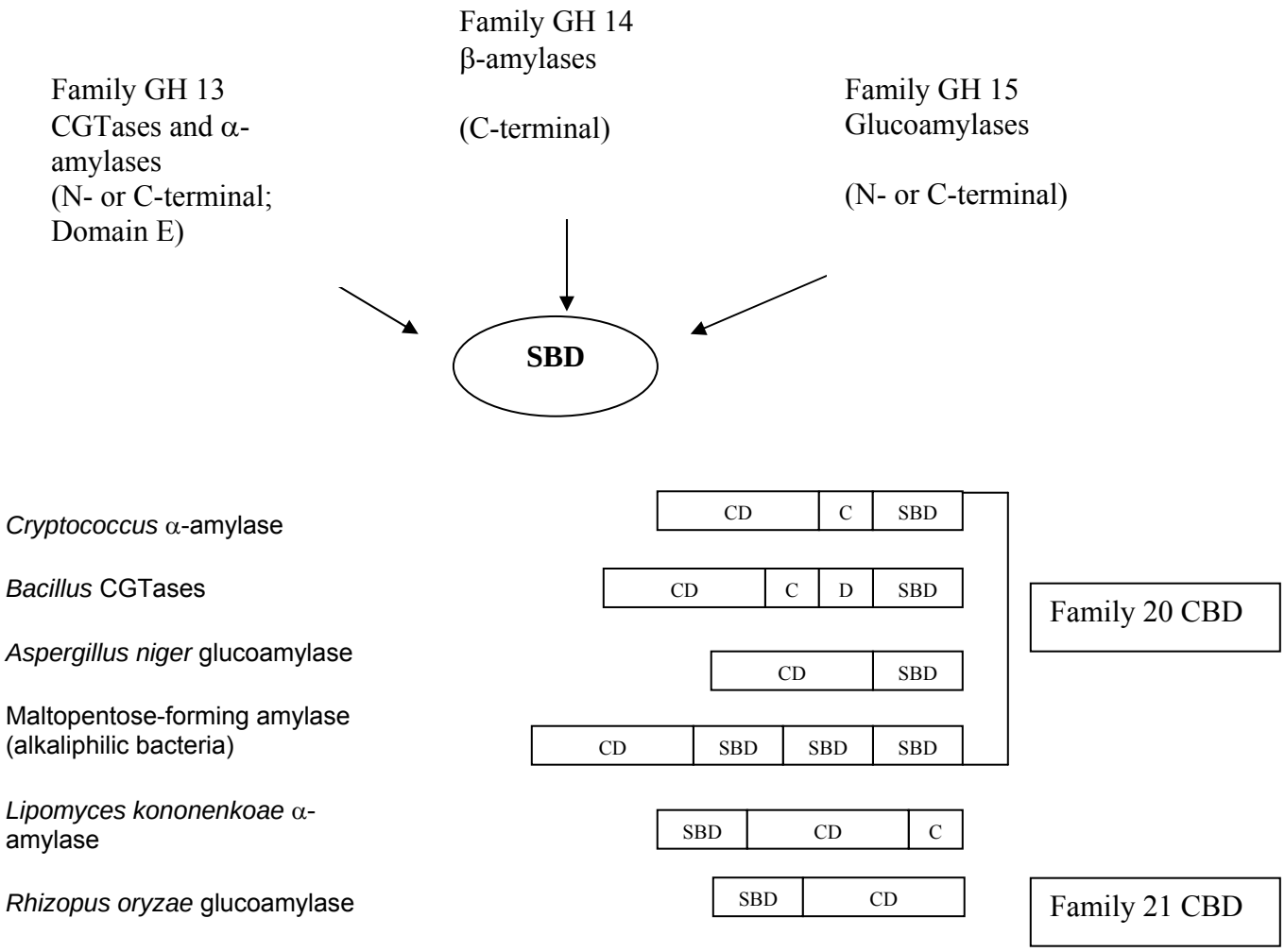


Fig. 3. Organization of starch binding domains in glucosyl hydrolase members CD-catalytic domain; SBD-starch binding domain; C-domain C; D-domain D; CBD-carbohydrate binding domain.

Fig. 4. Evolutionary tree of starch-binding domains from α-amylases, glucoamylases and CGTases.