STUDY OF CELLULOLYTIC AND HEMICELLULOLYTIC ENZYMES FROM A NOVEL ALKALOTHERMOPHILIC *THERMOMONOSPORA* SP.

A Thesis Submitted
To
University Of Pune

For
The Degree of Doctor of Philosophy
In
Microbiology

By

SHARMILI S. JAGTAP

Division of Biochemical Sciences

National Chemical Laboratory

PUNE -411 008

INDIA

NOVEMBER 2006

DEDICATED TO

MY FAMILY

“The whole of Science is nothing more

than a refinement of everyday thinking.”

- Albert Einstein

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

During the course of my thesis work there were many people who were instrumental in helping me. I would like to take this opportunity to thank all those people who made this thesis possible and an enjoyable experience for me. More than a duty, it gives me great pleasure to express my sincere thanks to my research guide Dr. Mala Rao, for her excellent guidance, fruitful discussions, valuable advices, constant support and encouragement throughout the course of this work and for having faith in my abilities to perform better. I consider myself extremely fortunate to have her as my research guide.

I am extremely grateful to my collaborators, Prof. Ramakumar and Manikandan Indian Institute of Science, Bangalore for their input in crystallization work. I am also thankful to Dr. D.B.Wilson for his timely help in amino acid sequencing work.

Thanks are also due to Dr. Vasanti Deshpande for her kind support during my Ph.D work.

I am grateful to Dr. Sushama Gaikwad and Dr. Vandana Ghormade for their needful help in some crucial experiments. I also acknowledge help from all the scientists in Biochemical Sciences division.

My special thanks go to my senior Dr. Sudeep George. I am deeply indebted to him for teacher-like guidance in the lab, for fine-tuning my experimental skills and for his selfless help and support. I sincerely thank my seniors Dr. CV and Dr. Jui for their kind support. I also thank them for their affection and kind cooperation.

I express my gratitude to Mrs. Indira Mohandas for her timely help rendered during the course of work. I am thankful to Mr. R. Lambharte for his untiring help in routine chores.

It gives me an immense pleasure to acknowledge my labmates Anamika, Anish, Vinod, Ajit, Aarohi and Gyan Prakash for their cheerful company and their cooperation. I would also like to express

my thanks to ex-labmates Nilesh, Nitin, Neha and trainees Kavita, Shubhangi, Mayura, Aashish, Ganesh, Yogita and Efrain.

I wholeheartedly thank my friends, Shabab, Sajid, Atul, Suresh, Poorva, Smitha, Meena, Jayanthi, Prashant, Nookaraju, Rohini, Vaiju, Namita, Medha, Monali and many others who made my stay at NCL pleasant.

I acknowledge with an appreciation that extends beyond any words for the love and support of my friends Bharati, Uma, Sudipa, Sonali, Feroz , Supriya, Yasmeen, Jyoti and Shubhangi.

I find no words to express my feelings for Sofia for her invaluable help during crisis time.

I owe my deepest gratitude to my parents and sisters for their eternal support and understanding of my goals and aspirations. Without them, I would not have been able to complete much of what I have done and become who I am. Their infallible love and support has always been my strength. Their patience and sacrifice will remain my inspiration throughout my life. The time spent with my two sweethearts, my lovable nephew Shree and niece Sonu (Apoorva) will always be cherished. My brothers Sangram and Vikram deserve special mention for their understanding and caring. I express my deep sense of gratitude to my uncle and specially my auntie Nisha Mavashi for their continuing encouragement and support.

I thank Dr. Sivaram, Director NCL for allowing me to submit my work in the form of a thesis. The fellowship awarded by University Grants Commission (UGC) Govt. of India is gratefully acknowledged.

Though, many have not been mentioned, everybody is cherished.

Sharmili Jagtap

CERTIFICATE

Certified that the work incorporated in the thesis

STUDY OF CELLULOLYTIC AND HEMICELLULOLYTIC ENZYMES FROM A NOVEL ALKALOTHERMOPHILIC *THERMOMONOSPORA* SP.

submitted by **Miss. Sharmili S. Jagtap** was carried out under my supervision at the Division of Biochemical Sciences, National Chemical Laboratory, Pune, India. Material obtained from other sources has been duly acknowledged in the thesis.

November 2006

Dr. (Mrs.) Mala Rao
(Research Guide)

DECLARATION BY THE CANDIDATE

I declare that the thesis entitled **"STUDY OF CELLULOLYTIC AND HEMICELLULOLYTIC ENZYMES FROM A NOVEL ALKALOTHERMOPHILIC *THERMOMONOSPORA* SP."** submitted by me for the degree of Doctor of Philosophy is the record of work carried out by me under the guidance of **Dr. (Mrs.) Mala Rao** and has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other university or other institute of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis.

November-2006 **Sharmili S. Jagtap**

ABBREVIATIONS

ANS	8-Anilinonapthalene-1-sulfonic acid
ATP	Adenosine 5'-triphosphate
CD	Circular dichroism
DEAE	Diethylaminoethyl
DNSA	Dinitrosalicylic acid
EDTA	Ethylene diamine tetra acetic acid
GdnCl	Guanidine hydrochloride
HPLC	high performance liquid chromatography
NBS	*N*-Bromosuccinimide
OD	Optical density
OPTA	*o*-phthalaldehyde
PAGE	Polyacrylamide gel electrophoresis
PEG	Polyethylene glycol
PHMB	*p*-hydroxymercuribenzoic acid
pI	Isoelectric point
PNPG	*p*-Nitrophenyl b-D-glucopyranoside
PNPX	*p*-Nitrophenyl b-D-xylopyranoside
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TEMED	N,N,N'N'- Tetramethyl ethylene diamine
TNBS	2,4,6-trinitrobenzenesulphonic acid
TSC	*Thermomonospora* sp. cellulase / 1,4 – β – D – Glucan Glucanohydrolase
WRK	Woodward's reagent K
Km	Michaelis Menton Constant
Vmax	Maximum velocity
APS	Ammonium persulphate
kDa	Kilo Dalton
ml	Milliliter
µl	Microliter
Å	Angstrom
AU	Atomic units
K	Kelvin
°C	Degree Celsius

ABSTRACT

Cellulose is a fibrous, insoluble, and crystalline polysaccharide made up of β 1-4 linked D-glucopyranosyl residues and constitutes the major structural component of plant cell walls and most abundant biopolymer on Earth with approximately 10^9 tons biosynthesized each year. Cellulose can be degraded by the coordinated action of cellulolytic enzymes, such as endoglucanase, cellobiohydrolase, and β-glucosidase. Cellulolytic microorganisms are abundant in nature and play an important role in the carbon cycle by recycling carbon dioxide fixed by photosynthesis. Efficient, inexpensive enzymatic conversion of cellulosic biomass has potential applications in juice and wine industries, in animal feed and feed supplements, in textile industry, in detergents and in the pretreatment of industrial wastes. *Thermomonospora* is a filamentous thermophilic soil bacterium and an important species degrading cellulose and hemicellulose in plant residues.

The present work has been organized under the following headings

1. General introduction
2. Purification and biochemical characterization of 1,4 – β – D – Glucan Glucanohydrolase
3. Conformation and microenvironment of the active site of 1, 4 – β – D – Glucan Glucanohydrolase
4. Unfolding and chaperone assisted folding of 1, 4 – β – D – Glucan Glucanohydrolase
5. Crystallization and preliminary x-ray characterization of 1, 4 – β – D – Glucan Glucanohydrolase and its amino acid sequence analysis

Summary

Chapter 1. General Introduction

This chapter presents the literature survey of the research carried out on enzymes involved in plant biomass degradation, especially on cellulolytic and hemicellulolytic enzymes with special respect to their occurrence, properties and industrial applications.

Chapter 2. Purification and biochemical characterization of 1,4 – β – D – Glucan Glucanohydrolase

In our laboratory a novel alkalothermophilic actinomycete *Thermomonospora* sp. was isolated from self-heating compost. The strain produces extracellular cellulases and xylanases. One of the extracellular cellulases (1,4 – β – D – Glucan Glucanohydrolase E.C.No.3.2.1.4) was purified to homogeneity from the culture filtrate by fractional ammonium sulphate precipitation (35-55%), DEAE Sephadex A50 and Biogel P100 column chromatography. The molecular weight of the purified enzyme was 14.2 kDa by MALDI - TOF analysis and is in agreement with SDS PAGE. The purified enzyme had a pI of 4.33 with pH and temperature optima 7 and at 50°C respectively. The CMCase was highly thermostable at 50°C for 12h and had a half-life of 1h at 60°C and 20 min at 80°C. The CMCase was stable in a wide pH range from 5 to 8. The secondary structure of the enzyme is made up of α helix regular 0.588, α helix distorted 0.226, β sheet regular0.001, β sheet distorted 0.002, β turn 0.062 and unordered structure 1.04.

Most microbes thought to play a prominent role in cellulose hydrolysis in nature, have evolved strategies that bring the cell close to the cellulose surface and give the cellulolytic organism "first access" to hydrolysis products. The removal of certain polysaccharides by one species or group of microbes may improve the accessibility of a second group to cellulose. In this regard, it is interesting that some cellulolytic bacteria can actively depolymerize certain hemicelluloses, particularly xylans.

This chapter describes the substrate specificity of 1,4 – β – D – Glucan Glucanohydrolase.

The purified enzyme had both endocarboxymethyl cellulase and endoxylanase activities. During purification, the ratio of carboxymethyl cellulase to xylanase activity remained constant indicating that both the activities reside on the same protein. Therefore, the question arises of whether the two activities of the enzyme are localized on the same or on two different active sites. The following data indicated that the enzyme at the same active site hydrolyzes CMC and xylan; (i) the enzyme showed same pH optima for CMC and xylan hydrolysis. (ii) The enzyme was stable at 50°C for 12h.for both substrates. (iii) both the activities were stable in wide pH range of 5 to 8. (iv) both activities were inhibited by O-phthalaldehyde. (v) the mode of action of the enzyme for hydrolysis of both the substrates was of endo type. However the Km for the two activities were widely different, Viz. 6

mg/ml for CMC and 3 mg/ml for Xylan. The kinetic analysis by the method of Keleiti was employed to analyze whether one or two sites are responsible for hydrolysis of CMC and xylan. The experimental data coincides well with the theoretical values calculated for the case of a single active site. Conformation and microenvironment at the active site was probed with fluorescent chemo affinity labeling using o-phthalaldehyde as the chemical initiator. Formation of isoindole derivative resulted in complete inactivation of the enzyme to hydrolyze both xylan and CMC as judged by fluorescence studies corroborating a single active site for the hydrolysis of xylan and CMC.

Chapter 3.Conformation and microenvironment of the active site of 1, 4 – β – D – Glucan Glucanohydrolase

The importance of various functional groups for the activity of 1,4 – β – D – Glucan Glucanohydrolase was investigated by the use of chemical modifiers with restricted amino acid specificity, under the conditions which did not alter the enzyme conformation. Conformation and microenvironment at the active site of 1,4 – β – D – Glucan Glucanohydrolase was probed with fluorescent chemo affinity labeling using o-phthalaldehyde as the chemical initiator. O-phthalaldehyde reacted with endoglucanase yielding a stable fluorescent derivative, which exhibited an absorbance characteristic of isoindole. Enzyme was completely inactivated by o-phthalaldehyde with concomitant formation of isoindole derivative indicating that the –SH and $–NH_2$ groups of enzyme are situated at or close to the active site of enzyme. Further studies were performed to locate the o-phthalaldehyde binding site in the enzyme. Modification of enzyme by 2,4,6-trinitrobenzene sulfonic acid (TNBS) a lysine–specific modifier resulted in inactivation of the enzyme. Kinetic analysis of TNBS modified enzyme implicated the presence of an essential lysine residue at the active site of enzyme. The inhibition of 1,4 – β – D – Glucan Glucanohydrolase by p- hydroxymercury benzoate (PHMB) indicated the presence of an essential Cys residue at the active site of the enzyme. The substrate protection studies confirmed the presence of these residues in the active site. The involvement of carboxyl groups in the catalytic action of the enzyme was investigated using by using the modifier Woodward's reagent K (WRK).

Chapter 4. Unfolding and chaperone assisted folding of 1, 4 – β – D – Glucan Glucanohydrolase

In the present investigation the conformation and microenvironment of the active site of 1,4 – β – D – Glucan Glucanohydrolase has been probed. The enzyme was completely inactivated by low concentrations of GdnCl which was used as an active site perturbant. Fluorescence studies revealed the inactivation of the enzyme by GdnCl precedes gross conformational change. Denaturation studies using GdnCl indicated that the enzyme folds through a partially folded state that resembles the molten globule. Fluorescence and delay experiments revealed that α crystallin interacts with the molten globule state of the enzyme and prevents its aggregation.

Chapter 5. Crystallization and preliminary x-ray characterization of 1, 4 – β – D – Glucan Glucanohydrolase and amino acid sequence analysis

Crystallization and preliminary X- ray characterization

The 1,4 – β – D – Glucan Glucanohydrolase was crystallized by hanging – drop vapour-diffusion method using commercial crystallization kits. Good quality crystals, suitable for crystallization were obtained using 0.2M sodium acetate, 0.1M sodium cacodylate pH 6.5, 18% PEG 8000.

The crystals were diffracted at 2.3 ^{0} A resolution. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit cell parameters a = 49.9, b = 79.5,

c = 99.7Å. The Matthew's coefficient VM was calculated to be 2.7 Å $^{3}Da^{-1}$

Structure solution is not feasible by the molecular – replacement method as the primary structure of the enzyme is not known, although the sequences of several trypsin –digested fragments are available, which we hope signals that a full sequence will be forthcoming. Nevertheless we are in the process of growing better quality crystals in order to obtain an atomic resolution data set which would help in solving the structure by *ab intio* methods

Amino acid sequence analysis

The research interest of this part of the thesis is to sequence the purified 1,4 – β – D – Glucan Glucanohydrolase. The amino acid sequence analysis of the peptides obtained after tryptic digestion was carried out. The C-terminal sequence is **ASHMDQTYYLCSDVNFV**

The other five peptides sequenced are **NYGCWLR, SMDAVGPWK, HVVFTIWK, LPDGQLCSGGLAEGGR, PLTWGSLDLVHR**

Of these **ASHMDQTYYLCSDVNFV** and **LPDGQLCSGGLAEGGR** showed sequence homology with the cellulase binding domain of other enzymes from *Themobifida fusca* and *Streptomyces coelicolor*. The conserved residues found in the multiple sequence alignment of these peptides, were cysteine, serine, phenylalanine, and glycine.

PUBLICATIONS

[1] **Sharmili Jagtap** and Mala Rao (2006) "Conformation and microenvironment of the active site of a low molecular weight 1,4–β–D–Glucan Glucanohydrolase from an alkalothermophilic *Thermomonospora* sp. : Involvement of a lysine and cysteine residues" Biochem. Biophys. Res. Comm. 347, 428-432

[2] **Sharmili Jagtap** and Mala Rao (2005) Purification and properties of a low molecular weight 1,4 – β – D – glucan glucohydrolase having one active site for carboxymethyl cellulose and xylan from an alkalothermophilic *Thermomonospora* sp. Biochem. Biophys. Res. Comm. 329, 111-116

[3] K. Manikandan, **Sharmili Jagtap**, Mala Rao and S. Ramakumar (2006) Crystallization and preliminary x- ray crystallization of a thermostable low- molecular weight 1,4 – β – D – glucan glucohydrolase from an alkalothermophilic *Thermomonospora* sp. ***Acta Cryst.*** F62, 385-387

[4] **Sharmili Jagtap** and Mala Rao , Unfolding and refolding of 1, 4 – β – D – Glucan Glucanohydrolase from *Thermomonospora* sp.: Interactions of α – Crystallin with the Molten Globule State induced by guanidine hydrochloride communicated

[5] "Involvement of carboxylic acid residues in the active site of a thermostable endoglucanase from an alkalothermophilic *Thermomonospora* sp." **Sharmili Jagtap** and Mala Rao, Manuscript under preparation.

[6] Indian patent applied - R. Anish, **S. S. Jagtap**, S. George, A. Ahmad and M. B. Rao. "A process for the preparation of alkali stable cellulase from an alkalothermophilic *Thermomonospora* sp. Indian patent application No. 1298/DEL/2004.

CONFERENCES / ABSTRACTS / POSTERS

[1] "Endoglucanase from a novel alkalothermophilic *Thermomonospora* sp." Poster presented in the 10th international conference of FAOBMB, at IISc Bangalore, India, from December 7-11, 2003

[2] Attended National Seminar on Crystallography, at NCL, Pune from January 8-11, 2003

[3] Attended workshop on "Macromolecular Crystallography" on January 12, 2003, at NCL, Pune.

[4] "Conformation and microenvironment of the active site of endoglucanase from *Thermomonospora* sp." Poster presented at Biotech 2004, October 13-15, 2004.

[5] "Purification and properties of a low molecular weight 1,4 – β – D – glucan glucohydrolase having one active site for carboxymethyl cellulose and xylan from an alkalothermophilic *Thermomonospora* sp." Oral presentation at Biotron'05, a national level symposium, at Hyderabad on January 28-29, 2005.

[6] "Biochemical characterization of endoglucanases from cellulose utilizing Thermomonospora sp. and their application" Mala Rao, Anish Ramakrishnan, Sharmili Jagtap Poster presented at National conference on Forest Biodiversity Resources: Exploitation, Conservation and Management, at Chennai on March 21-22,2006

CHAPTER-I

GENERAL INTRODUCTION

Cellulose is the most abundant biopolymer on earth and approximately 10^7 tons is synthesized per annum. It is a fibrous, insoluble, and crystalline. It is a linear polysaccharide made up of β -1-4 linked D-glucopyranosyl residues and constitutes the major structural component of plant cell walls. Each glucose residue is rotated by 180° relative to its neighbours, so that the basic repeating unit is cellobiose. The chain length varies between 100 and 14000 residues. The fully equatorial conformation of β -linked glucopyranose residues stabilizes the structure, minimizing its flexibility.

The end of the glucan chain with an anomeric carbon that is not linked to another glucose residue is referred to as the reducing end of the polymer. The other end of the polymer is the non-reducing end. Immediately after synthesis the cellulose chains coalesce into highly crystalline cellulose microfibrils held together by hydrogen bonds, hydrophobic interactions and van der Waals forces. This highly organized packing of the cellulose chains makes it resistant to hydrolysis. The width of the micro-fibrils depends on the source of the cellulose.

Structural unit

Fig.1 Cellulose is a linear polymer of β-(1 - 4)-D-glucopyranose units.

The structure of cellulose

The substrate cellulose has a complex structure despite a simple chemical composition consisting of D-glucose residues linked by β-1,4-glucosidic bonds. Cellulose is difficult to degrade because the molecules form tightly packed, extensively hydrogen bonded regions called crystalline cellulose. The crystalline regions are believed to be separated by less ordered amorphous regions, but these still contain many hydrogen bonds. It's crystalline and makes native cellulose a recalcitrant substrate for enzymatic hydrolysis. Microorganisms meet this challenge with the aid of a multi-enzyme system. Aerobic bacteria produce numerous individual, extra-cellular enzymes with binding modules for different cellulose

conformations. Specific enzymes act in synergy to elicit effective hydrolysis (Lynd *et al.,* 2002). In contrast, anaerobic bacteria possess a unique extracellular multienzyme complex, called cellulosome (Forsberg *et al.,* 1981). All organisms known to degrade cellulose efficiently produce a battery of enzymes with different specificities, which act together in synergism. Cellulolytic microorganisms are abundant in nature and play an important role in the carbon cycle by recycling carbon dioxide fixed by photosynthesis. All organisms known to degrade cellulose efficiently produce a battery of enzymes with different specificities, which act together in synergism. Cellulases have been found in insects, plants and fungi while bacteria producing cellulases are found wherever cellulose is present such as in compost piles, soil, rotting wood, etc (Andre *et al.,* 2003; Berghem and Petterson .,1974; Hakamada *et al.,* 2002; Kirby *et al.,*1997;Wilson and Irwin,1997) Many animals, including ruminants, utilize the cellulose present in their food; however, they do not produce cellulases but rely on cellulolytic microorganisms, primarily bacteria, to hydrolyze the cellulose.

The cellulase system is considered to comprise three hydrolytic enzymes (Bhat., 2000)

i) the endo-(1,4)-β-D-glucanase (synonyms: endoglucanase, endocellulase, carboxymethyl cellulase or 1,4-β-D-glucan glucanohydrolase [EC 3.2.1.4], which cleaves β-linkages at random, commonly in the amorphous parts of cellulose;

ii) the exo-(1,4)- D-glucanase or 1,4-β- D-glucan cellobiohydrolases (synonyms: cellobiohydrolase, exocellulase, microcrystalline cellulase, Avicelase [EC 3.2.1.91]), which releases cellobiose from either the nonreducing or the reducing end, generally from the crystalline parts of cellulose; and

iii) the β-glucosidase or β- D-glucoside glucohydrolase (synonym: cellobiase [EC 3.2.1.21], which releases glucose from cellobiose and short-chain cellooligosaccharides. Although β-glucosidase has no direct action on cellulose, it is regarded as a component of cellulase system because it stimulates cellulose hydrolysis.

Cellulase Synergism

A remarkable feature of the cellulose system is synergism; that is, the action of a mixture of two or more cellulases is greater than the sum of the action of each individual component. Individual cellulases alone will not degrade crystalline cellulose beyond about 5%, no matter how much enzyme is added to the reaction. From the currently available data it can be postulated that synergism in crystalline cellulose occurs when two or more cellulases attack different sites in cellulose and, as a result of their attack, new sites are created for each other. For example, when a non-reducing end attacking exocellulase binds and processively cleaves a cellulose chain, it disrupts the hydrogen bonds to neighboring chains creating some new sites at which endocellulases can attack; one or more reducing ends may also be exposed during this process. Thus, this exocellulase can synergize with both endocellulases and reducing end attacking exocellulases. Endocellulases create both types of ends with each cleavage so they can synergize with both types of exocellulase. It has been shown experimentally in synergistic mixtures containing an endo- and an exocellulase that hydrolysis by the endoglucanase is stimulated as much as hydrolysis by the exocellulase (Irwin *et al.,* 1993). A schematic representation of the synergism by noncomplexed and complexed cellulases system is presented in Fig. 2

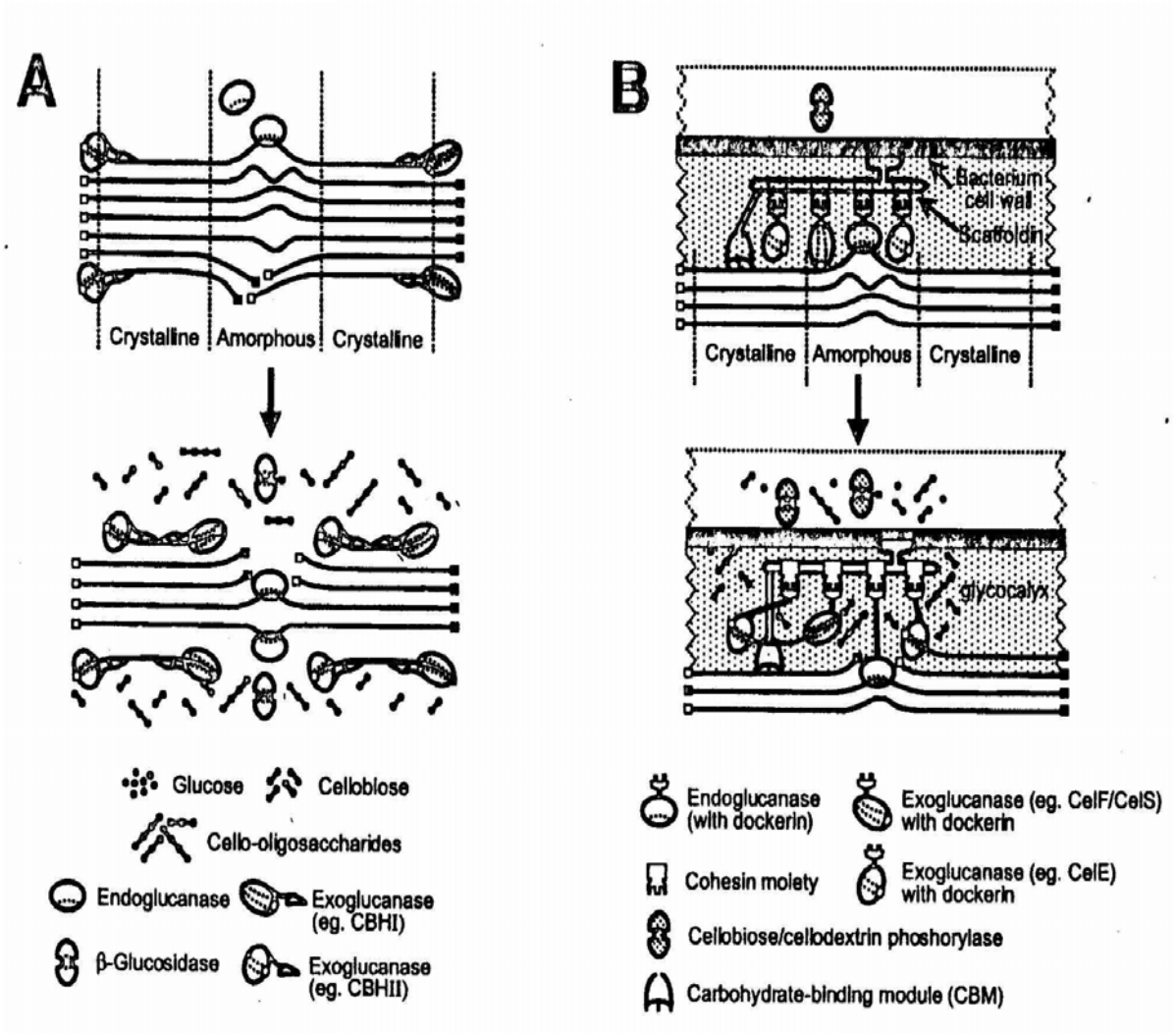


Fig. 2 Schematic representation of the hydrolysis of amorphous and microcrystalline cellulose by noncomplexed (A) and complexed (B) cellulase systems. The solid squares represent reducing ends, and the open squares represent nonreducing ends. Amorphous and crystalline regions are indicated. Cellulose, enzymes, and hydrolytic products are not shown to scale

Table-1

Characteristics of some fungal endoglucanases

Fungus	Molecular Weight (Da)	pI	Optimum temperature (°C)	Optimum pH
Aspergillus aculeatus [Murao *et al.,* 1988]	25000	4.8	50	4.5
	38000	3.4	65	4.0
	66000	4.0	70	5.0
	68000	3.5	60	2.5
Aspergillus niger [Okada, 1988]	31000	3.67	–	4.0
Coriolus versicolor [Idogaki and Kitamoto, 1992]	29500	–	55	5.0
Dichomitus squalens [Rouau and Foglietti,1985]				
I	42000	4.8	55–60	4.8–5.0
II	56000	4.3	55–60	4.8
III	47000	4.1	55–60	4.6–4.8
Humicola insolen [Hayashida *et al.,* 1988]	57000	–	50	5.0
Humicola grisea [Hayashida *et al.*,1988]	63000	–	50	5.0
	58500	6.5	–	5.0
Mucor circinelloides [E.Morag *et al.*,1990]	27000	–	55	4.0–6.0
Mucor circinelloides [Nakamura *et al.,* 2001]	41000	–	45–55	6.0
Meruliporia incrasata	25000	>3.6	50	

[Kleman-Leyer *et al.*, 1994]				
	49000	>3.6	–	–
	57000	>3. 6	–	–
Rhizopus oryzae [Murashima *et al.,* 2002]				
I	41000	–	55	5.0–6.0
II	61000	–	55	5.0–6.0
Thermoascus aurantiacus [Parry *et al.,* 2002]	34000	3.7	70–80	4.0–4.4
Trichoderma koningii [Wood *et al.,* 1978]	13000	4.72	–	–
	31000	5.09	–	–
	48000	4.32	–	–
Trichoderma reesei [Bhikhabhai *et al.,* 1984]	20000	7.5	–	–
	43000	4.0	–	–
	48000	5.5	–	–
	55000	4.5	–	–
	56000	5.0	–	–
	67000	6.5	–	–
Trichoderma viride QM 9414 [Vorage n *et al.,*1988]	23500	7.7	–	5.5
	45000	6.9	–	4.0
	50000	5.3	–	5.1
	52000	3.5	–	4.5
	57000	4.4	–	4.5

Bacteria

The Table no. 2 comprises the cellulolytic enzyme systems from the bacterial of different genera (Robonson *et al.,* 1989)

Table-2

Cellulases of Bacterial Origin

Genus	Species
Clostridium	*thermocellum*
	cellulolyticum
	Stercorarium
Cellulomonas	*fimi*
	uda
	flavigena
	fermentans
Bacillus	*subtilis DLG*
	subtilus N-24
	subtilus PAP115
	polymyxa
	licheniformis
	Cereus
Bacillus N-4	
Bacillus 1139	
Thermomonospora	*fusca*
	curvata
Ruminococcus	*albus*
	flavefaciens
Bacteroides	*succinogens*
	cellulosolvens
Erwinia	*chrysanthemi*
Acetivibrio	*cellulolyticus*

It should be noted, however, that cellulolytic bacteria also exist among the *Streptomyces* (Ishaque *et al.,* 1980; Van Zyl,1985; MacKenzie *et al.,* 1984; MacKenzie *et al.,* 1984), *Cellvibrio* (Wynne and Pemberton.,1986; Oberkotter and Rosenberg, 1978)*, Pseudomonas* (Yamane *et al.,*1970; Wolfe *et al.*, 1986; Schell, 1987; Lejeune *et al.,*1988) *Rhizobium,* and *Microbispora.*(Yu and Hungate, 1979)

Cellulose degradation by Anaerobic Bacteria

Cellulosome

Cellulosomes are large, surface-bound, extracellular multienzyme complexes responsible for the degradation of crystalline cellulose and associated plant cell wall polysaccharides.First identified in the thermophilic anaerobe *Clostridium thermocellum* (Lamed *et al.,*1983a; Lamed *et al.,* 1983b;) cellulosomes have now been characterized in a variety of other anaerobic bacteria, including *Clostridium celluloyticum* (Pages *et al.,* 1996; Gal *et al.,* 1997), *Clostridium cellulovorans* (Doi *et al.,* 1994), *Clostridium josui* (Kakiuchi, *1998) Acetivibrio cellulolyticus* (Ding *et al.,* 1999; Xu *et al.,* 2003), and *Bacteroides cellulosolvens* (Ding *et al.,* 2000; Xu *et al.,* 2004), as well as ruminal bacteria (Ding *et al.,* 2001) and anaerobic fungi (Steenbakkers *et al.,* 2001). The general quaternary architecture of cellulosomes includes numerous catalytic subunits constituting a variety of cellulases and hemicellulases, which are organized on a large, multimodular, noncatalytic scaffoldin subunit to form a single macromolecular machine able to efficiently and synergistically breakdown crystalline cellulose (Bhat *et al.,* 1992; Bhat *et al.,* 1994). A highly conserved noncatalytic module, termed dockerin, of the catalytic subunits is responsible for anchoring each enzyme to the scaffoldin subunit (Bayer *et al.,* 1994; Tokatlidis *et al.,* 1991; Tokatlidis *et al.,* 1993). Table 3 summarizes the details of components of cellulosome.

Table-3

Known components of the cellulosome of *Clostridium thermocellum*

Cellulosome Components	Description	Modular Structure	Molecular Mass (kDa)
CipA	Scaffoldin	2(Coh1)-CBM3a-7(Coh1)-X2-DocII	197
CelJ	Cellulase	X-Ig-GH9-GH44-DocI-X	178
CbhA	Cellobiohydrolase	CBM4-Ig-GH9-2(Fn3)-CBM3b-DocI	138
XynY	Xylanase	CBM22-GH10-CBM22-DocI-CE1	120
CelH	Endoglucanase	GH26-GH5-CBM9-DocI	102
CelK	Cellobiohydrolase	CBM4-Ig-GH9-DocI	101
XynZ	Xylanase	CE1-CBM6-DocI-GH10	92
CelE	Endoglucanase	GH5 - DocI - CE2	90
CelS	Exoglucanase	GH48 - DocI	83
CelF	Endoglucanase	GH9-CBM3c-DocI	82
CelN	Endoglucanase	GH9-CBM3c-DocI	82
CelQ	Endoglucanase	GH9-CBM3c-DocI	82
CelO	Cellobiohydrolase	CBM3a-GH5- DocI	75
XynA, XynU	Xylanase	GH11-CBM4-DocI-NodB	74
CelD	Endoglucanase	Ig - GH9 - DocI	72
XynC	Xylanase	CBM22-GH10-DocI	70
XynD	Xylanase	CBM22-GH10-Doc1	70
ManA	Mannanase	CBM-GH26-DocI	67
CelT	Endoglucanase	GH9 - DocI	65
CelB	Endoglucanase	GH5 - DocI	64
CelG	Endoglucanase	GH5 - DocI	63
CseP	Unknown	UN - DocI	62
ChiA	Chitinase	GH18 - DocI	55
CelA	Endoglucanase	GH8 - DocI	53
XynB, XynV	Xylanase	GH11-CBM4-DocI	50
LicB	Lichenase	GH16 - DocI	38

Abbreviations : CE, carbohydrate esterase; CBM, substrate binding module; Coh1, typ-I cohesin module; DocI, typ-I dockerin module; FAE, ferulic acid esterase; GH, glycosyl hydrolase (family); Ig, immunoglobulin-like module; UN, unknown module; X, hydrophobic module with unknown function.

Classification and structure of cellulases

The protein sequences were classified in glycosyle hydrolase families (GHF) based on amino acid sequence similarities (Henrissat and Bairoch, 1996; Coutiho and Henrissat, 1999). The overall 3D structure and the stereo specificity of hydrolysis are conserved within a family. Cellulases have been assigned to 12 of the Glycosyl Hydrolases (GH) families: 5 – 9, 12, 26, 44, 45, 48, 61 and 74 (http://afmb.cnrs-mrs.fr/CAZY) (Henrissat and Bairoch, 1996; Henrissat and Davies, 1997). The Carbohydrate-Active enZYmes Server (CAZY) provides a classification of glycoside hydrolases. Structural data on individual cellulase modules have been accumulating over the last few years, with three-dimensional structures solved in nine families of catalytic modules and seven families of CBMs (Bourne *et al.*, 2000). Amino acid sequence comparisons, supplemented by hydrophobic cluster analysis, have allowed classification of glycoside hydrolases into more than 60 families, which, on the basis of crystallographic investigations, can now be grouped into "clans" of related structures. (Henrissat, 1991; 1997) A `clan' is a group of families that are thought to have a common ancestry and are recognized by significant similarities in tertiary structure together with conservation of the catalytic residues and catalytic mechanism. The growing number of three dimensional structures solved for glycosyl hydrolases and / or improved sequence comparison strategies have revealed the relationship between some glycosyl hydrolases families which can be grouped in clans (Table 4). Many cellulases genes have been cloned and sequenced (Henrissat *et al.*, 1989). Henrissat and colleagues have grouped the bulk of these genes into elevenfamilies based on sequence homology and hydrophobic cluster analysis.

Table 4

Clan grouping of glycosyl hydrolase families

Clan	Families grouped
GH-A	1,2,5,10,17,30,35,39,42
GH-B	7,16
GH-C	11,12
GH-D	27,36
GH-E	33,34

Most cellulases have a modular structure with a catalytic module and a cellulose-binding module (CBM) usually joined by a glycosylated and presumably flexible linker peptide (Gilkes *et al.*, 1991). Removal of the CBM results in a significantly reduced activity of the enzymes on crystalline cellulose probably because of a decreased binding capacity, but the activity on soluble cellulose oligomers is retained (Gilkes *et al.*, 1988, Tilbeurgh *et al.*, 1986, Tomme *et al.*, 1988). Inversely, isolated CBMs retain most of the binding capacity on crystalline cellulose but are devoid of catalytic action (Tilbeurgh *et al.*, 1986, Tomme *et al.*, 1988). Enzymatic activity is sometimes affected when the interdomain linker has been shortened or completely deleted suggesting that it should be of sufficient length and/or flexibility to ensure an independent action of the two functional modules (Gilkes *et al.*, 1992, Shen *et al.*, 1991). Linkers are short amino acid sequences that connect the different domains found in cellulases. Usually, linkers comprise flexible disordered chains rich in proline and hydroxy amino acid residues (serine and threonine), as well as glycine and alanine. Their length can vary from 5-6 to 100 residues, although most often is limited by the range 20 to 50 residues. In many linkers repeated motifs of four-seven residues can be identified, where some positively or negatively charged, or hydrophobic residues are inserted within PTS - rich sequence (Proline- Threonine- serine).

The linker sequence in *C. fimi* CenA is PTPTPTPTT (PT) 7 (Shen *et al.*, 1991 and in *T. reesei* CBHI is NPSGGNPPGGNPPGTTTTRRPATTTGSSPGP (Srisodsuk *et al.*, 1993), while *T. fusca* E4 has the linker sequence such as PEEGEEPGGGEGPGGGEEPG (Jung *et al.*, 1993).

Carbohydrate binding Modules/ Domains (CBMs/CBDs)

Carbohydrate binding modules (CBMs) target the enzymes to specific polysaccharides. The resultant increase in the concentration of the hydrolases greatly enhances the catalytic efficiency of these enzymes on insoluble and recalcitrant substrate.

Table 5

Some characteristics of the major CBD families

Family	Typical size	Typical origin, reference to available structures
I	33-40	The CBD family characteristic to fungi such as T. reesei. *Humicola insolens, Fusarium oxysporum.* Over 30 known members. Structure of *T. reesei* CBHI CBD published (Kraulis *et al.,* 1989).
II	90-108	A CBD type found in *Cellulomonas fimi, Thermomonospora fusca, Pseudomonas fluorescens* and other bacteria. Over 30 known members, divided into two subfamilies. Structure of C. *Jimi Cex* CBD published (Xu *et al.,* 1995).
III	132-172	Found in enzymes from *C. thermocellum, C. cellulocorans, Bacillus lautus* and other bacteria. Also found in cellulosomes. Over 20 known members, divided into two subfamilies. Structure of *Cl. thermocellum* CipB CBD published (Tormo *et al.,* 1996).
IV	125-168	So far identified in enzymes from *C.fimi, C. celluloly/icum* and some other bacteria. Less than ten known members. Structure of the soluble oligosaccaride CBD of the *C.fimi* CenC published (Johnson *et al.,* 1996).

All CBMs so far characterized have structures based on different β- sheet topologies. CBMs have been grouped into a number discrete families based upon amino acid sequence similarity (Coutinho & Henrissat 1999). Boraston *et al.* (2000) have classified the structures from 22 CBM families into seven 'fold families' (Table 6). Within several of these sequence-based families, subtle differences in structure lead to diverse ligand specificity (Coutinho and Henrissat 1999, Simpson *et al.*, 2000, Boraston *et al.*, 2002, Czjzek *et al.*, 2001). Thus, CBMs in complex with their target sugars present excellent systems for

dissecting the molecular determinants that define the structural basis for protein–carbohydrate recognition.

Folds in CBM

The *β*-sandwich

In terms of both total number of families and entries in databases, the dominant fold among CBMs is the *β*-sandwich (fold family 1). This fold comprises two *β*-sheets, each consisting of three to six antiparallel β-strands.

The *β*-trefoil

This fold contains 12 strands of *β*-sheet, forming six hairpin turns. A *β*-barrel structure is formed by six of the strands, attendant with three hairpin turns. The other three hairpin turns form a triangular cap on one end of the *β*-barrel called the 'hairpin triplet'. The subunit of this fold, called here a trefoil domain, is a contiguous amino acid sequence with a four *β*-strand, two-hairpin structure having a trefoil shape. Each trefoil domain contributes one hairpin (two *β*-strands) to the *β*-barrel and one hairpin to the hairpin triplet. The fold of the resulting molecule has a pseudo-3-fold axis.

OB (Oligosaccharide Binding) fold

In contrast, members of fold families 3–5, which are small 30–60-amino-acid polypeptides containing only *β*-sheet and coil , show less diversity in their ligand specificities with folds that appear more specialized to the recognition of cellulose and/or chitin. The majority of these CBMs have planar carbohydrate-binding sites comprising aromatic residues.

Hevein fold

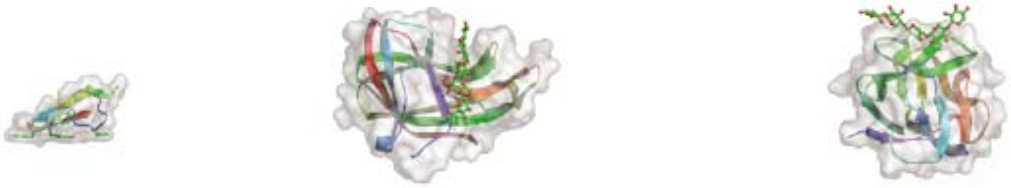
Hevein domains are small (approx. 40 amino acids) CBMs originally identified as chitin-binding proteins in plants. The fold comprises predominantly coil, but does have two small *β*-sheets and a small region of helix

Table 6

CBM fold families

Fold family	Fold	CBM families
1	β- Sanwich	2,3,4,6,9,15,17,22,27,28,29,32,34,36
2	β- Trefoil	13
3	Cysteine knot	1
4	Unique	5,12
5	OB fold	10
6	Hevein fold	18
7	Unique; contains hevein –like fold	14

Three-dimensional (3D) structures of CBMs have shown that the topography of the binding sites in these proteins reflects the macroscopic nature of the target ligand. This has led to the classification of these modules into types A, B or C (Fig. 3). Type A CBMs, which interact with crystalline cellulose, display a planar hydrophobic surface that interacts with adjacent chains of the 010 face of the crystal lattice (Czjzek *et al.*, 2001, Simpson *et al.*, 2002). Type B CBMs bind to less-structured plant structural polysaccharides, accommodating single sugar chains in a groove that extends along the complete length of the protein and are the commonest class of CBMs (Nagy *et al.*, 1998; Kormos *et al.*, 2000; McLean *et al.*, 2000; Boraston *et al.*, 2001; Charnock *et al.*, 2002,). Type C CBMs contain small binding sites that interact with mono- or disaccharides. A characteristic of CBMs that distinguishes them from lectins is a scarcity of hydrogen bonds between these proteins and their target ligands; instead, binding is dominated by hydrophobic interactions (Coutinho & Henrissat 1999, Boraston *et al.*, 2001, Xie *et al.*, 2001 Charnock *et al.*, 2002, Notenboom *et al.*, 2002). The lack of hydrogen bonds facilitates the accommodation of heterogeneous ligands in which the backbone saccharide polymer is often decorated with an array of different sugars and organic acids (McLean *et al.*, 2000), and may contribute to the necessarily dynamic nature of the interaction.



Type A **Type B** **Type C**

Fig. 3 A classification for carbohydrate-binding modules. (Adapted from Davis *et al.*, 2004)

Catalytic mechanisms of glycoside hydrolases

All cellulases hydrolyze the β-1,4-glucosidic bond between glucosyl moieties by a general acid catalysis requiring a proton donor and nucleophile /base. They release the products either by overall retention or inversion of the anomeric configuration at C_1 (Fig. 4).Glycoside hydrolases cleave glycosidic bonds using acid-base catalysis. Catalysis may be performed with either inversion or net retention of the anomeric configuration of the substrate. All glycosidases may also be classified based upon whether they retain or invert the anomeric configuration during catalysis. There are two stereo chemically different mechanisms of hydrolysis for cellulases as shown n Fig.4. The two basic mechanisms were proposed by Koshland (Kirby *et al.*, 1997). The catalytic mechanisms have been well characterized and frequently reviewed (Kleman-Leyer and Kirk 1994, Murashima *et al.*, 2002, Parry *et al.*, 2002, Wood and McCrac 1978). Inversion is a simple single displacement reaction. A catalytic acid gives protonic assistance to leaving group departure (glycoside hydroxyls have a high pKa and thus make poor leaving groups) whilst a catalytic base is required to deprotonate a water for nucleophilic substitution at the anomeric centre. The acid and base are typically located some 7-13Å apart in order to accommodate the nucleophilic water `below' the pyranoside ring. One carboxyl side is protonated and acts as a catalytic acid, donating its proton to the glycosidic oxygen of the leaving group. The other carboxyl side chain acts as a catalytic base removing a proton from the water molecule which attacks the C_1 carbon thus inverting the linkage from β toα. On many cellulase systems, the identification, indeed the existence, of the catalytic base remains controversial.

The retention mechanism is a double displacement essentially as outlined by Koshland in 1953. A covalent glycosyl-enzyme intermediate is formed, and subsequently hydrolyzed, via

oxocarbenium ion-like transition states. This requires two essential residues, an enzymatic nucleophile and a catalytic acid/base which first serves as a classical Bronsted acid, protonating the leaving group to assist departure then functions as a base, deprotonating the incoming water nucleophile for the second step. The nucleophile and acid/base are always found some 5-6 Å apart on all systems studied thus far. Since catalytic mechanism is dictated by the location of functional groups on the protein, the stereochemistry of catalysis is conserved within each family (Fig. 4)

.

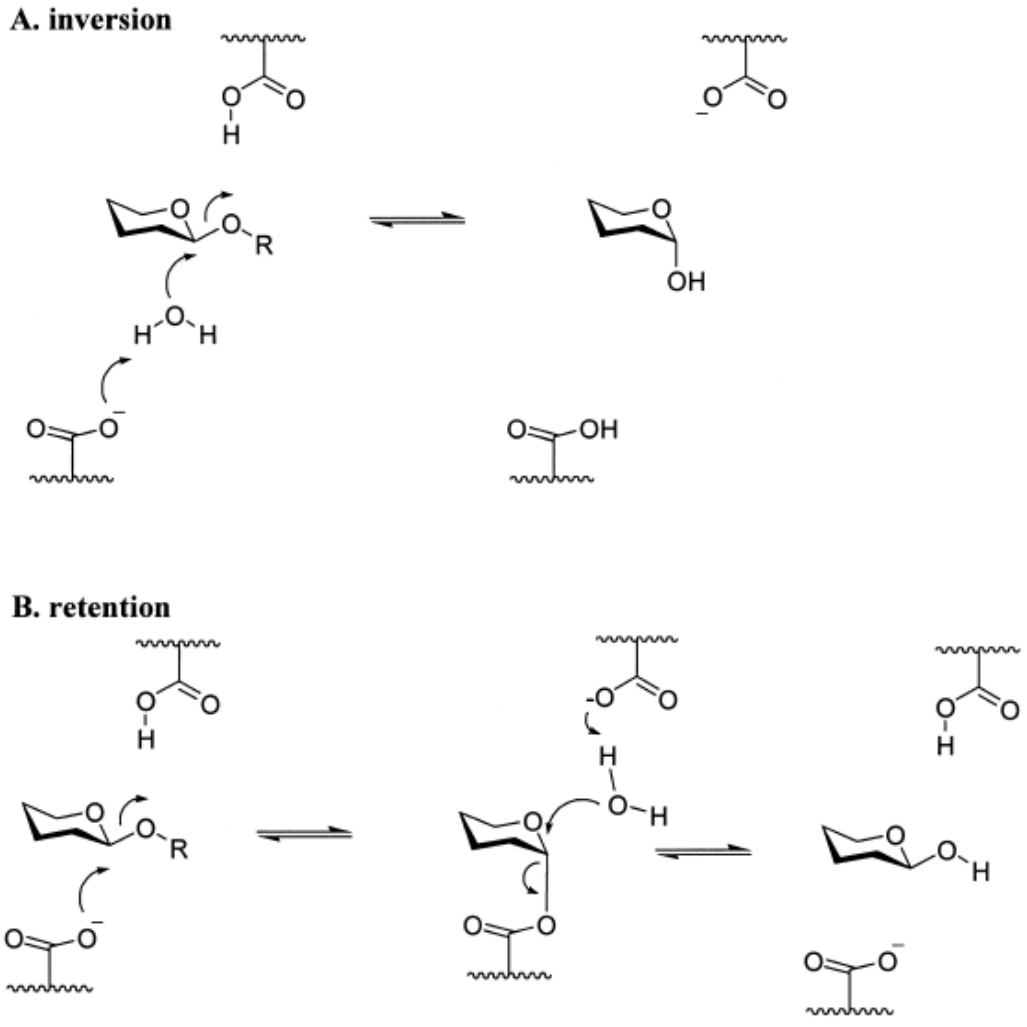


Fig. 4: The two catalytic mechanisms of glycosyl hydrolases A: Inversion, B: Retension

Regulation of cellulase production

In *T. Reesei* the production of cellulases is induced only in the presence of the substrate cellulose but suppressed when easily utilizable sugars, such as glucose, are available. As early as 1962, sophorose (β-1,2-glucobiose) was identified as a strong inducer of cellulases in *T. reesei* (Mandels *et al*., 1962). However it has not been demonstrated conclusively that sophorose is the natural inducer of cellulase production. Cellobiose, δ-cellobiose-1,5-lactone and other oxidized products of cellulose hydrolysis, or even xylobiose resulting from xylan hydrolysis, have not been ruled out as the natural inducer(s). Moreover the possibility that cellobiose functions as an inducer is more complex because at high levels it inhibits cellulase production (Kubicek *et al*., 1998). It should be noted that not only the production of cellulases but also the production of hemicellulases is induced, presumably reflecting the interwined occurrence of these polymers in nature (Margolles-Clark *et al*., 1997). A study of four filamentous fungi revealed that extracellular cellulase was repressed at intracellular ATP concentration above 10^{-7}mb/ml and that the cyclic AMP (cAMP) played a role in derepression of enzyme synthesis (Wang *et al*., 1996). Basal levels of cellulase production presumably allow the production of the inducer through limited cellulose hydrolysis, which in turn mediates further induction of cellulase production. The mechanism by which sophorose or other inducers stimulate transcription through the transcriptional activators ACEI and ACEII is not clear yet. Expression of cellulase genes of *T. fusca* is also regulated at two levels: induction by cellobiose and catabolite repression in the presence of glucose (Wilson, 1992). CelR represses cellulase production in the absence of cellulose and cellobiose. However, cellobiose acts as a inducer and inactivated CelR, thereby facilitating its dissociation from promoters allowing transcription of cellulase genes (Spiridonov *et al*., 1999). Catabolite repression of cellulase genes occurs in the presence of glucose and may be regulated by cAMP levels, as indicated by studies done with *Thermobifida curvata* (Wilson *et al*., 1999; Wood *et al*., 1984). Cellulosic substrates, as well as cellobiose and xylose, at moderate levels of 0.05 to 0.2 g/liter, serve as inducers for cellulase production (Rodriguez *et al*., 1996).

Ecological aspects of cellulose-degrading communities

Different habitats in which cellulose is widely available, by their differing characteristics have fostered the development of cellulose utilization strategies that differ in enzyme architecture and presentation, rate and extent of cellulolysis, ancillary hydrolytic activities, fate of hydrolytic products, and interactions among cellulolytic and non cellulolytic microbes. In soils, cellulose is available primarily in the form of litter (dead plant material) that is relatively recalcitrant due to the high lignin content of terrestrial plants. A lack of fixed nitrogen and other nutrients may secondarily limit microbial growth, and the low moisture content of soils often favors the growth of fungi as the dominant cellulolytic biota (Lynch, 1988). The fungal strategy for cellulolysis involves extracellular cellulases that work alongside lignolytic enzyme systems. In nature, cellulose utilization is carried out not by pure cultures of microorganisms but by multiple cellulolytic species coexisting with each other and with many noncellulolytic species. While cellulolytic species may compete directly for cellulose (Chen *et al*., 2001), both cellulolytic and noncellulolytic species can compete for cellodextrin products of cellulose hydrolysis. Of particular interest in the context of improving the bioconversion of cellulosic biomass are reports of synergistic interactions among fibrolytic and non-fibrolytic bacteria grown on authentic plant cell wall material (Fondevila *et al*., 1996). In therse examples, more complete utilization of cellulose and/or hemicelluloses sometimes has been observed, apparently by the simultaneous or sequential depolymerization of different, intimately associated cell wall polysaccharides. Instead, it appears likely that removal of certain plant cell wall polysaccharides by one species or group of microbes may improve the accessibility of a second group to cellulose or to hemicelluloses. In this regard, it is interesting that some cellulolytic bacterial can actively depolymerize certain hemicellulose (particularly xylans) and pectins but not effectively utilize the component monosaccharides and oligosaccharides, even in pure culture (*i.e*., in the absence of competition from noncellulolytic species (Fondevila *et al*., 1994). Moveover, the soil bacterium *Cellulomonas* sp. strain ATCC 21399 synthesizes array of hydrolytic enzymes when grown on cellulose but not when grown on xylan, starch or other substrates (Poulsen *et al*., 1988). This suggests that in nature these cellulolytic species utilize xylanases and pectinases primarily as tools to gain access to cellulose; i.e., they have sacrificed a reasonably abundant energy source (xylan or pectin) in exchange for an oppourtunity to

exploit an even more abundant energy source (cellulose) that is utilizable by fewer competitors. Enzymatic cleavage of cell wall linkages between cinnamic acid and arabinoxylans by *Ruminococcus albus* and *Butyrivivrio fibrisolvens* has been proposed as another example of an activity whose primary function is to enhance accessibility of an organism to cellulose (McSweeney *et al.*, 1998). The ease with which cellulolytic microbes establish mutualistic interactions with noncellulolytic microbes has important implications in natural environments.

Biotechnological applications

Cellulosic material is the only sustainable source of fuels because of their relatively low cost and plentiful supply. The potential large-scale production of liquid fuels and chemical feed stocks from renewable biological sources is now receiving significant attention worldwide. The biodegradation and bioconversion of lignocelluloses into useful products and biological alleviation of pollution from lignocellulosic wastes is an enormous environmental challenge. Lignocellulosic biomass such as municipal solid waste, agricultural and forest residues and paper waste is mainly composed of cellulose, hemicellulose and lignin. The application of enzymes in detergent, leather and paper industries demands identification of highly stable enzymes active at extreme pH and temperature. The search for extremophilic organisms is one of the means for obtaining enzymes with properties suitable for industrial applications.

Cellulase preparations have considerable ecological and industrial importance. Efficient, inexpensive enzymatic conversion of cellulosic biomass has following potential applications (Buchert *et al.*, 1997, Niehaus *et al.*, 1999, Nakamura *et al.*, 2001, Van wyk *et al.*, 2001, Zhou *et al.*, 2001, Ingram *et al.*, 1995, Lynd *et al.*, 1991, Horikoshi 1996, Miettinen-Oinonen and Suominen 2002, Collins *et al.*, 2003).

- In production of single-cell protein for use in animal feed, feed supplements to increase digestibility.
- In the color extraction of juice from fruits and vegetables and
- In the production of ethanol for use as a fuel or chemical feedstock.
- In the current industrial processes, in detergents causing color brightening and softening,
- Cellulases also have been shown to have potent applications in detergent industry to remove damaged fibers (piles) from cotton fabrics. They are also used for "biostoning" to replace the stone washing process for denim fabric as well as to change the texture of other fabrics.
- in the pretreatment of biomass that contains cellulose to improve nutritional quality of forage
- in the pretreatment of industrial wastes and
- Cellulases are used to lower energy consumption in mechanical pulp production by improving the beatability, drainage and solubility of chemical pulps.

The biotechnological application of cellulases demands the identification of highly stable enzyme preparations active and stable at high pH and temperature. Extremophilic microbes isolated from diverse environments represent a source of such enzymes, which can withstand the rigors of industrial process. Thermophilic organisms frequently have enzymes displaying greater thermal stability than those of mesophiles. (Stutzenberger 1972; Wilson 1988) There are number of reports on thermophiles producing cellulases (Dey *et al.*, 1992; Okazaki *et al.*, 1984). A few alkalothermophilic Bacillus capable of producing enzymes of industrial importance are also known (Okazaki *et al.*, 1984; Poulsen and Peterson, 1988; Rajaram and Varma 1990).

Thermomonospora species

Thermomonospora belongs to the order actinomycetales of the actinomycetes group and is a filamentous soil bacterium. It is generally thermophilic with optimum growth temperatures ranging from 48 to 55°C and aerobic in nature. They appear to be major degraders of cellulose and hemicellulose in heated plant residues such as compost heaps, rotting hay, manure pile, and paper mill waste. *Thermomonospora* is distinct from the genus *Thermoactinomyces* that contains organisms that produce heat-resistant forespores, which Thermomonospora does not. Thus far the best-characterized systems are those of *T.fusca* (Meyer and Humphrey 1982; Crawford and McCoy 1972; Calza *et al.*, 1985; Ghangas and Wilson, 1987; Lin and Wilson 1987; Moreira *et al.*, 1981; Stutzenberger 1971; Hagerdal *et al.*, 1978; Fennington *et al.*, 1983) and *T.curvata* (Fennington *et al.*, 1983; 1984). Concurrent with growth, these organisms produce highly active extracellular cellulases that degrade both amorphous and crystalline cellulose and that do not appear to be present in a complex (Hagerdal *et al.*, 1978).

Table-4

Properties of purified cellulases from *Thermomonospora* sp. (Irwin *et.al.*, 1998)

Enzyme	Mol. Wt. (kDa)	Stereochemistry	Family
E1	101.2	Inversion	E
E2	43.0	Inversion	B
E2cd	30.0	Inversion	B
E3	65.0	Retension	B
E4	90.2	-	E
E5	46.3	Retension	A
E5cd	34.4	Retension	A
E6	106.0	-	-

Thermobifida is a bacterial genus where *T. fusca* (previously *Thermomonospora fusca*) is a complete cellulase system. Its cellulases are spread over GH-families 5, 6, 9, and 48. The first cellulase with both a catalytic domain and a CBM to be crystallized was a truncated form of Cel9A (E4-68, lacking one C-terminal CBM) from *T. fusca* (Sakon *et al.*, 1997; Irwin *et al.*, 1998). Important studies regarding synergy and substrate binding has been done for Cel5A, Cel6B and Cel48A (Jeoh *et al.*, 2002; Jung *et al.*, 2003). Cel6B has a documented preference for non-reducing ends (Jung *et al.*, 2003). Three-dimensional structure is reported for Cel6A and Cel9A (Sakon *et al.*, 1997).

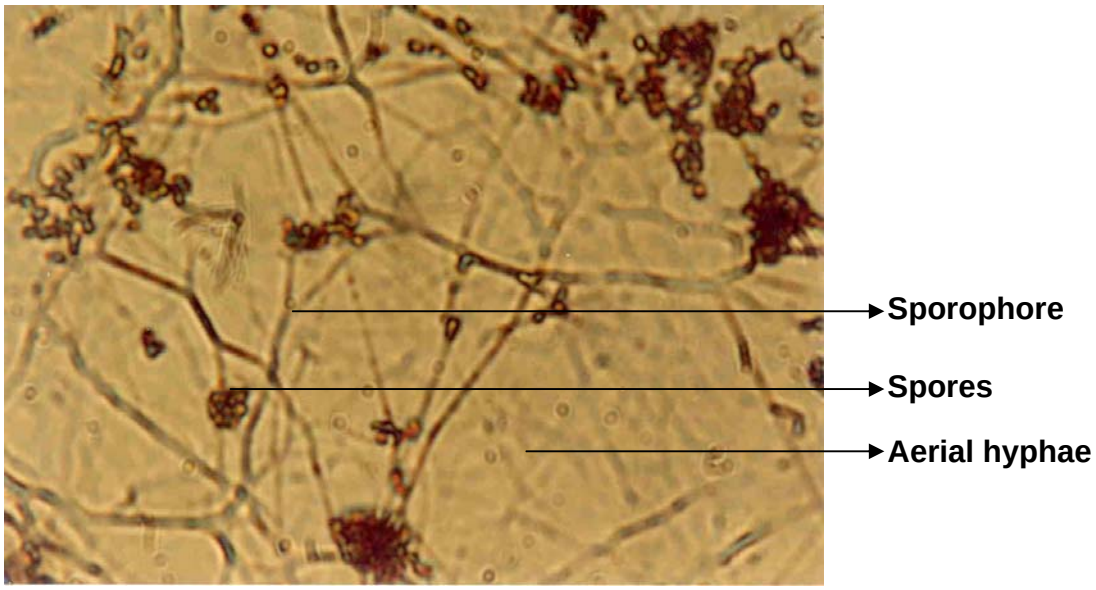


Fig. 5 Morphology of *Thermomonospora* sp. under phase contrast microscope (45X) showing sporophore, spore and aerial hyphae.

The morphology of *Thermomonospora* sp. under phase contrast microscope showed dichotonomously branched sporophore, aerial hyphae with single spore. It is a filamentous thermophilic soil bacterium and important species degrading cellulose (Wilson 1995; Zhang *et al.*, 1995)

Extensive screening at NCL has resulted in the isolation of novel thermoalkalophilic actinomycetes, a novel alkalothermophilic *Thermomonospora* sp. from soil from Barabanki district of Uttar Pradesh (George, 2001). The organism grows at 50°C and at pH 9 and produces extracellular CMCase (Carboxymethyl cellulase) maximally after 120h.(Fig.6 the xylanases from this organism have been studied in detail (George *et al.*, 2001) while cellulases have shown to be useful in textile industry (Anish *et al.*, 2007).

Cellulolytic enzyme from *Thermomonospora* sp. was studied in the present investigation

REFERENCES

Andre G., Kanchanawong, P., Palma, R., Cho H., Deng, X., Irwin, D., Himmel M.E., Wilson, D.B., and Brady J. W., 2003. Computational and experimental studies of the catalytic mechanism of Thermobifida fusca cellulase Cel6A (E2), Protein Eng., Vol. 16 (2), 125-134.

Anish, R., Rahman, M.S., and Mala Rao, 2007, Application of cellulase from an Alkalothermophilic Thermomonospora sp. In Biopolishing of Denims. Biotehnol. Bioengg. 96: 48-56.

Bayer, E. A., Morag, E., and Lamed, R. 1994, The cellulosome a treasure-trove for biotechnology, Trends Biotechnol. 12, 379-386.

Berghem, L. E., Petterson, L. G., 1974. The Mechanism of Enzymatic Cellulose Degradation: Isolation and some properties of a β-Glucosodase from Trichoderma viride, Eur. J. Biochem.. 46, 293-300.

Bhat, K. M., and Wood, T. M. 1992, The cellulase of anaerobic bacterium Clostridium thermocellums isolation, dissociation, and reassociation of the cellulosome, Carbohydr. Res. 227, 293-300.

Bhat, M., 2000, Cellulases and related enzymes in biotechnology. Biotechnol Adv. 18, 355-383

Bhat, S., Goodenough, P. W., Bhat, M. K., and Owen, E, 1994, Isolation of four major subunits from Clostridium thermocellum cellulosome and their synergism in the hydrolysis of crystalline cellulose, Int. J. Biol. Macromol. 16, 335-342.

Bhikhabhai, R., Johnsson, G, Petterson, G, 1984, Isolation of cellulolytic enzymes from Trichoderma reesei QM 9414 . Appl. Biochem, 6, 336-345.

Boraston, A.B., D. N. Bolam, H. J. Gilbert and G. J. Davies. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition Biochem. J. 382: 769-781.

Bourne, Y., and B. Henrissat. 2001. Glycoside hydrolases and glycosyltransferases: families and functional modules. Curr. Opin. Struct. Biol. 11:593-600.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

Buchert, J., J. Pere, L. Oijusluoma, L. Rahkamo, L. Viikari. 1997. Cellulases - tools for modification of cellulosic materials. In: Niches in the World of Textiles World Conference of the Textile Institute, Manchester, England, 284–290.

Calza, R. E., D. C. Irwin, and D. B. Wilson. 1985. Purification and characterization of two β-1, 4-endoglucanases from Thermomonospora fusca Biochem. 24:7797-7804.

Charnock, S. J., D. N. Bolam, D. Nurizzo, L. Szabo, V. A. McKie, H. J. Gilbert and G. J.Davies. 2002. Promiscuity in ligand-binding: the three-dimensional structure of a Piromyces carbohydrate-binding module, CBM29-2, in complex with cello- and mannohexaose. Proc. Natl. Acad. Sci. USA. 99:14077-14082.

Charnock, S. J., D. N. Bolam, J. P. Turkenburg, H. J. Gilbert, L. M. Ferreira, G. J. Davies, and C. M. Fontes. 2000. The X6 "thermostabilizing" domains of xylanases are carbohydrate-binding modules: structure and biochemistry of the Clostridium thermocellum X6b domain. Biochem. 39: 5013–5021.

Chen, J., and P. J. Weimer. 2001. Competition among three predominant ruminal cellulolytic bacteria in the absence or presence of non-cellulolytic bacteria. Microbiol. 147:21–30.

Collins T., M. Meuwis, C. Gerday and G. Feller. 2003. Activity, Stability and Flexibility in Glycosidases Adapted to Extreme Thermal Environments, J. Mol. Biol. 328: 419–428.

Coutinho P.M. and B. Henrissat.1999. The modular structure of cellulases and other carbohydrate-active enzymes: an integrated database approach. In Genetics, Biochemistry and Ecology of Cellulose Degradation. Ohmiya K., K.Hayashi, K. Sakka, Y. Kobayashi, S. Karita & T. Kimura eds, Uni Publishers Co., Tokyo : 15-23

Crawford, D. L. and E.McCoy. 1972. Cellulases of Thermomonospora fusca and Streptomyces thermodiastaticus Appl. Microbiol. 24, 150-152.

Czjzek, M., D. N. Bolam, A.Mosbah, J.Allouch, C. M. Fontes, L. M. Ferreira et al. 2001. The location of the ligand-binding site of carbohydrate binding modules that have evolved from a common sequence is not conserved. J. Biol. Chem. 276: 48580–48587

Davies, G., M. L. Sinnott, and S. G. Withers. 1997. Glycosyl Transfer. In Comprehensive Biological Catalysis Sinnott, M. L., ed., Academic Press, London. 1: 119-209.

Dey, D., J. Hinge, A. Shendye, M. Rao. 1992. Purification and properties of extracellular endoxylanase from an alkalophilic thermophilic Bacillus sp. Can. J. Microbiol. 38: 436-442.

Ding, S. Y., E. A. Bayer, D. Steiner, Y. Shoham and R. Lamed. 1999. A novel cellulosomal scaffoldin from Acetivibrio cellulolyticus that contains a family 9 glycosyl hydrolase, J. Bacteriol. 181: 6720-6729.

Ding, S. Y., E. A. Bayer, D.Steiner, Y. Shoham, and R. Lamed. 2000. A scaffoldin of the Bacteroides cellulosolVens cellulosome that contains 11 type II cohesins, J. Bacteriol. 182: 4915- 4925.

Ding, S. Y., M. T. Rincon, R. Lamed, J. C. Martin, S. I. McCrae, V.Aurilia, Y. Shoham, E. A.Bayer, and H. J. Flint. 2001. Cellulosomal scaffoldin-like proteins from Ruminococcus flavefaciens, J. Bacteriol. 183: 1945-1953.

Doi, R. H., M. Goldstein, S. Hashida, J. S.Park and M. Takagi. 1994. The Clostridium celluloVorans cellulosome. Crit. Rev. Microbiol. 20: 87-93.

Fennington G., D. Neubaer and F. Stutzenberger. 1983. Enhanced cellulase production in mutants of Thermomonospora Adenosine 3, 5-cyclic monophosphate levels in *Thermonomonspora curvata* during cellulase biosynthesis Biotechnol. Bioengg. 25: 2271-2276.

Fennington G., D. Neubaer, and F. Stutzenberger. 1983. Enhanced cellulase production in mutants of *Thermomonospora* Adenosine 3, 5-cyclic monophosphate levels in *Thermonomonspora curvata* during cellulase biosynthesis Biotechnol. Bioengg. 25: 2271-2276.

Fennington G., D. Neubaer, and F. Stutzenberger. 1984. Cellulase biosynthesis in a catabolite repression-resistant mutant of *Thermomonospora curvata.* Appl. Environ. Microbiol. 47: 201-204.

Fondevila, M., and B. A. Dehority. 1994. Degradation and utilization of forage hemicellulose by rumen bacteria, singly and in coculture or added sequentially. J. Appl. Bacteriol. 77: 541–548.

Forsberg, C.W., T. J. Beveridge, and A. Hellstrom. 1981. Cellulase and Xylanase Release from *Bacteroides succinogenes* and its importance in the Rumen Environment. Appl. Environ Microbiol. 42: 886–896.

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

Garfin, D.E., 1990. One –Dimensional Gel Electrophoresis, Methods Enzymol. 182: 425-441.

George S. P. and M. B. Rao. 2001. Conformation and polarity of the active site of xylanase I from Thermomonospora sp. as deduced by fluorescent chemoaffinity labeling: Site and significance of a histidine residue Eur. J. Biochem. 268: 2881-2888.

George, S., A. Ahmad and M.Rao. 2001. Studies on carboxymethyl cellulase produced by an alkalothermophilic actinomycete. Bioresource Technol**.** 77: 171-175.

Ghangas, G. S. and D. B. Wilson. 1987. Expression of a Thermomonospora fusca Cellulase Gene in *Streptomyces lividans* and *Bacillus subtilis.* Appl. Environ. Microbiol. 53: 1470-1475.

Gilkes, N. R., B. Henrissat, D. G. Kilburn, R. C., Jr. Miller, and R. A.Warren. 1991. Domains in microbial beta-1, 4-glycanases: sequence conservation, function, and enzyme families. Microbiol. Mol. Biol. Rev. 55: 303-315.

Gilkes, N. R., E. Jervis, B. Henrissat, B. Tekant, R. C., Jr. Miller, R. A. Warren, and D. G. Kilburn. 1992. The adsorption of a bacterial cellulase and its two isolated domains to crystalline cellulose. J. Biol. Chem. 267: 6743-6749.

Gilkes, N. R., R. A. Warren, R. C., Jr. Miller, and D. G. Kilburn 1988. Precise excision of the cellulose binding domains from two Cellulomonas fimi cellulases by a homologous protease and the effect on catalysis J. Biol. Chem. 263: 10401-10407.

Grigorevski de Lima A. L., R. P. Nascimento, E. P. S. Bon and R. R. R. Coelho 2005. *Streptomyces drozdowiczii* cellulase production using agro-industrial by-products and its potential use in the detergent and textile industries. Enz Microb. Technol. 37: 272-277.

Hagerdal B.G.R., J.D. Ferchak and E.K. Pye 1978. Cellulolytic enzyme system of *Thermoactinomyces* sp. grown on microcrystalline cellulose Appl. Environ. Microbiol. 36: 606-612.

Hakamada Y., K. Endo, S. Takizawa, T. Kobayashi, T. Shirai, T. Yamane, S. Ito, 2002. Enzymatic properties, crystallization, and deduced amino acid sequence of an alkaline endoglucanase from Bacillus circulans .Biochemica et Biophysica Acta. 1570. 174-180.

Hayashida S, K. Otta, K. Mo 1988. Cellulases of *Humicola insolen*s and *Humicola grisea.* Methods Enzymol 160: 323–32.

Henrissat B. and A. Bairoch 1996. Updating the sequence-based classification of glycosyl hydrolases. Biochem J. 316: 695-696.

Henrissat, B. and A. Bairoch. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. Biochem. J. 293: 781-788.

Henrissat, B., M. Claeyssens, P. Tomme, L. Lemesle and J.P. Mornon. 1989. Gene, Cellulase families revealed by hydrophobic cluster analysis. 81: 83-95.

Horikoshi, K. 1996. Alkaliphiles: from an industrial point of view. FEMS. Microbiol. Rev. 18: 259-270.

Idogaki H. and Y. Kitamoto. 1992. Purification and some properties of a carboxymethyl cellulase from *Coriolus versicolor*. Biosci. Biotech. Biochem. 56: 970–1.

Ingram L.O. and J.B. Doran. 1995. Conversion of cellulosic materials to ethanol FEMS Microbiology Reviews. 16: 235-241.

Irwin D, L. Walker, M. Spezio and D Wilson. 1993. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects Biotech Bioengg. 42:1002 - 1013.

Ishaque, M., and D. Kluepfel. 1980. Cellulase complexes of a mesophilic *Streptomyces* strain Can. J. Microbiol. 26,183-189.

Johnson, P.E., M.D. Joshi, P. Tomme, D.G. Kilburn, L.P. McIntosh. 1996. Structure of the N-terminal cellulose binding domain of *Cellulomonas fimi* CenC determined by nuclear magnetic resonance spectroscopy. Biochemistry 35: 14381-14394.

Jung E.D., G. Lao, D. Irwin, B.K. Barr, A. Benjamin and D.B. Wilson. 1993. DNA sequences and expression in *Streptomyces lividans* of an exoglucanase gene and an endoglucanase gene from *Thermomonospora fusca* Appl. Environ. Microbiol. 59: 3032-3043.

Kirby J., J.C. Martin, A.S. Daniel, and H.J. Flint. 1997. Dockerin like sequences in cellulases and xylanases from the rumen cellulolytic bacterium *Ruminococcus flavefaciens*. FEMS Microbiol. Lett. 149: 213 – 219

Kleman-Leyer K. M and T. K. Kirk. 1994. Three native cellulose-depolymerizing endoglucanases from solid–substrate cultures of the brown rot fungus *Meruliporia* (*Serpula*) *incrassata*. Appl Environ Microbiol. 60: 2839–45.

Kormos, J., P. E. Johnson, E. Brun, P. Tomme, L. P. McIntosh, C. A. Haynes, and D. G Kilburn. 2000. Binding site analysis of cellulose binding domain CBD(N1) from endoglucanse C of *Cellulomonas fimi* by site-directed mutagenesis. Biochemistry. 39: 8844–8852.

Koshland, D. E. 1953. Stereochemistry and the mechanism of enzymatic reactions. Biol. Rev. 28: 416-436.

Kraulis, P., M. Clore, M. Nilges, T.A. Jones, G. Pettersson, J. Knowles, and A. Gronenborn. 1989. Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. Biochemistry 28: 7241-7257.

Kubicek, C. P., and M. E. Penttila¨. 1998. Regulation of production of plant polysaccharide degrading enzymes by *Trichoderma*, *In* Harman G. E. and C. P. Kubicek (ed.), *Trichoderma* and *Gliocladium*, vol. 2 ed. Taylor & Francis Ltd., London, United Kingdom. 49–72.

Lamed, R., E. Setter, and E. A. Bayer. 1983a. Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*, *J. Bacteriol.* 156: 828-36.

Lamed, R., R. Kenig, E. Setter and E. A. Bayer 1983b. The cellulosome: a discrete cell surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulolose binding and various cellulolytic activities, *Biotechnol. Bioeng. Symp.* 13: 163-181.

Lao G, G.S. Ghangas, E.D. Jung and D.B. Wilson. 1991. DNA sequences of three beta-1,4-endoglucanase genes from *Thermomonospora fusca*. J Bacteriol 173: 3397–3407.

Lehtio, J., J. Sugiyama, M.Gustavsson, L. Fransson, M. Linder and T. T. Teeri. 2003. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. Proc. Natl Acad. Sci. USA, 100: 484–489.

Lejeune, A., S. Courtios, and C. Colson, 1988. Characterization of an Endoglucanase from *Pseudomonas fluorescens* subsp. *cellulosa* Produced in *Escherichia coli* and Regulation of the Expression of Its Cloned Gene Appl. Environ. Microbiol. 54: 302-308.

Lin, E. and D. B. Wilson, 1987. Regulation of β 1,4-endoglucanase synthesis in *Thermomonospora fusca*. Appl. Environ. Microbiol. 53: 1352-1357.

Linder, M., G. Lindeberg, T.Reinikainen, T. T. Teeri and G. Pettersson. 1995. The difference in affinity between two fungal cellulose-binding domains is dominated by a single amino acid substitution. FEBS Lett. 372: 96–98.

Lynch, J. M. 1988. The terrestrial environment, p.103–131. *In* Lynch J. M. and J. E. Hobbie (ed.), Microorganisms in action: concepts and applications in microbial ecology, 2nd ed., Blackwell Scientific Publishers, Oxford, United Kingdom.

Lynd, L.R., J.H. Cushman, R.J. Nicholas and C.E. Wyman. 1991. Fuel ethanol from cellulosic biomass. Science. 251: 1318-1323.

Lynd, L.R., P.J. Weimer, W.H. Zyl, and I.S. Pretorius 2002, Microbial Cellulose Utilization: Fundamentals and Biotechnology, Micro. Mol. Biol. Rev., 66: 506-577.

MacKenzie CR, D. Bilous and K. G. Johnson. 1984. Purification and characterization of an exoglucanase from *Streptomyces flavogriseus*. Can J Microbiol 30: 1171–1178.

MacKenzie, C. R., D. Bilous, and K. G. Johnson. 1984. *Streptomyces flavogriseus* cellulase: Evaluation under various hydrolysis conditions Biotechnol. Bioengin. 26, 590-594.

Mandels, M., F. W. Parrish, and E. T. Reese. 1962. Sophorose as an inducer of cellulase in *Trichoderma reesei.* J. Bacteriol. 83: 400–408.

Margolles-Clark, E., M. Ilme´n, and M. Penttila¨. 1997. Expression patterns of ten hemicellulase genes of the filamentous fungus *Trichoderma reesei* on various carbon sources. J. Biotechnol. 57:167–179.

McCarter, J. D. and S. G. Withers. 1994. Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. 4: 885-892.

McCarthy, A.J. and T. Cross. 1984. A Taxonomic Study of *Thermomonospora* and other Monosporic Actinomycetes. J. Gen.Microbiol. 130: 5-25.

McLean, B. W., M. R. Bray, A. B. Boraston, N. R. Gilkes, C. A. Haynes and D. G. Kilburn. 2000. Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. Protein Eng. 13: 801–809.

McSweeney, C. S., A, Dulieu, and R. Bunch. 1998. *Butyrivibrio* spp. And other xylanolytic microorganisms from the rumen have cinnamoyl esterase activity. Anaerobe. 4: 57–65.

Merril,C.R., 1990. Gel- Staining Techniques, Methods Enzym. 182: 477- 485.

Meyer, H. P. and A. E. Humphrey. 1982 Cellulase production by wild and a new mutant strain of *Thermomonospora* sp. Biotechnol. Bioengg. 24: 1901-1904.

Miettinen-Oinonen A. and P. Suominen. 2002. Enhanced production of *Trichoderma reesei:* Endoglucanases and use of the new cellulase Preparations in producing the stonewashed effect on denim fabric. Applied and Environmental Microbiology. 68: 3956–3964.

Miller, G.M., 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31: 426-428.

Morag, E., E.A. Bayer and R. Lamed. 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose degrading enzymes. J.Bac. 172: 6098 – 6105

Moreira, A. R., J. A. Phillips and A. E. Humphrey. 1981. Production of cellulases by *Thermomonospora* sp. Biotechnol. Bioengg. 23: 1339-1347.

Moreira, A. R., J. A. Phillips and A. E. Humphrey. 1981. Production of cellulases by *Thermomonospora* sp. Biotechnol. Bioengin. 23: 1339-1347.

Murao S., R. Sakamoto and M. Arai 1988. Cellulases of *Aspergillus aculeatus*. Methods Enzymol 160:274–99.

Murashima K, T. Nishimura, Y. Nakamura, J. Koga, T. Moriya, N. Sumida et al. 2002. Purification and characterization of new endo-1,4--d-glucanases from *Rhizopus oryzae*. Enzyme Microb Technol 30: 319–20.

Nagy, T., P. Simpson, M. P. Williamson, G. P. Hazlewood, H. J. Gilbert and L. Orosz. 1998. All three surface tryptophans in type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. FEBS Letters, 429: 312–316.

Nakamura Y, T. Moriya, Y. Baba, K. Yanai, N. Sumida, T. Nishimura et al. 2001. Endoglucanase and cellulase preparations containing the same. European Patent 1123974.

Nakamura, H., H. Kubota, T. Kono, A. Isogai and F. Onabe. 2001. Modification of pulp properties by cellulase treatment and application of cellulase to wastepaper deinking and mechanical pulp refining. In: 68th Pulp and Paper Research Conference. Proceedings of the Pulp and Paper Research Conference, 18–19/06, Japan. 2–5.

Niehaus, F., C. Bertoldo, M.Kahler and G.Antranikian. 1999. Extremophiles as a source of novel enzymes for industrial applications. Appl. Microbiol. Biotechnol. 51: 711–729.

Notenboom, V., A. B. Boraston, S. J. Williams, D. G. Kilburn, and D. R. Rose. 2002. High-resolution crystal structures of the lectin-like xylan binding domain from *Streptomyces lividans* xylanase 10A with bound substrates reveal a novel mode of xylan binding. Biochemistry, 41: 4246–4254.

Oberkotter, L. V. and F. A. Rosenberg. 1978. Extracellular endo-beta-1,4-glucanase in *Cellvibrio vulgaris*. Appl. Environ. Microbiol. 36:205-209.

Okada G. 1988. Cellulase of *Aspergillus niger*. Methods Enzymol. 160:259–64.

Okazaki W., T. Akiba, K. Horikoshi, and R. Akahoshi. 1984. Production and properties of two types of xylanases from alkalophilic thermophilic *Bacillus* sp. Appl. Microbiol. Biotechnol. 19:335-340.

P. M.Coutinho and B. Henrissat. 1999. Carbohydrate active enzymes: an integrated database approach. In Recent Advances in Carbohydrate Bioengineering. Gilbert, H. J., G. J. Davies, B.Henrissat and B. Svensson eds., The Royal Society of Chemistry, Cambridge. 3–12.

Pages, S., A. Belaich, C. Tardif, C. Reverbel-Leroy, C. Gaudin, and J. P. Belaich, 1996. Interaction between the endoglucanase CelA and the scaffolding protein CipC of the *Clostridium cellulolyticum* cellulosome. J. Bacteriol. 178:2279-2286.

Parry N. J., D. E. Beever, E. Owen, W. Nerinckx, M. Claaeyssens and J. V. Beeumen. 2002. Biochemical characterization and mode of action of a thermostable endoglucanase purified from *Thermoascus aurantiacus*. Arch Biochem Biophys. 404:243–53.

Poulsen O. M. and L. W. Peterson. 1988. Growth of *Cellulomonas* sp. ATCC 21399 on different polysaccharides as sole carbon source induction of extracellular enzymes. Appl. Mirobiol. Biotechnol. 29:480 – 484

Rajaram, S., and A. Varma. 1990. Production and characterization of xylanase from *Bacillus thermoalkalophilus* grown on agricultural wastes. Appl. Microbiol. Biotechnol. 34: 141-144.

Robson, L. M. and G. H. Chambliss. 1989. Review: Cellulases of bacterial origin. Enzyme Microb. Technol.11:626-644

Rodríguez, H., F. Alea, and E. Kyslı´kova. 1996. Regulation of cellulolytic activity in *Cellulomonas* sp. IIbc. Bioresource Technol. 55:79–82.

Rouau X, and M. J. Foglietti 1985. Purification and partial characterization of three endo-glucanases from *Dichomitus squalens*. Carbohydr. Res. 142:299–314.

Saha B. C. 2004. Production, purification and properties of endoglucanase from a newly isolated strain of *Mucor circinelloides* Process Biochemistry 39:1871–1876

Schell, M. A. 1987. Purification and Characterization of an Endoglucanase from *Pseudomonas solanacearum*. Appl. Environ. Microbiol. 53:2237-2241

Shen H., M. Schmuck, I. Pilz, N. R. Gilkes, D. G. Kilburn, R. C. Miller Jr and R. A. Warren. 1991. Deletion of the linker connecting the catalytic and cellulose-binding domains of endoglucanase A (CenA) of *Cellulomonas fimi* alters its conformation and catalytic activity. J. Biol. Chem. 266:11335-11340.

Simpson, H. D. and F. Barras. 1999. Functional analysis of the carbohydrate-binding domains of *Erwinia chrysanthemi* Cel5 (endoglucanase Z) and an Escherichia coli putative chitinase. J. Bacteriol. 181:4611–4616.

Simpson, P. J., H. Xie, D. N. Bolam, H. J. Gilbert, and M. P. Williamson. 2000. The structural basis for the ligand specificity of family 2 carbohydrate-binding modules. J. Biol. Chem. 275:41137–41142.

Simpson, P. J., S. J. Jamieson, M. Abou-Hachem, E. N. Karlsson, H. J. Gilbert, O. Holst and M. P. Williamson. 2002. The solution structure of the CBM4-2 carbohydrate binding module from a thermostable *Rhodothermus marinus* xylanase. Biochemistry. 41:5712–5719.

Sinnott, M. L. 1990. Catalytic mechanism of enzymic glycosyl transfer. Chem. Rev. 90:1171-1202.

Spiridonov, N. A. and D. B. Wilson. 1999. Characterization and cloning of *celR*, a transcriptional regulator of cellulase genes from *Thermomonospora fusca*. J. Biol. Chem. 274:13127–13132.

Srisodsuk M, T. Reinikainen, M. Penttila and T. T. Teer. 1993. Role of the interdomain linker peptide of *Trichoderma reesei* cellobiohydrolase I in its interaction with crystalline cellulose J. Biol. Chem. 268:20756-20761

Steenbakkers, P. J., X. L. Li, E. A. Ximenes, J. G. Arts, H. Chen, L. G. Ljungdahl, and H. J. Op Den Camp. 2001. Noncatalytic docking domains of cellulosomes of anaerobic fungi. J. Bacteriol. 183:5325-5333.

Stutzenberger, F. J. 1971. Cellulase Production by *Thermomonospora curvata* Isolated from Municipal Solid Waste Compost. Appl. Microbiol. 22:147-152

Stutzenberger, F. J., 1972. Cellulolytic activity of *Thermomonospora curvata* nutritional requirements for cellulase production. Appl. Microbiol. 24:77-82.

Szabo, L., S. Jamal, H. Xie, S. J. Charnock, D. N. Bolam, H. J. Gilbert and G. J. Davies. 2001. Structure of a family 15 carbohydrate-binding module in complex with xylopentose. evidence that xylan binds in an approximate 3-fold helical conformation. J. Biol. Chem. 276:49061–49065.

Teather R. M. and P. J. Wood. 1982. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. Appl. Environ. Microbiol. 43:777–780.

Tokatlidis, K., P. Dhurjati and P. Beguin. 1993. Properties conferred on *Clostridium thermocellum* endoglucanase CelC by grafting the duplicated segment of endoglucanase CelD. Protein Eng. 6:947-952.

Tokatlidis, K., S. Salamitou, P. Beguin, P. Dhurjati and J. P. Aubert. 1991 Interaction of the duplicated segment carried by *Clostridium thermocellum* cellulases with cellulosome components. FEBS Lett. 291:185-188.

Tomme, P., H. Van Tilbeurgh, G. Pettersson, J. Van Damme, J. Vandekerckhove, J. Knowles, T. Teeri, and M. Claeyssens. 1988. Studies of the cellulolytic system of *Trichoderma reesei* QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. Eur. J. Biochem. 170:575-581

Tomme, P., R. A. J. Warren, R. C. J. Miller, D. G. Kilburn and N. R. Gilkes. 1995. in *Enzymatic Degradation of Insoluble Polysaccharides* (Saddler, J. N. , and Penner, M., eds). 618:142-163, American Chemical Society, Washington, D. C.

Tormo, J., R. Lamed, A. J. Chirino, E. Morag, E. A. Bayer, Y. Shoham and T. A. Steitz. 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. 15:5739–5751.

van Tilbeurgh, H., P. Tomme, M. Claeyssens, R. Bhikhabhai and G. Pettersson. 1986. Limited proteolysis of the cellobiohydrolase I from *Trichoderma reesei* Separation of functional domains FEBS Lett. 204:223-227

Van wyk, J., A. Mogale and T. Seseng. 2001 Bioconversion of wastepaper to sugars by cellulase from *Aspergillus niger*, *Trichoderma viride*. Bioresource Technology 17:34-33

Van Zyl, W. H. 1985. A Study of the cellulases produced by three mesophilic actinomycetes grown on bagasse as substrate Biotechnol. Bioengg.. 27, 1367-1373

Véronique R., C. Mirjam, S. Martin, P. Panine and B. Henrissat. 2002. Dimension, Shape, and Conformational Flexibility of a Two Domain Fungal Cellulase in Solution Probed by Small Angle X-ray Scattering J. Biol. Chem. 277:40887-40892

Voragen A. G., J. G. Beldman and F. M. Rombouts. 1988. Cellulases of a mutant strain of *Trichoderma viridie* QM 9414. Methods Enzymol. 160:243–51.

Wang, D. Z., Y. Zu, and P. Gao. 1996. Studies of the regulation of cellulase systems by ATP and cAMP in mycelial fungi. Weishengwu Xuebao. 36:12–18.

Wilson D. B. 1995. Cloning of Thermomonospora fusca cellulase genes in E. coli and S. lividans. Part B, Methods in Molecular Genetics. 6:367-374.

Wilson, D. B. 1992. Biochemistry and genetics of actinomycete cellulases. Crit. Rev. Biotechnol. 12:45–63.

Wilson, D. B. and D. C. Irwin. 1999. Genetics and properties of cellulases. Adv. Biochem. Eng. Biotechnol. 65:1–21.

Wilson, D.B. 1988. Cellulases of *Thermomonospora fusca.* In: Biomass, Part A: Cellulose and Hemicellulose. In: Wood, W.A., Kellogg, S.T. (Eds.), Met. in Enz., vol. 160. Academic Press, San Diego, CA, pp. 314-323.

Wolfe, B. R., T. A. Mudry, B. R. Glick and J. J. Pasternak. 1986. Isolation of Endoglucanase Genes from *Pseudomonas fluorescens* subsp. *cellulosa* and a *Pseudomonas* sp. Appl. Environ. Microbiol. 51:1367-1369

Wood T. M. and S. I. McCrac. 1978. The cellulase of *Trichoderma koningii*. Purification and properties of some endoglucanase components with special reference to their action on cellulose when acting alone and in synergism with the cellobiohydrolase. Biochem J. 171:61–72.

Wood, W. E., D. G. Neubaer and F. J. Stutzenberger. 1984. Cyclic AMP levels during induction and repression of cellulase biosynthesis in *Thermomonospora curvata*. J. Bacteriol. 160:1047-1054

Wood, W. E., D. G. Neubauer and F. J. Stutzenberger. 1984. Cyclic AMP levels during induction and repression of cellulase biosynthesis in *Thermomonospora curvata*. J. Bacteriol. 160:1047–1054.

Wynne E. C. and J. M. Pemberton. 1986. Cloning of a Gene Cluster from *Cellvibrio mixtus* which Codes for Cellulase, Chitinase, Amylase, and Pectinase. Appl. Environ. Microbiol. 52:1362–1367.

Xie, H., D. N. Bolam, T. Nagy, L. Szabo, A. Cooper and P. J. Simpson. 2001. Role of hydrogen bonding in the interaction between a xylan binding module and xylan. Biochemistry. 40:5700–5707.

Xie, H., H. J. Gilbert, S. J. Charnock, G. J. Davies, M. P. Williamson, and P. J. Simpson. 2001. *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22-2: the role of conserved amino acids in ligand binding. Biochemistry. 40:9167–9176.

Xu, G.Y., E. Ong, N. Gilkes, D. Kilburn, D. R. Muhandiram, M. Harris-Brandts, J. Carver, L. Kay and T. Harvey. 1995. Solution structure of a cellulose-binding domain from *Cellulomonas fimi* by nuclear magnetic resonance spectroscopy. Biochemistry. 34:6993-7009.

Xu, Q., E. A. Bayer, M. Goldman, R. Kenig, Y. Shoham and R. Lamed. 2004. Architecture of the *Bacteroides cellulosolvens* cellulosome: description of a cell surface-anchoring scaffoldin and a family 48 cellulase. J. Bacteriol. 186:968-977.

Xu, Q., W. Gao, S. Y. Ding, R. Kenig, Y. Shoham, E. A. Bayer and R. Lamed. 2003. The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. J. Bacteriol. 185:4548-4557.

Yamane K., H. Suzuki, M. Hirotani, H. Ozawa, and K. Nisizawa. 1970. Effect of nature and supply of carbon sources on cellulase formation in *Pseudomonas fluorescens* var. cellulosa J. Biochem. (Tokyo) 67:9-18.

Yamane, K., H. Suzuki, and K. Nisizawa. 1970. Purification and properties of extracellular and cell-bound cellulase components of *Pseudomonas fluorescens* var. cellulosa J. Biochem. 67:19-35

Yu, I. and R. E. Hungate. 1979. The extracellular cellulases of *Ruminococcus albus*. Ann. Rech. Vet. 10:251-254

Zhang S., B. K. Barr and D. B. Wilson. 2000. Effects of noncatalytic residue mutations on substrate specificity and ligand binding of *Thermobifida fusca* endocellulase Cel6A. Eur. J. Biochem. 267:244 - 252.

Zhang S., D. C. Irwin and D. B. Wilson. 2000. Site-directed mutation of noncatalytic residues of *Thermobifida fusca* exocellulase Cel6B. Eur. J. Biochem. 267:3101 - 3115.

Zhang S., G. Lao and D. B. Wilson. 1995. Characterization of a *Thermomonospora fusca* exocellulase. Biochemistry. 34:3386-3395.

Zhou, L., K Yeung and C. Yuen. 2001. Combined cellulase and wrinkle free treatment on cotton fabric. J. Dong Hua University. 18:11–15.

CHAPTER-II

Purification and Biochemical Characterization of 1,4-β-D-Glucan Glucanohydrolase

ABSTRACT

In our laboratory a novel alkalothermophilic actinomycete *Thermomonospora* sp. was isolated from self-heating compost. The strain produces extracellular cellulases and xylanases. One of the extracellular cellulases (1,4-β-D-Glucan Glucanohydrolase E.C.No.3.2.1.4) was purified to homogeneity from the culture filtrate by fractional ammonium sulphate precipitation (30-55%), DEAE Sephadex A50 and Biogel P100 column chromatography. The molecular weight of the purified enzyme was 14.2 kDa by MALDI - TOF analysis and is in agreement with SDS PAGE. The purified enzyme had a pI of 4.33 with pH and temperature optima 7 and at 50°C respectively. The CMCase was highly thermostable at 50°C for 12 h and had a half-life of 1 h at 60°C and 20 min at 80°C. The CMCase was stable in a wide pH range from 5 to 8. The secondary structure of the enzyme is made up of α helix regular 0.588, α helix distorted 0.226, β sheet regular0.001, β sheet distorted 0.002, β turn 0.062 and unordered structure 1.04.

The purified enzyme had both endocarboxymethyl cellulase and endoxylanase activities. During purification, the ratio of carboxymethyl cellulase to xylanase activity remained constant indicating that both the activities reside on the same protein. Therefore, the question arises of whether the two activities of the enzyme are localized on the same or on two different active sites. The Km for the two activities were widely different, Viz. 6 mg/ml for CMC and 3 mg/ml for Xylan. The kinetic analysis by the method of Keleiti was employed to analyze whether one or two sites are responsible for hydrolysis of CMC and xylan. The experimental data coincides well with the theoretical values calculated for the case of a single active site. Conformation and microenvironment at the active site was probed with fluorescent chemo affinity labeling using o-phthalaldehyde as the chemical initiator. Formation of isoindole derivative resulted in complete inactivation of the enzyme to hydrolyze both xylan and CMC as judged by fluorescence studies corroborating a single active site for the hydrolysis of xylan and CMC.

INTRODUCTION

Cellulose is a fibrous, insoluble and crystalline polysaccharide made up of β-1-4 linked D-glucopyranosyl residues and constitutes the major structural component of plant cell walls. It is the most abundant biopolymer on earth and approximately 10^7 tons is synthesized per annum. Cellulose can be degraded by the coordinated action of cellulolytic enzymes, such as endoglucanase, cellobiohydrolase, and β-glucosidase (Bhat, 2000). Xylan is the second most abundant biopolymer after cellulose and the major hemicellulosic polysaccharide found in plant cell walls. It is a heterogeneous polysaccharide consisting of β-1-4 linked D-xylosyl residues on the backbone but also containing arabinose, glucuronic acid and arabinoglucuronic acids linked to D- xylose backbone. Xylanases are enzymes that degrade the xylan backbone into small oligomers (Kulkarni *et al*., 1999). Most microbes thought to play a prominent role in cellulose hydrolysis in nature, have evolved strategies that bring the cell close to the cellulose surface and give the cellulolytic organism "first access" to hydrolysis products (Robson and Chambliss, 1989; Lynd *et al*., 2002). The removal of certain polysaccharides by one species or group of microbes may improve the accessibility of a second group to cellulose (Poulsen and Peterson, 1988). In this regard, it is interesting that some cellulolytic bacteria can actively depolymerize certain hemicelluloses, particularly xylans (Wilson, 1988; Kirby *et al*., 1997; Wang *et al*., 2003). In nature, these cellulolytic species utilize xylanases and pectinases to gain access to cellulose (McSweeny *et al*., 1998). *Thermomonospora* is an important species degrading cellulose and hemicellulose in plant residues. It produces extracellular cellulases and xylanases maximally after 120 h. (McCarthy and Cross, 1984; Zhang *et al*., 1995) The present chapter describes the purification and properties of 1,4-β-D-Glucan Glucanohydrolase (TSC) with molecular weight of 14.2 kDa from *Thermomonospora* sp. Based on the kinetic analysis, experimental evidence showed that the purified enzyme has a single active site for the substrates CMC and xylan.

MATERIALS AND METHODS

Materials

DEAE Sephadex A50, Biogel P100, EDTA were obtained from Sigma Chemical Co. USA. Molecular weight markers for gel filtration chromatography, and SDS PAGE were obtained from Biorad. Sephacryl S200 was purchased from Amersham Pharmaceuticals. All buffer salts were from standard commercial sources and of highest quality available. All other chemicals used in the present work were of analytical grade.

Methods

Microorganism and culture conditions

Thermomonospora sp. is an alkalothermophilic actinomycete having optimum growth at pH 9 and 50°C used in this study was isolated from self-heating compost from Barabanki district, Uttar Pradesh, India (George *et al*., 2001)

Cellulase production in liquid culture

Culture was grown in Reese's Medium with 2% Wheat Bran, initially in 10ml for 48h at 50°C on a shaker at 200 rpm. The seed culture was inoculated into 100 ml Reese's Medium with 4% Cellulose Powder (Composition in Table 1). The flasks were incubated at 50°C on a shaker at 200 rpm.

Table-1

Composition of Reese's media

Constituents	For 100 ml
Tween 80	0.10 ml
K_2HPO_4	0.20 g
$(NH_4)_2SO_4$	0.70 g
Urea	0.15 g
$MgSO_4$. $7H_2O$	0.03 g
Peptone	0.125 g
Yeast Extract	1.00 g
$FeSO_4.7H_2O$ (1%)	50.0 µl
$MnSO_4.H_2O$ (1%)	15.6 µl
$ZnSO_4.7H_2O$ (1%)	14.0 µl
$CoCl_2$ (1%)	20.0µl

Enzyme Assay

The enzyme activity (xylanase and caboxymethylcellulase) was measured by incubating 1 ml of assay mixture containing the suitably diluted enzyme in 0.05 M phosphate buffer pH 7 with 0.5ml of substrate (1% xylan for xylanase activity and 1% CMC for caboxymethylcellulase activity) at 50°C for 30 min. Enzyme and reagent blanks were also simultaneously incubated with the test samples. The reducing sugar released was determined by the dinitrosalicylic acid method (Miller, 1959).

One unit of xylanase activity was defined as the amount of enzyme that produced 1µmole of xylose equivalent per minute from xylan, under assay conditions. One unit of carboxymethyl cellulase activity was defined as the amount of enzyme that produced 1µmole of glucose equivalent per minute from CMC, under assay conditions. Protein concentration was determined according to the method of Bradford, using bovine serum albumin as standard (Bradford, 1976).

Production and purification of 1,4- β-D-Glucan Glucanohydrolase (TSC)

Thermomonospora sp. was grown for 120 h in a modified media containing 1% yeast extract, 4% cellulose paper powder and 0.1% Tween 80. The pH of the medium was adjusted to 9 after autoclaving with sterile 10% Na_2CO_3. The culture and insoluble cellulose paper powder were removed by centrifuging at 5,000 rpm for 10 min. All purification steps were carried out at 4°C unless otherwise stated.

Ammonium sulphate precipitation:

Protein precipitation by salting out technique using ammonium sulphate $(NH_4)_2SO_4$ was carried out with constant gentle stirring. The amount of solid ammonium sulphate to be added was calculated according to the formula of Jagannathan *et al*. (1956).

$$X = \frac{50\ (S_2 - S_1)}{1 - 0.28\ (S_2)}$$

Where X = Amount of ammonium sulphate to be added for 100 ml.

S_1 = Initial concentration of ammonium sulphate.

S_2 = Required saturation of ammonium sulphate.

The culture filtrate (200 ml) was subjected to fractional ammonium sulphate precipitation (30 to 55%). The culture filtrate was initially saturated to 30% by the slow addition of 38.8 g of ammonium sulphate with gentle stirring. The solution was allowed to stand overnight at 4°C. The precipitated solution was centrifuged at 10,000 rpm for 10 min at 4°C and the pellet was discarded while the supernatant was subjected to 55 % saturation by the addition of 23.6 g of ammonium sulphate. The precipitate was recovered by centrifuging it at 10,000 rpm for 10 min at 4°C. The pellet was dissolved in minimum amount of 0.05 M sodium phosphate buffer, pH 7 and dialyzed against 100 volumes of the same buffer with several changes for 24 h.

Ion exchange chromatography:

The dialyzed fraction was applied to DEAE-Sephadex A-50 column (6 x 20cm), previously equilibrated with 0.05 M sodium phosphate buffer at pH 7. The sample was loaded at a rate of 20 ml/h. The column was initially washed with 0.05 M sodium phosphate buffer, pH 7 at a rate of 20 ml/h to remove the unadhered and loosely adhered proteins. It was observed that cellulase does not elute out in unadhered fraction. The elution of the adhered proteins was carried out by a linear gradient of sodium chloride (0-1.0M) in 0.05 M sodium phosphate buffer, pH 7. Fractions of 5 ml each were collected and the fractions having maximum specific activity were pooled and concentrated by ultrafiltration through Amicon UM-10 membrane.

Gel filtration chromatography:

The concentrated sample was loaded onto a Sephacryl S-200 column (2.5 x 100 cm), which was equilibrated with 0.05 M, pH 7 sodium phosphate buffer. Elution was carried out by using the same eluant at a flow rate of 10 ml/hr, and 2 ml fractions were collected. The fractions having maximum specific activity were pooled and concentrated by ultrafiltration through Amicon UM-10 membrane. The purity of enzyme was checked by 12% SDS PAGE (Laemmli) followed by sliver staining.

12% SDS PAGE:

Composition of resolving gel for 10ml

3.35ml water, 2.5ml 1.5M Tris-Cl buffer pH 8.8, 0.1ml of 10% SDS, 4.0ml of Acryl amide and bisacrylamide, 50µl of freshly prepared 10% APS, 5µl TEMED.

Composition of stacking gel for 10ml

6.1 ml water, 2.5 ml 0.5 M Tris-Cl buffer pH 6.8, 0.1 ml of 10% SDS, 1.3 ml of Acryl amide and bisacrylamide, 50 µl of freshly prepared 10% APS, 5 µl TEMED.

Determination of molecular weight

The Molecular weight of the enzyme was determined by gel filtration chromatography using Sephacryl S200 previously equilibrated with equilibrated with 50 mM Phosphate buffer pH 7, which was calibrated using the following marker proteins, Albumin (67,000), Ovalbumin

(43,000), Chymotrypsinogen A (25,000) and Ribonuclease A (13,700). The void volume of the column was determined by the elution volume of Blue Dextran.

The presence of subunits in the protein was determined by 12% SDS PAGE. Molecular weight standards used were Phosphorylase b (97,400), Bovine Serum Albumin (66,200), Ovalbumin (45,000), Carbonic Anhydrase (31000), Soyabean trypsin inhibitor (21,500) and Lysozyme (14,400).

Matrix–associated laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry analysis of the purified enzyme sample was carried out using applied Biosystems instrument.

Circular dichroism measurements and determination of secondary structure

CD spectra were recorded in a Jasco –J715 spectrophotometer at ambient temperature using a cell of 1 mm path length. Replicate scans were obtained at 0.1nm resolution, 0.1 nm bandwidth and a scan speed of 50 nm/min. Spectra were average of 6 scans with the baseline subtracted spanning from 260 to 190 nm in 0.1 nm increments. The CD spectrum of 1,4-β-D-Glucan Glucanohydrolase (1 mg/ml) was recorded in 10 mM sodium phosphate buffer pH 7. The data obtained from CD spectra was used for the determination of the secondary structure of the enzyme.

Amino acid analysis

The amino acid analysis was carried out using amino acid analysis kit AccQ- Fluor supplied by Waters Corporation. Salt free lyophilized 1,4-β-D-Glucan Glucanohydrolase (50 µg) was hydrolyzed using 6 N HCl with constant boiling, in vacuum sealed hydrolyzing tubes for 24 h at 110°C. The sample was then derivatized by 6-aminoquinolyl-N- hydroxysuccinimidyl carbamate (AQC) in borate buffer at pH 9. 10 picomoles of the hydrolysate were loaded on AccQ- Tag column equipped with a fluorescent detector. Total cysteine was determined according to Boyer (1954) and total tryptophan according to Spande and Wikop (1967)

Determination of isoelectric point

The isoelectric focusing of the enzyme was determined by using Bio Rad Rotofor unit. Prior to assembly, the ion exchange membranes were equilibrated overnight in the appropriate electrolyte solutions. The anode and cathode electrolyte chambers were assembled and filled with electrolytes immediately to prevent the membrane from drying. The anode was filled with acidic electrolyte 0.1 M H_3PO_4. The cathode was filled with 0.1 M NaOH. The final concentration of Bio- lytes (pH range 2-12) was 0.5%. The sample in 5 mM phosphate buffer pH 7.0 was loaded through the ports using a 50ml syringe with 1-1/2 inch 19 gauge needle. The air bubbles were eliminated from electrode ports. The ports were connected to the chiller set at 4°C. The power supply was set to 12W constant power. The run was completed in 3.5 h. All 20 fractions were simultaneously aspirated from the cell and delivered to collection tubes. After harvesting, pH and cellulase activity of individual fractions was checked.

Activity towards different substrates

The purified enzyme was checked for its activity against different substrates. Ten milligrams of substrates such as chitin, guar gum, filter paper, gum arabic and avicel were incubated with a suitably diluted enzyme in a total volume of 1 ml in 0.05 M sodium phosphate buffer, pH 7, at 50°C for 30 min. The reducing sugars released were estimated by dinitrosalicylic acid method. The activity of the supernatant towards PNPX and PNPG was determined by incubating 1 ml of 0.3 % substrate with 0.1 ml of suitably diluted enzyme in 0.05 M phosphate buffer, pH 7, for 10 min at 40°C. The liberated *P*-nitrophenol was measured at 400 nm. The enzyme activity was calculated according to Mishra *et al.* (1984).

Influence of pH on activity and stability

The optimum pH of the enzymes were determined by estimating the enzyme activity under standard assay conditions at different pH values ranging from pH 4- 10. 0.05 M of citrate phosphate buffer (pH 4), acetate buffer (pH 5), sodium phosphate buffer (pH 6-8), and carbonate buffer (pH 9-10) were used. The pH stability of the enzyme was measured by incubating 5 IU of enzyme at 50°C in buffer of desired pH for 1h.

Influence of temperature on activity and stability

The optimum temperature of CMCase and xylanase activities of the purified enzyme was determined by assaying at different temperatures ranging from 40-100°C. The temperature stability was determined by incubating 5 IU of enzyme at different temperatures. The samples were removed at regular intervals and the residual activity was estimated under standard assay conditions.

End product analysis of hydrolysis of CMC and xylan

In order to study the pattern of hydrolysis by TSC, 5 mg of CMC was incubated with 0.5 IU of enzyme in a reaction volume of 1 ml for various time intervals (3 and 7 h) at 50°C. The reaction was terminated by the acidification of the reaction mixture with 1N HCl to pH 6.0 followed by 3 min incubation in boiling water bath. The samples were then centrifuged to remove unhydrolyzed CMC. The same procedure was carried out for hydrolysis of xylan.

HPLC

A Sugar – PAK column –300mm x 6.6mm (ID), (Waters, Analytical) was used for resolution of cello-oligosaccharides. Samples were filtered through a 0.45-micron membrane just before injection to remove any particulates. Sample of 100µl was loaded on the column. The column was then eluted with $1x10^{-4}$ M Ca- EDTA buffer at a flow rate of 0.4 ml/min at 70°C. The time period of run was 20 min. The detector was set at 30°C and the products were determined by the measurement of refractive index changes measured in mill volts. The retention times and the areas of the individual peaks after correction for the baseline were compared with the standard cello- oligosaccharides run under identical conditions. The same procedure was followed for the end product analysis of hydrolysis of xylan.

Kinetic analysis

A kinetic method elaborated by Keleti *et al.* (1987) was used to evaluate whether TSC hydrolyzing both CMC and xylan has one or two active sites. The initial velocities are calculated as µmole min^{-1} ml $^{-1}$ in the simultaneous presence of both substrates at concentrations obeying the relationship.

$$\mathbf{[CMC]\ Km^{Xylan} + [xylan]\ Km^{CMC} = K = 10\ (mg/ml)^2}$$

Where the value of K was chosen arbitrarily to achieve concentration ranges around Km values. For computations Km ^{xylan} = 3 mg/ml and Km ^{CMC} = 6 mg/ml determined separately were used. The Michaelis constants and maximum velocities (V_{max}) of the cellulase and xylanase activities of the enzyme were estimated from Lineweaver- Burk plots o their initial velocities of saccharification of the substrates.

If we consider enzyme with one active center acting concomitantly with substrate S_1 and the alternative substrate S_2, the two substrates will compete with each other for the same active site. (Scheme 1) (Keleti *et al*.,1987).

$$\mathbf{E + S_1 \rightleftharpoons ES_1 \rightarrow E + P_1} \qquad \mathbf{E + S_2 \rightleftharpoons ES_2 \rightarrow E + P_2}$$

In this case the equation for the initial velocity ($V_{0)}$ is represented by

$$\mathbf{V_0 = (V_1K_2\,[S_1] + V_2K_1\,[S_2] \,/\, K_1K_2 + K_1\,[S_2] + K_2\,[S_1])}$$

Where V_1 and V_2 are maximum velocities, and K_1 and $K_{2,}$ the dissociation constants for CMC and xylan respectively.

Assuming two different active sites for the substrates S_1 and S_2 the mechanism of action can be represented by Scheme 2

$$\mathbf{E + S_1 \rightleftharpoons ES_1 \rightarrow E+P_1} \qquad \mathbf{E +S_2 \rightleftharpoons ES_2 \rightarrow E + P_2}$$

$$\mathbf{ES_1 + S_2 \rightleftharpoons ES_1S_2 \rightarrow ES_1 + P_2} \qquad \mathbf{ES_2 + S_1 \rightleftharpoons ES_1S_2 \rightarrow ES_1 + P_1}$$

If the two active sites are independent, the initial velocity can be represented by

$$\mathbf{Vo} = \frac{\mathbf{V_1K_2\,[S_1] + V_2K_1\,[S_2] + (V_1 + V_2)\,[S_1]\,[S_2]}}{\mathbf{K_1K_2 + K_1\,[S_2] + K_2\,[S_1] + [S_1][S_2]}}$$

If the two substrates mutually influence each other's binding

$$V_0 = V_1 K_2 [S_1] + V_2 K_1 [S_2] / K_1 K_2 + K_1 [S_2] + K_2 [S_1] + [S_1] [S_2]$$

$$+$$

$$(V_1 + V_2) [S_1] [S_2] / K_1 K_3 + K_4 [S_2] + K_3 [S_1] + [S_1] [S_2]$$

Where K_3 and K_4 are the dissociation (or Michaelis) constants of S_2 and S_1 in the presence of S_1 and S_2, respectively

Chemo- affinity labeling studies using OPTA

TSC (50 μg), in 0.05 M phosphate buffer pH 7 was incubated with o- phthalaldehyde (OPTA) in methanol at 25°C. The formation of enzyme – isoindole derivative was followed spectrophotometrically by monitoring the increase in fluorescence at 420 nm with excitation wavelength fixed at 338nm. The aliquots were removed at intervals to check the residual carboxymethyl cellulase and xylanase activity of the enzyme.

RESULTS AND DISCUSSION

Purification of 1,4-β-D-Glucan Glucanohydrolase (TSC)

After 120 h of growth, the culture and insoluble cellulose paper powder was removed by centrifuging at 10000 rpm for 10 min. The clarified broth was concentrated by ultra filtration through UM 10 membranes. The concentrated broth was subjected to fractional ammonium sulphate precipitation (30-55%). The precipitate formed was recovered by centrifugation at 10000 rpm for 10 min. The precipitate was dissolved in minimum amount of 50 mM Phosphate buffer pH 7 and dialyzed against the 100 volumes of the same buffer with several changes for 24h. The dialyzed fraction was applied to DEAE Sephadex A50 previously equilibrated with 50 mM Phosphate buffer pH 7. The elution was carried out by a linear gradient of sodium chloride (0 M-1.0 M) in of 50 mM Phosphate buffer pH 7. The Fig. 1 shows fractionation pattern on DEAE Sephadex A50. The cellulase and xylanase activities were almost completely coincident with each other and with the protein peak. The Table 1 summarizes the steps of purification of TSC.

The fractions having maximum specific activity were pooled and concentrated by Amicon UM 10 Membrane and applied to Biogel P100 column. The elution was carried out by using the same eluant at a flow rate of 10 ml/hr, and 2 ml fractions were collected (Fig.2). The fractions having maximum specific activity were pooled and concentrated.

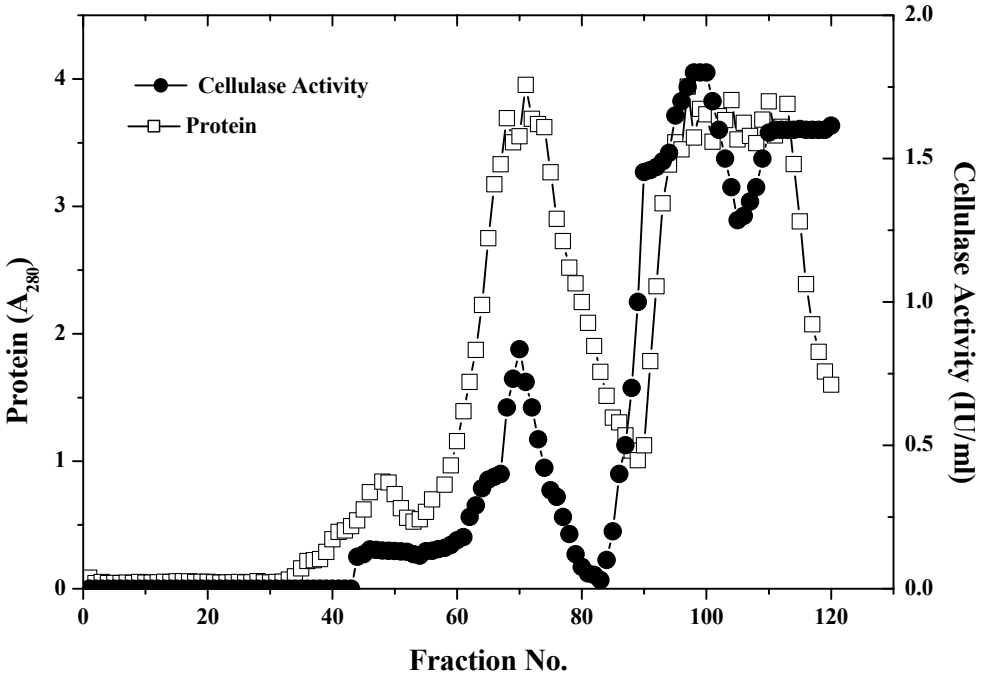


Fig. 1 Anion exchange chromatography using DEAE Sephadex A50. The dialyzed enzyme sample obtained after 30-55 % ammonium sulphate precipitation was loaded onto DEAE Sephadex A50 column and TSC along with other adhered proteins was eluted by using a linear gradient of 0 – 1M NaCl.

Determination of molecular weight

The homogeneity and purity of the enzyme was analyzed by12% SDS PAGE followed by sliver staining and the enzyme was found to be electrophoretically homogeneous, revealing a single band (Fig.3A) and the relative mass of the purified TSC was calculated to be 14.5 kDa using the calibration curve of markers on SDS PAGE (Fig.3B). The molecular weight of the purified TSC as determined by gel filtration chromatography was estimated to be 14.4 kDa (Fig. 4). The molecular weight of the purified TSC as determined by gel filtration (14.4 kDa) and SDS PAGE (14.5 kDa) were comparable, indicating the absence of subunits in the protein.

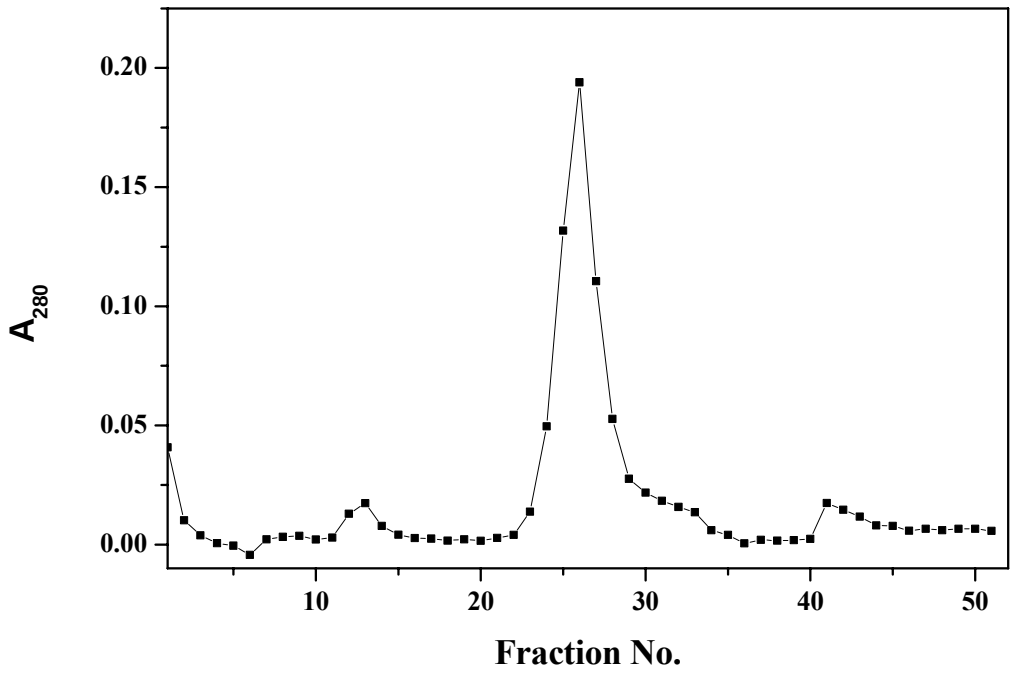


Fig. 2 Fractionation pattern of TSC obtained after DEAE Sephadex A50 on Biogel P100 column. The fractions were checked for enzyme activity and absorbance at 280 nm.

Table-2

Purification of TSC from *Thermomonospora* sp.

Step	**Total protein (mg)**	**Total Enzyme Activity (IU)**				**Ratio**		
		xylanase	**Specific activity (IU/mg)**	**cellulase**	**Specific activity (IU/mg)**	**Xylanase: cellulase**	**% Recovery**	**Fold purification**
Culture filtrate	640	11700	18.28	700	1.1	16.61	100	1
Ammonium sulphate precipitation	208	8784	42.23	528	2.54	16.63	75	2.30
DEAE Sephadex	35.4	6055	171.04	366.2	10.34	16.53	51	9.40
Biogel P100	0.25	242.52	970.09	14.7	58.79	16.49	9	53.44

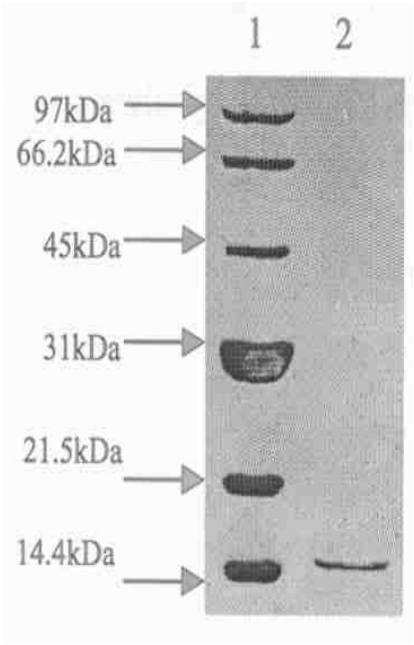


Fig. 3A 12% SDS PAGE of standard molecular marker proteins in Lane 1 and Lane 2 contained purified 1,4-β-D-Glucan Glucanohydrolase was visualized by silver staining.

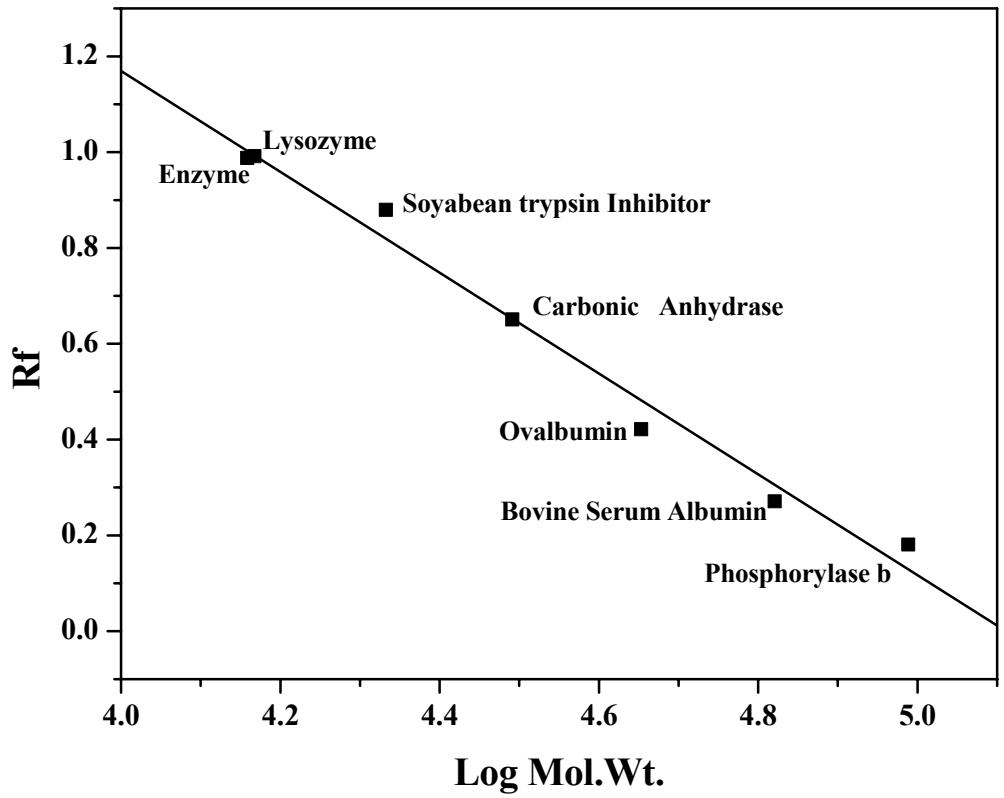


Fig.3B Calibration curve for 12% SDS PAGE. Electrophoretic mobilities of the purified TSC and of the reference proteins on SDS polyacrylamide gel were plotted versus their log molecular weights.

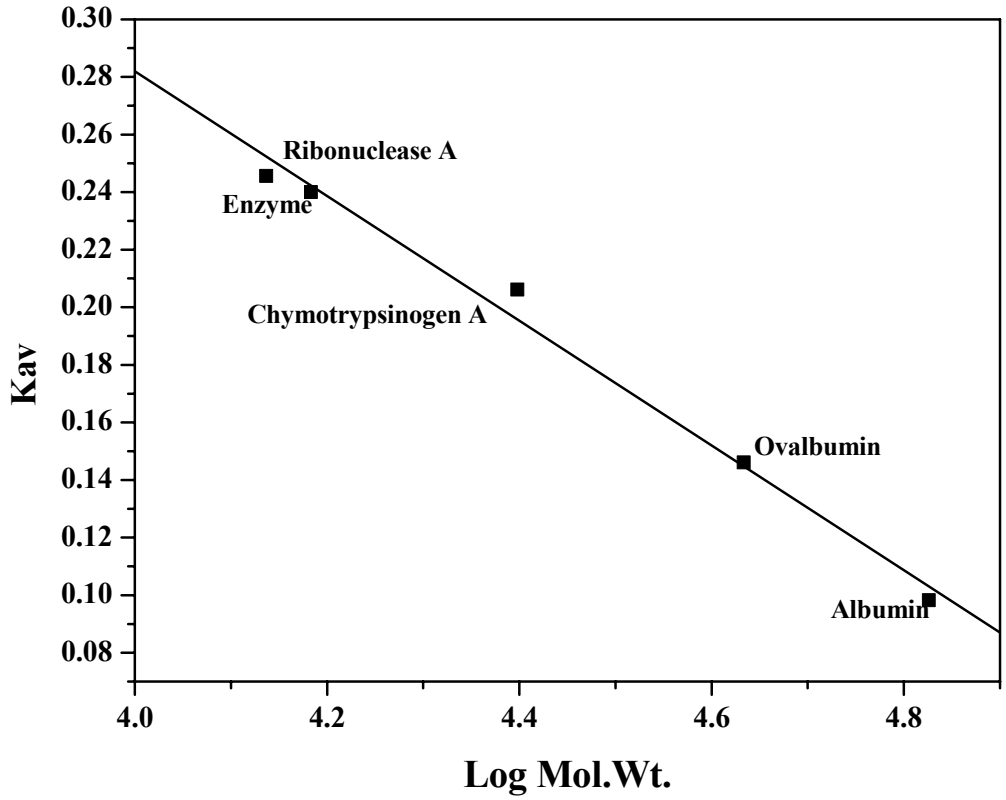


Fig. 4 Calibration curve for Gel Filtration Chromatography using Sephacryl S200. The molecular weight of TSC was (Mr: 14.4 kDa)

The purified TSC was also analyzed by MALDI-TOF mass spectrometry, which revealed with a sharp peak of an average molecular mass of 14.2 kDa confirming the homogeneity of the enzyme (Fig.5). The endogluacanases from *Thermomonospora* reported to date are of the molecular weight in the range of 30 to 108 kDa (Stutzenberger, 1971; Calza *et al*., 1985; Irwin *et al*., 1993). This is the lowest molecular weight endoglucanase reported till date from a novel alkalothermophilic *Thermomonospora.*

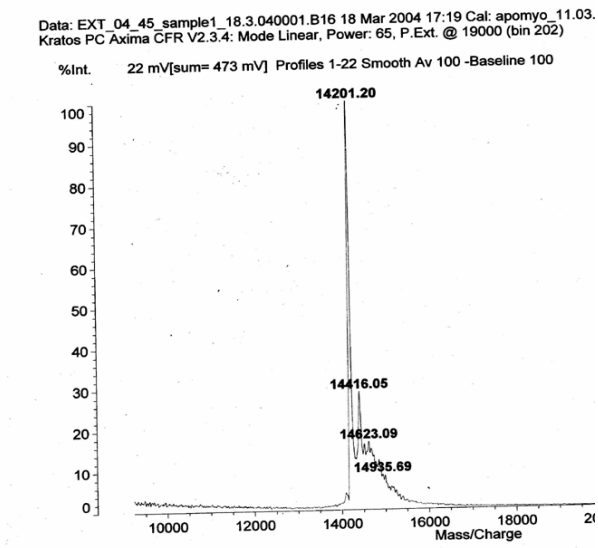


Fig. 5 MALDI-TOF mass spectrometry of purified TSC.

The amino acid composition of TSC is given in Table 3

Table-3

Amino acid composition of 1,4-β-D-Glucan Glucanohydrolase

Amino acid	No. of residues/mol
Aspartic acid and Asparagine	17
Threonine	9
Serine	8
Glutamic acid and glutamine	10
Proline	5
Glycine	13
Alanine	11
Valine	7
Methionine	2
Isoleucine	4
Leucine	11
Tyrosine	4
Phenylalanine	4
Lysine	3
Histidine	4
Arginine	5
Tryptophan	4
Cysteine	3
Total	124

The estimated molecular weight of TSC from amino acid composition is 13.6 kDa.

Circular dichroism measurements and determination of secondary structure

The CD spectrum of TSC in the far-UV region (200 - 260 nm) exhibited a strong negative ellipticities in the region 215-222 nm and a weaker one at 208 nm, characteristic of a protein having an α- helix (Fig. 6).

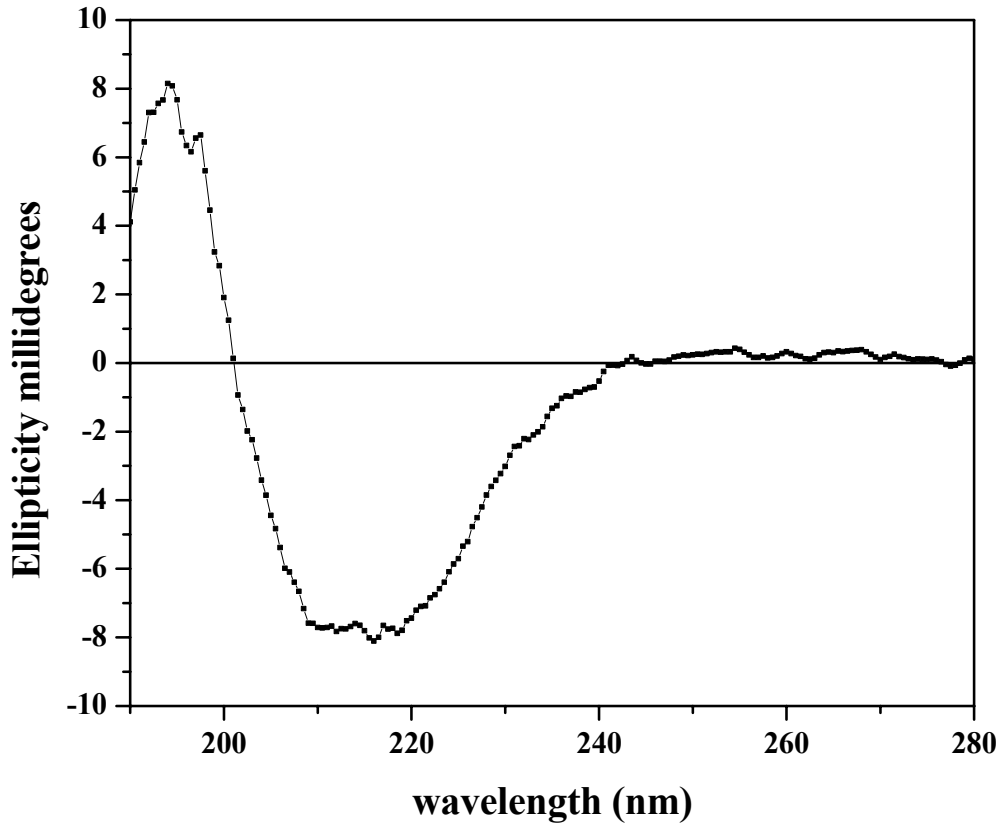


Fig.6 CD spectra of purified 1,4-β-D-Glucan Glucanohydrolase. Far-UV CD spectra were recorded for native enzyme from 280-190 nm at 25°C. Each spectrum represents the average of six scans.

The analysis of the secondary structure of the enzyme obtained using CDPRO shows the presence of:

α helix regular	0.588
α helix distorted	0.226
β sheet regular	0.001
β sheet distorted	0.002
β turn	0.062
unordered structure	1.04

Determination of isoelectric point

The isoelectric point of the enzyme was determined to be 4.33 by isoelectric focusing, indicating that it is an acidic protein.

Activity towards different substrates

The purified enzyme hydrolyzes xylan but does not show any activity towards chitin, guar gum, filter paper, gum arabic and avicel.

Influence of pH on activity and stability

The enzyme was active in a wide range of pH from 5 to 9 with optimum activity at pH 7 for both cellulase and xylanase activities. It was stable in an expansive range of pH 5 to 8 with more than 85% activity (Fig.7).

Influence of temperature on activity and stability

The temperature optimum of the purified enzyme was 50°C (Fig.8). The enzyme was highly thermostable, retaining complete activity of hydrolyzing xylan and CMC at 50°C for more than 15hr and half-life of 20 min at 80°C (Fig.9).

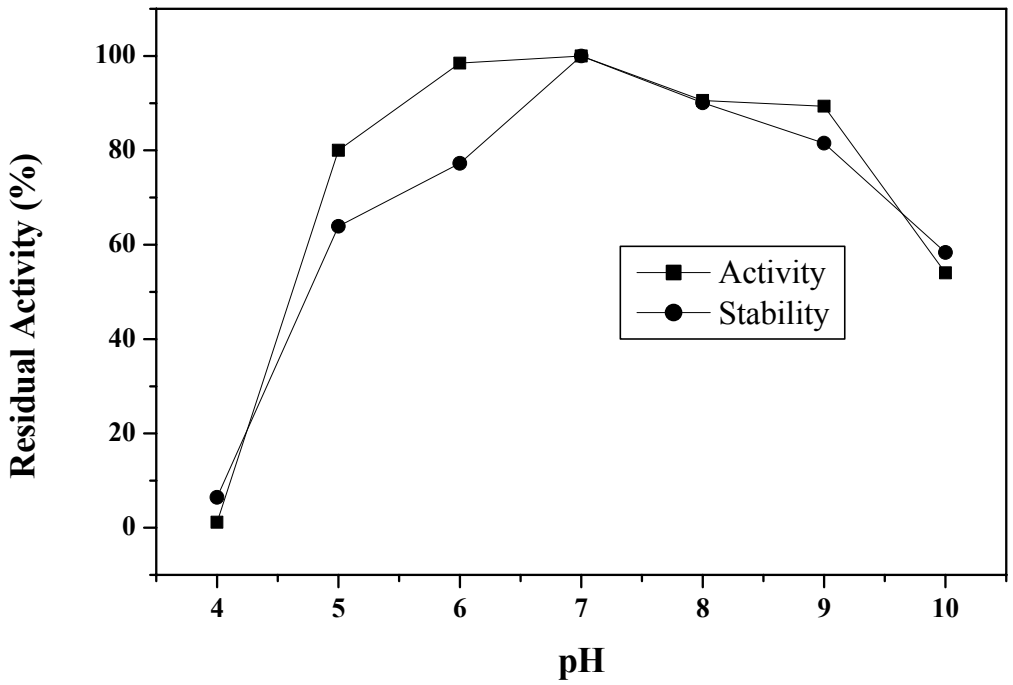


Fig.7 The pH Optimum (■) and pH stability (●) of TSC. The pH stability of the enzyme was measured by incubating 5IU of TSC for an hour at 50°C in buffer of desired pH. The following buffer systems were used: citrate phosphate buffer (pH 4), acetate buffer (pH 5), sodium phosphate buffer (pH 6-8), and carbonate buffer (pH 9-10).

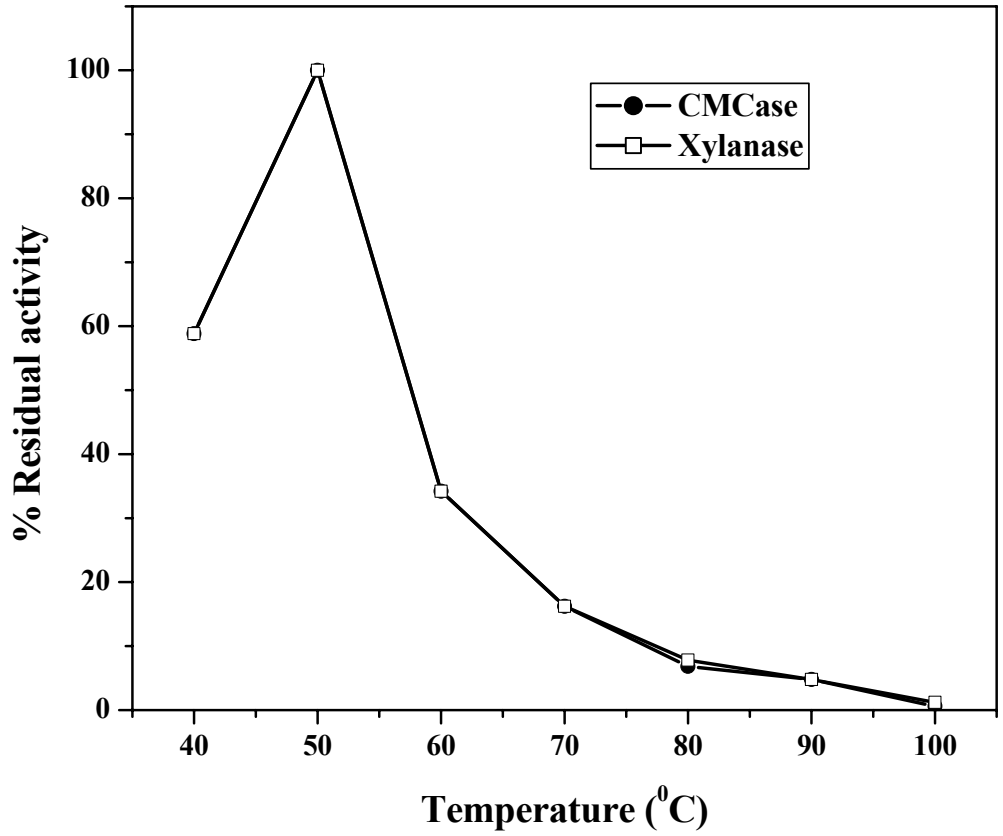


Fig.8 The optimum temperature for CMCase (•) and xylanase (□) activity of TSC in 0.05 M phosphate buffer (pH 7). The activity of the enzyme was determined in the range of (40-100°C)

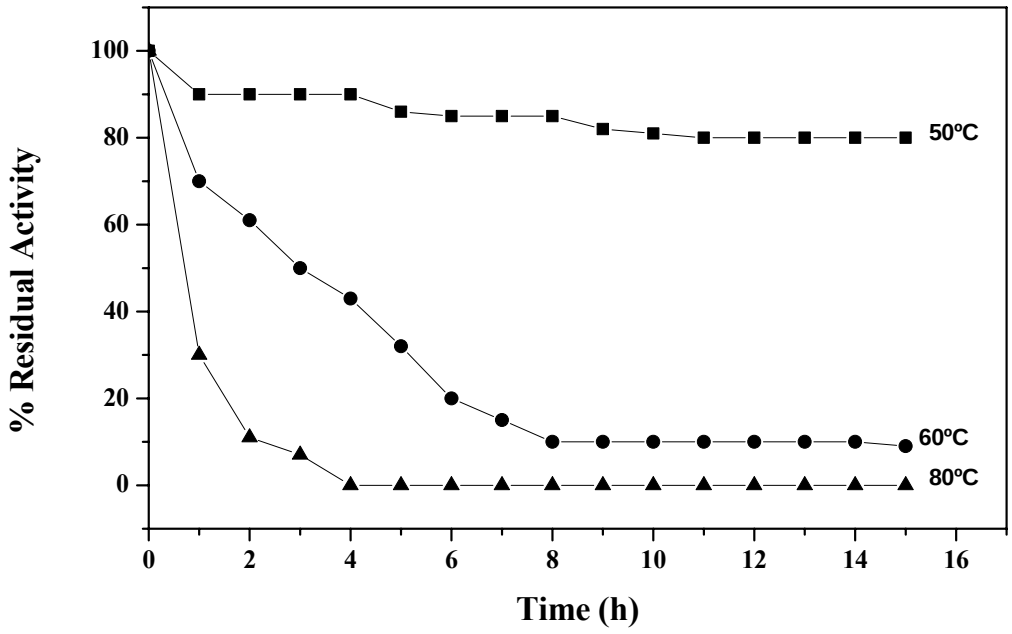


Fig.9 Thermal stability of TSC: 5IU of TSC was incubated in 0.05M phosphate buffer (pH 7) at a) 50°C (■), b) 60°C (●) and c) 80°C (▲) for different time intervals and the residual activity was determined.

End product analysis of hydrolysis of CMC and xylan

The products formed by hydrolysis of CMC and xylans by purified enzyme were analyzed separately by HPLC. The HPLC profiles show the formation of cellooliogsaccharides and glucose from hydrolysis of CMC and xylo-oligosaccharides and xylose from xylan (Fig. 10). The enzyme produced comparatively larger amounts of cellooliogsaccharides and glucose from CMC in 7h while xylo-oligosaccharides and xylose from xylan suggesting the endo mode of action (Fig.11).

A **B**

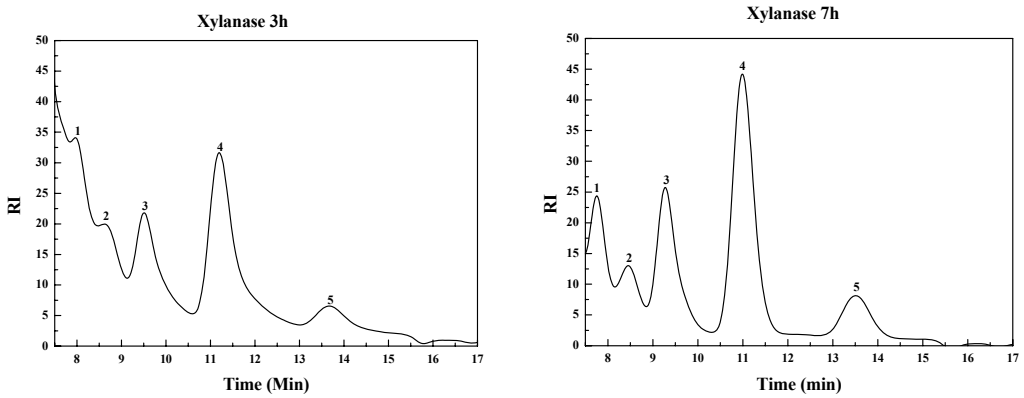


Fig. 10 HPLC analysis of the end products of hydrolysis of xylan by TSC using sugar - PAK column. The column was eluted with $1x10^{-4}$M Ca- EDTA buffer at a flow rate of 0.4 ml/min at 70°C.

1: Xylopentaose

2: Xylotetraose

3: Xylotriose

4: Xylobiose

5: Xylose

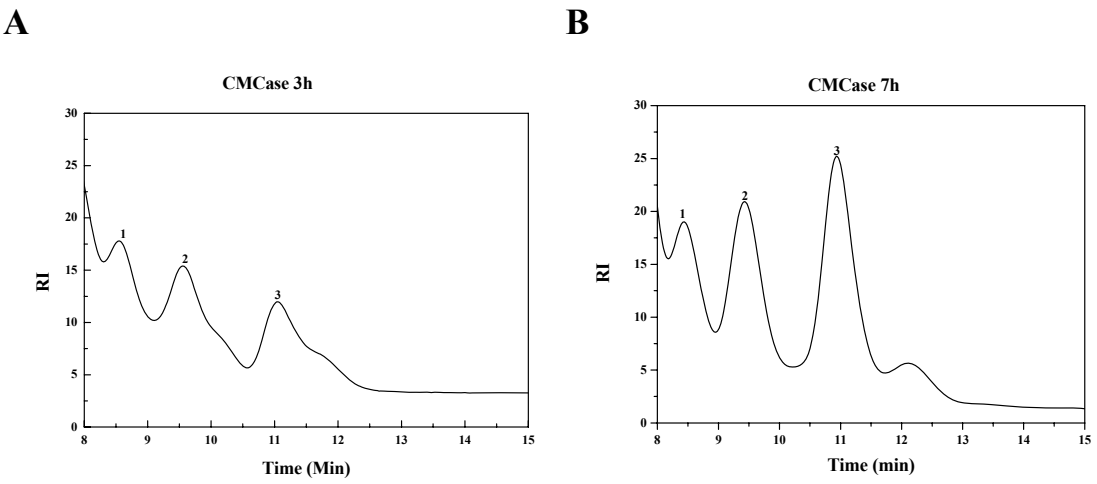


Fig. 11 HPLC analysis of the end products of hydrolysis of CMC by TSC using sugar - PAK column. The column was eluted with $1x10^{-4}$ M Ca- EDTA buffer at a flow rate of 0.4 ml/min at 70°C.

1: Cellotriose

2: Cellobiose

3: Glucose

Kinetic studies of carboxymethyl cellulase and xylanase activities:

During purification, the ratio of carboxymethyl cellulase to xylanase activity remained constant indicating that both the activities reside on the same protein. Therefore, the question arises of whether the two activities of the enzyme are localized on the same or on two different active sites. The following data indicated that the enzyme at the same active site hydrolyzes CMC and xylan; (i) the enzyme showed same pH optima for CMC and xylan hydrolysis. (ii) The enzyme was stable at 50°C for 12h.for both substrates. (iii) both the activities are stable in wide pH range of 5 to 8. (iv) both activities were inhibited by O-Pthaldialdehyde. (v) The mode of action of the enzyme for hydrolysis of both the substrates is of endo type. However the Km for the two activities are widely different, Viz. 6 mg/ml for CMC and 3 mg/ml for Xylan. The kinetic analysis by the method of Keleiti (Keleti *et al*., 1987) was employed to analyze whether one or two sites are responsible for hydrolysis of CMC and xylan (Fig.12)

If CMC and xylan were hydrolyzed at two different active sites, the rate of hydrolysis of two substrates should be equal to the sum of their individual rates. Conversely if there is single site, these two substrates will show competitive inhibition with respect to each other and hence the rate of hydrolysis by the enzyme with the mixture of substrates would be lower than the sum of the rates with individual substrates. The Table 4 summarizes the calculated and the actual experimental values.

The observed values of overall rate of hydrolysis agreed better with the theoretical values calculated from Eqn.1 than with those calculated from Eqn.2 and Eqn.3 (Fig.12). The results of the kinetic analysis are consistent with assumption of a single active site.

Table-4

Overall reaction velocities of TSC with mixtures of various concentrations of xylan and CMC

Individual Rates				**Rate of Mixtures**	
Xylan		**CMC**		**Observed**	**Calculated**
Conc. mg/ml	**A_{540}**	**Conc. mg/ml**	**A_{540}**	**A_{540}**	**A_{540}**
1	0.0984	10	0.1376	0.0921	0.2360
2	0.1368	8	0.1223	0.1055	0.2591
3	0.1626	6	0.1128	0.1128	0.2754
4	0.1874	4	0.1076	0.1188	0.3062
5	0.2061	2	.0921	0.1233	0.3294

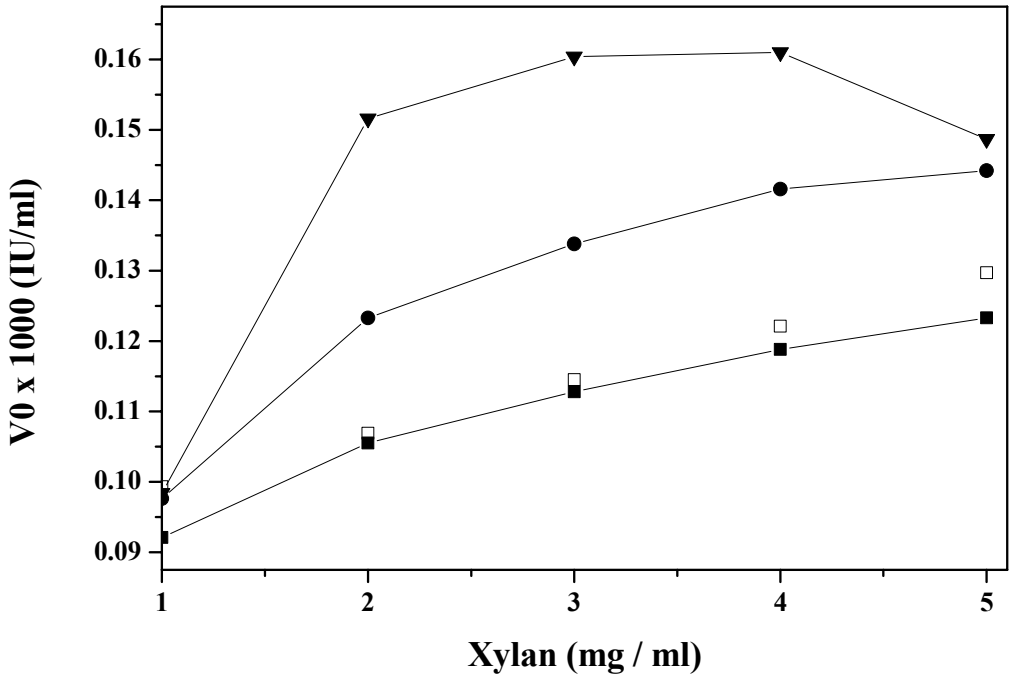


Fig. 12 Plot of V_0 V_S [S] at different concentrations of CMC and xylan. In each case, 0.2 units / μg of the enzyme were added to reaction mixture. The concentration of only one substrate is indicated on the abscissa. [■] Experimental points, [□] single active site Eqn.I, [●] two substrates mutually influencing each other's binding Eqn.III, [▼] two independent active sites Eqn.II.

Chemo- affinity labeling studies:

Conformation and microenvironment at the active site of 1,4-β-D-Glucan Glucanohydrolase was probed with fluorescent chemo affinity labeling using o-phthalaldehyde as the chemical initiator. O-phthalaldehyde is bifunctional reagent forms isoindole derivative by cross-linking the proximal thiol and amino groups (Simons and Johnson., 1978) O-phthalaldehyde reacted with TSC yielding a stable fluorescent derivative, which exhibited an absorbance characteristic of isoindole (420nm) with the complete loss of activity of the enzyme to hydrolyze both CMC and xylan. Complete inactivation of the TSC by O-phthalaldehyde might be due to the formation of isoindole derivative by cross-linking the proximal thiol and amino groups located at or near the active site of the enzyme.

Cellulases from fungi and anaerobic bacteria have been discovered which show the interrelationship with xylanase in its induction, regulation at transcriptional level and inactivation by inhibitors (Yaoi and Mitsuishi, 2002; Aro *et al*., 2001; Morag *et al*., 1990; Royer and Nakas, 1990; Sanchoa *et al*., 2003; Aro *et al*., 2003). It has been reported that few endoglucanases [1,4-β-D-Glucan Glucanohydrolase] components of *Trichoderma viride* (Shikata and Nisizawa,1975) and *Irpex lacteus* show xylanase activity (Kanda *et al*.,1976).However, number of bifunctional enzymes which have both cellulase and xylanase activity known to date are rather low, and information on their properties is scarce. The results of the present study suggest that an interrelationship exists between the xylanase and cellulase activity with probable ecological significance. The identical pH and temperature optima, similar stability, inhibition by OPTA reagent and kinetic analysis are in favor of the single active site for the hydrolysis of carboxymethyl cellulose and xylan by 1,4-β-D-Glucan Glucanohydrolase

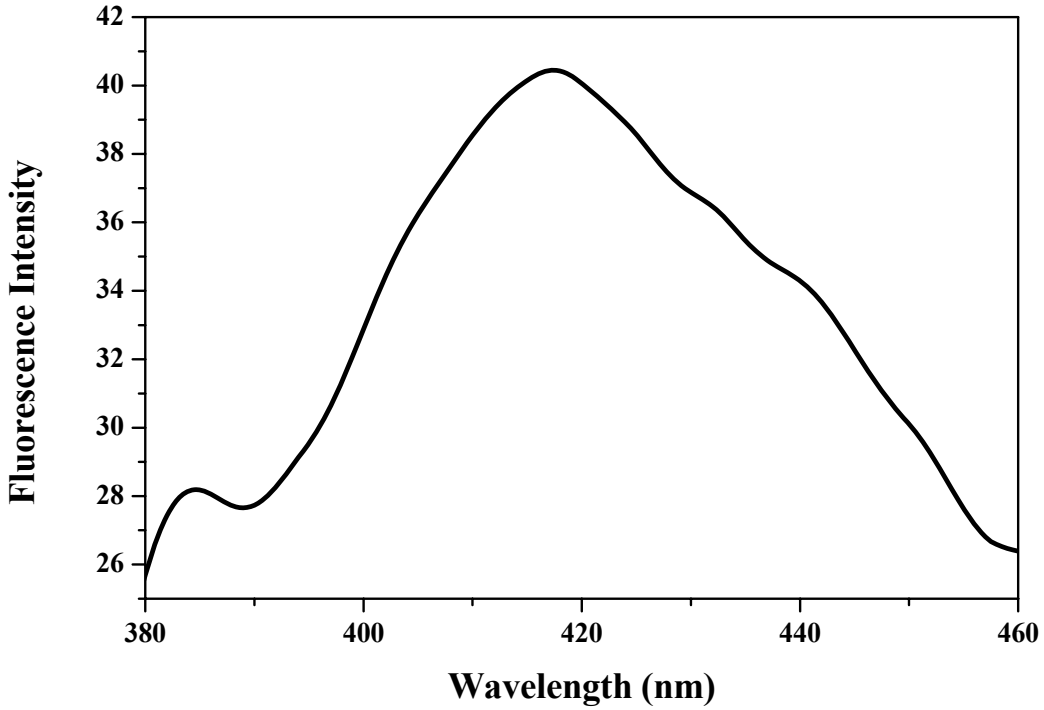


Fig. 13 Isoindole fluorescence of endoglucanase modified by OPTA. TSC sample (50μg) was incubated with 1mm of OPTA for 10 min at 25°C. The isoindole fluorescence was monitored at λ_{excit} 338nm and λ_{emiss} 415nm.

SUMMARY

A low molecular weight 1,4-β-D-Glucan Glucohydrolase from an extracellular culture filtrate of *Thermomonospora* sp. was purified to homogeneity. The molecular weight of the purified enzyme was 14.2 kDa by MALDI -TOF analysis and is in agreement with SDS PAGE and gel filtration chromatography. The purified enzyme had both endocarboxymethyl cellulase and endoxylanase activities. A kinetic method was employed to study the active site of the enzyme that hydrolyses both carboxymethyl cellulose and xylan. The experimental data coincides well with the theoretical values calculated for the case of a single active site. Conformation and microenvironment at the active site was probed with fluorescent chemo affinity labeling using o-phthalaldehyde as the chemical initiator. Formation of isoindole derivative resulted in complete inactivation of the enzyme to hydrolyze both xylan and CMC as judged by fluorescence studies corroborating a single active site for the hydrolysis of xylan and CMC.

REFERENCES

Aro, N., A. Saloheimo, M. Ilmen, M.Penttila. 2001. ACEII, a novel transcriptional activator involved in regulation of cellulase and xylanase genes of *Trichoderma reesei* J. Biol. Chem., 276:24309-24314.

Aro, N., M. Ilmén, A. Saloheimo, M. Penttilä. 2003. ACEI of *Trichoderma reesei* is a repressor of cellulase and xylanase expression, Appl. Environ. Microbiol, 69:56-65.

Bhat, M. 2000. Cellulases and related enzymes in biotechnology, Biotechnol. Adv. 18:355-383.

Boyer, P.D. 1954.Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials. J. Am. Chem. Soc. 76:4331- 4337.

Bradford, M.M.1976 A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

Calza, R.E., D.C.Irwin and D.B.Wilson. 1985. Purification and characterization of two β, 1-4-endoglucanase from *Thermomonospora fusca,* Biochemistry, 24:7797-7804.

George, S., A. Ahmad and M. Rao. 2001. Studies on carboxymethyl cellulase produced by an alkalothermophilic actinomycete. Biores. Technol**.** 77:171-175.

Irwin, D.C., M. Spezio, L. P. Walker and D. B. Wilson.1993.Activity studies of eight purified cellulases: specificity, synergism and binding domain effects. Biotech. Bioengg. 42: 1002-1013.

Jagannathan,V., K. Singh and M. Damodaran.1956.Carbohydrate metabolism in citric acid fermentation. 4. Purification and properties of aldolase from *Aspergillus niger***.** Biochem. J. 63:94-105.

Kanda, T., K. Wakabayashi and K. Nisizawa. 1976. Xylanase activity of an endo-cellulase type of carboxymethyl cellulase type from *Irpex lacteus*. J.Biochem. 79:989-995.

Keleti, T. R., Leocini, R. Pagani, E. Marinello.1987. A kinetic method for distinguishing whether an enzyme has one or two active sites for two different substrates. Eur. J.Biochem., 170:179-183.

Kirby, J.,J. C. Martin, A. S. Daniel and H. J. Flint.1997.Dockerin like sequences in cellulases and xylanases from the rumen cellulolytic bacterium *Ruminococcus flavefaciens*. FEMS Microbiol. Lett. 149:213-219.

Kulkarni,N., A. Shendye, and M. Rao. 1999. Molecular and biotechnological aspects of xylanases. FEMS Microbio. Rev. 23:411-456.

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

Lynd, L.R., P. J. Weimer, W.H. Zyl, I. S. Pretorius. 2002. Microbial Cellulose Utilization: Fundamentals and Biotechnology, Micro. Mol. Biol.Rev. 66:506-577.

McCarthy, A. J. and T. Cross. 1984. A Taxonomic Study of *Thermomonospora* and other Monosporic Actinomycetes, J. Gen. Microbiol. 130:5-25.

McSweeny, C. S., A. Dulieu and R. Bunch. 1998. *Butyrivibrio* spp. and other xylanolytic microorganisms from the rumen have the cinnamoyl esterase activity. Anaerobe 4: 57-65.

Miller, G. M. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31:426-428.

Mishra,C., S. Keskar, and M. Rao. 1984 . Production and properties of extracellular endoxylanase from *Neurospora crassa* Appl. Envir. Microbiol. 48: 224-228.

Morag, E., E. A. Bayer, R. Lamed. 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose degrading enzymes. J.Bac.172:6098 - 6105.

Poulsen, O.M., and L.W.Peterson. 1988. Growth of *Cellulomonas* sp. ATCC 21399 on different polysaccharides as sole carbon source induction of extracellular enzymes. Appl. Mirobiol. Biotechnol. 29:480-484.

Robson, L. M. and G.H.Chambliss.1989. Review: Cellulases of bacterial origin, Enzyme Microb. Technol. 11:626-644.

Royer , J.C. and J. P. Nakas. 1990. Interrelationship of xylanase induction and cellulase induction of *Trichoderma longibrachiatum.* Appl. Enviorn. Microbiol. 56:2535-2539.

Sanchoa, A. I., C. B. Fauldsa, B. Svenssonb, B. Bartoloméc, G. Williamsona and N. Jugea. 2003. Cross-inhibitory activity of cereal protein inhibitors against a-amylases and xylanases Biochim. Biophys. Acta. 1650:136-144.

Shikata, S. and K.Nisizawa. 1975. Purification and properties of an exocellulase component of a novel type from *Trichoderma viride*. J. Biochem.78:499-512.

Simons, S. S., and D.F. Johnson. 1978. Reaction of o-Phthalaldehyde and thiols with primary amines: Fluorescence properties of 1- alkyl (and aryl) thio-2 -alkylisoindoles. Anal. Biochem. 90:705-725.

Spande, T. F. and B. Witkop. 1967. Determination of the tryptophan content of proteins with *N*-bromosuccinimide. Methods Enzymol. 11: 498-506.

Stutzenberger, F. J. 1971. Cellulase production by *Thermomonospora curvata* isolated from municipal solid waste, Appl. Microbiol. 22:147-152.

Wang, J., M. Ding, Y.H. Li, Q.X. Chen, G. J. Xu and F. K. Zhao. 2003. A monovalent anion affected multi-functional cellulase EGX from the mollusca, *Ampullaria crossean*, Protein Expression and Purification. 31:108-114.

Wilson, D.B., Cellulases of *Thermomonospora fusca*. In: Biomass, Part A: Cellulose and Hemicellulose. In: W. A. Wood, S.T.Kellogg, (Eds.), Academic Press, San Diego, CA, 1988, Methods Enzymol.160:314-323.

Yaoi, K. and Y. Mitsuishi .2002. Purification, characterization, cloning, and expression of a novel xyloglucan-specific glycosidase, oligoxyloglucan reducing end-specific cellobiohydrolase. J.Biol.Chem. 277:48276-48281.

Zhang, S., G. Lao and D. B. Wilson.1995. Characterization of *Thermomonospora fusca* exocellulase. Biochemistry 34:3386-3395.

CHAPTER-III

Conformation and Microenvironment of the Active Site of 1, 4-β-D-Glucan Glucanohydrolase

ABSTRACT

Conformation and microenvironment at the active site of 1,4-β-D-Glucan Glucanohydrolase was probed with fluorescent chemo affinity labeling using o-phthalaldehyde. OPTA has been known to form a fluorescent isoindole derivative by cross-linking the proximal thiol and amino groups of cysteine and lysine. Modification of lysine of the enzyme by TNBS and of cysteine residue by PHMB abolished the ability of the enzyme to form an isoindole derivative with OPTA. Kinetic analysis of the TNBS and PHMB – modified enzyme suggested the presence of essential lysine and cysteine residue respectively at the active site of the enzyme. The substrate protection of the enzyme with carboxymethylcellulose (CMC) confirmed the involvement of lysine and cysteine residues in the active site of the enzyme. Multiple sequence alignment of peptides obtained by tryptic digestion of the enzyme showed cysteine is one of the conserved amino acids corroborating the chemical modification studies.

INTRODUCTION

In an attempt to understand the catalytic mechanism of an enzyme, it is essential to study the structural elements and the three dimensional conformation of the active site. Chemical modification of the reactive amino acid side chains in the active site helps to identify residues that are important for catalysis. The utility of chemical modification is greatly enhanced by its use in conjunction with site directed mutagenesis, which mutually supplement each other.

Important amino acid residues for the catalytic activity of cellulases

Family 5 cellulases

Three-dimensional structures of family 5 enzymes from *Clostridium thermocellum*, *Clostridium cellulolyticum*, *Bacillus agaradhaerens*, and a *Thermobifida fusca* β-mannanase allowed the identification and dissection of the catalytic residues in family 5. The information about some representative cellulases from family 5 is given in Table.1

Table-1

Family 5 cellulases

Enzyme	Organism	Catalytic acid/base	Catalytic nucleophile	Reference
Cel5B (CelC)	*C. thermocellum*	Glu140	Glu280	Wang *et al*.,1993
CelA	*C. cellulolyticum*	Glu170	Glu307	Dominguez *et al*., 1995
Cel5A	*B. garadhaerens*	Glu139	Glu228	Davies *et al*.,1998
Cel5A (CelZ)	*E. chrysanthemi*	Glu133	Glu220	Hilge *et al*.,1998
E1	*A. cellulolyticus*	Glu162	Glu282	Sakon *et al*.,1996

Family 6 cellulases

Family 6 enzymes perform catalysis with inversion of anomeric configuration. The first significant site-directed analysis of a family 6 enzyme, which remains a model study of its type, is the analysis of potential catalytic residues in the *Cellulomonas fimi*

endoglucanase Cel6A (formerly CenA). Table 2 summarizes the potential catalytic residues identified in family 6.

Table-2

Family 6 cellulases

Enzyme	Organism	Catalytic acid	Putative, but disputed, catalytic base	pKa modulator	Reference
Cel6A (CBHII)	*T. reesei*	Asp221	Asp401	Asp263	Rouvinen *et al*., 1990
Cel6A (CBHII)	*H. insolens*	Asp226	Asp405	Asp268	Varrot *et al*., 1999
Cel6A (E2)	*T. fusca*	Asp117	Asp265	Asp156	Wolfgang *et al*., 1999
Cel6B (endo B)	*H. insolens*	Asp139	Asp316	Asp92	Davies *et al*., 2000
Cel6A (CenA)	*C. fimi*	Asp252	Asp392	Asp287	[Howard *et al*., 1995

Family 7 cellulases

Cellulases from family 7 perform catalysis with net retention of configuration. The roles of these residues, Glu212-Asp214-Glu217 in *T. reesei* Cel7A (formerly CBHI), were investigated by mutation of these residues to their isosteric amide counterparts. Table 5 summarizes the catalytic active residues from the solved family 7 structures. Further high resolution crystal structures of these site-directed variants by the Uppsala group revealed how a cellulose chain was bound along the 50Å substrate binding tunnel. Active site variants Glu212Gln and Glu217Gln were crystallized together with cellooligosaccharides and the structures reveal glucose moieties covering subsites 37 to +4 (Mackenzie *et al*., 1998). The binding mode observed corresponds to that expected during productive binding of a cellulose chain and supports the hypothesis that hydrolysis by *T.reesei* Cel7A proceeds from the reducing end of a cellulose chain (Divne *et al*., 1998). Family 7 also contains endoglucanases in addition to the cellobiohydrolases. One of the most exciting developments in recent times is the construction of `glycosynthases': mutant glycoside hydrolases for oligosaccharide synthesis (Mackenzie *et al*., 1998). A powerful synthetic tool has been developed

through mutagenesis of the *H. insolens* family 7 endoglucanase. The information about some representative cellulases from family 7 is given in Table.3

Table-3

Family 7 cellulases

Enzyme	Organism	Catalytic acid/base	Catalytic nucleophile	Reference
Cel7A (CBH I)	*T. reesei*	Glu217	Glu212	Ståhlberg *et al*., 1996
Cel7B (EG I)	*T. reesei*	Glu201	Glu196	Kleywegt *et al*., 1997
Cel7B (EG I)	*H. insolens*	Glu202	Glu197	Davies *et al*., 1997
Cel7B (EG I)	*F. oxysporium*	Glu202	Glu197	Davies *et al*., 1997

Family 8 cellulases

Family 8 enzymes perform catalysis with inversion of anomeric configuration. The catalytic residues in family 8 were first identified by Ozaki et al. using a Bacillus sp. KSM- 330 cellulase (Ozaki *et al*., 1994). The structure of *C. thermocellum* Cel8A (formerly CelA) was later solved (Alzari *et al*., 1996) and the homologous Glu95 could easily be identified. However, the authors questioned the assignment of Asp152 as `catalytic' since they find that Asp278 could also be a possibility. Table 4 summarizes the data.

Table-4

Family 8 cellulases

Enzyme	Organism	Catalytic acid	Catalytic base	Reference
Cel8A (CelA)	*C. thermocellum*	Glu95	Asp152	Alzari *et al*.,1996
Endo K	*Bacillus* sp.	Glu130	Asp191	Ozaki *et al*., 1994

Family 9 cellulases

Family 9 cellulases perform catalysis with inversion of anomeric configuration. This is a most important family because all currently known plant cellulases belong to family 9 (Henrissat et *al*., 2001).

Table-5

Family 9 cellulase

Enzyme	Organism	Catalytic acid	Catalytic base	Reference
Cel9A (CelD)	*C. thermocellum*	Glu555	Asp201	Juy et *al*., 1992
Cel9A (E4)	*T. fusca*	Glu424	Asp58	Sakon et *al*., 1997
CelZ	*C. stercorarium*	Glu447	Asp84	Sakon et *al*., 1997

The putative catalytic residues of family 9 cellulase from *Clostridium stercorarium* CelZ (Riedel et *al*., 1999), Asp84 and Glu447, which are located within the N-terminal domain of the modular protein, were replaced by site directed mutagenesis. A minimized CelZ derivative (CelZCP) comprising the catalytic domain and the adjacent cellulose binding domain (CBD) family IIIc domain CP was used as target for mutagenesis. Six mutant enzymes and the unmodified CelZCP protein were compared with respect to thermal stability and activity, substrate specificity, product profile and synergism. Replacement of either one or both catalytic residues completely eliminated the ability of CelZ to attack insoluble Avicel preparations, whereas the endo-activity measured via hydrolysis of CMC was retained upon substitution of the catalytic base Asp84. This may reflect catalytic rescue by the carboxylates present on CMC. The `endo-active' *C. stercorarium* CelZ mutants who are defective in activity on insoluble cellulose were used in co-operativity studies with the *C. stercorarium* `exo'-glucanase CelY. As expected, synergy was found to be dependent on the endo-activity of CelZ. Mutants Asp84Gly and Asp84Glu were able to enhance the degradation of crystalline cellulose significantly, although no products could be released from this substrate by individual action of the mutants. Table 5 summarizes the catalytically active residues of this family.

Family 45 cellulases

Family 45 enzymes operate with inversion of anomeric configuration (Schou et *al*., 1993). The oligosaccharide complexes reveal solvent water suitably placed for the participation in a single displacement, inverting the reaction mechanism. A large conformation change takes place upon substrate binding. This `lid flipping' has the effect of increasing the hydrophobic environment of the catalytic proton donor, enclosing the active site at the point of cleavage, and bringing a third aspartate (Asp114) in close proximity of the substrate. Site-directed mutagenesis of the catalytic residues has been used to confirm their significance in catalysis (Table 6).Members of this family of endoglucanases have higher activity on amorphous cellulose than on CMC, but they do not degrade the crystalline regions of cellulose (Schülein et *al*., 1997).

Table-6

Family 45 cellulase

Enzyme	Organism	Catalytic acid	Catalytic base	Helper
Cel45A (endo V)	*H. insolens*	Asp10	Asp121	Asp114
Site-directed variants		Asn10	Asn121	Asn114

Structure/function relationships are one of the central issues in the investigation of biological macromolecules. The conformational integrity of an enzyme is essential for its activity. Investigations involving chemical modification of an enzyme can potentially yield insights into structure function relationships. Structure-function relationships are one of the central issues in the investigation of biological macromolecules. The conformational integrity of an enzyme is essential for its activity. In the present chapter, the residues involved in catalysis and substrate binding have been determined. We have also attempted to correlate the loss in catalytic activity to the conformational changes occurring in TSC. The conformational changes in the active site were probed by utilizing the fluorescent properties of *o*- phthalaldehyde (OPTA). OPTA is a fluorescent chemoaffinity label (FCAL), which forms a fluorescent isoindole derivative by cross-linking the thiol and amine groups of cysteine and lysine, respectively. Data presented here suggests that carboxyl acid group residues, lysine and cysteine are the essential residues in the active site of TSC.

MATERIALS AND METHODS

Materials

PHMB, TNBS, WRK and OPTA were obtained from Sigma Chemical Co. USA. All buffer salts were from standard commercial sources and of highest quality available. All other chemicals used in the present work were of analytical grade.

Methods

Production and purification of TSC

TSC was produced and purified as described in Chapter II. The carboxymethyl cellulase activity and protein was determined according to Miller and Bradford (Miller, 1959; Bradford 1976) respectively.

Modification of TSC with o- phthalaldehyde

Enzyme sample (5µg) was incubated with 1mM of OPTA for 10 min at 25°C. The formation of enzyme-isoindole derivative was followed spectrophotometrically by monitoring the increase in fluorescence at 420 nm with excitation wavelength fixed at 338nm. The aliquots were removed at intervals to check the residual carboxymethyl cellulase activity of the enzyme (Simons and Johnson, 1978)

Modification of TSC with TNBS

TSC (10 µg) was incubated with varying concentrations of 2,4,6- trinitrobenzenesulfonic acid (4 to 10 mM) in water, in the presence of 1 ml 4% sodium bicarbonate in the reaction volume of 3 ml at 37°C in the dark. Aliquots were withdrawn at suitable time intervals to check the residual activity of enzyme. 1ml of 10 % SDS solution was added to solubilize the protein upon addition of 0.5 ml of 1 M HCl. The absorbance of solution at 340 nm was read against blank containing water instead of protein solution. An extinction coefficient of 1.4×10^4 $M^{-1}cm^{-1}$ was used to calculate the number of amino groups (Habeeb, 1996). The degree of inactivation in presence of different amount of CMC was also determined. Enzyme samples (10 µg each) were incubated with different concentrations of CMC for 10 min before addition of inhibitors. Control tubes with only enzyme; only inhibitor and inhibitor/ substrate were incubated under identical conditions.

Inactivation kinetics of TSC by PHMB

TSC (20 µg) was incubated with varying concentrations of p-hydroxy-mercury benzoic acid (PHMB) in 50 mM phosphate buffer pH 7 at 25°C and assayed for carboxymethyl cellulase activity at different time intervals. Control tubes having only enzyme or only inhibitor or inhibitor and substrate were incubated under identical conditions. Substrate protection studies were performed by incubating the enzyme with the substrate CMC for 10 min prior to the addition of the modifier.

PHMB titration of TSC

TSC (7×10^{-5} M) in 50Mm phosphate buffer, pH 7.0 was titrated with 10 µl of PHMB (1mM) and the progress of the reaction was monitored spectrophotometrically at 250 nm. Simultaneously aliquots of the reaction mixture were withdrawn to assay the residual activity. The number of Cys residues modified was determined by the method of Boyer (Boyer, 1954).

Circular dichroism measurements

CD spectra were recorded in a Jasco-J715 spectrophotometer at ambient temperature using a cell of 1mm path length. Replicate scans were obtained at 0.1nm resolution, 0.1 nm bandwidth and a scan speed of 50 nm/min. Spectra were average of 6 scans with the baseline subtracted spanning from 260 to 190 nm in 0.1 nm increments. The CD spectrum of TSC
(1 mg/ml) was recorded in 10 mM sodium phosphate buffer pH 7.

Modification of carboxyl group residues by Woodward's Reagent K

TSC (10 µg) in 50 mM potassium phosphate buffer (pH 6.0) was incubated with different concentrations of WRK (5-30 mM) at 25°C. Aliquots of the reaction mixture were withdrawn at indicated time intervals and the excess reagent was quenched with a final concentration of sodium acetate buffer (pH 5.0; final concentration 250 mM). The residual CMCase activity was measured and expressed as percentage of a control.

RESULTS AND DISCUSSION

Formation of isoindole derivative at the active site

Conformation and microenvironment at the active site of TSC was probed with fluorescent chemo affinity labeling using O-phthalaldehyde as the chemical initiator. O-phthalaldehyde is bifunctional reagent forms isoindole derivative due to its reaction with -SH and -NH_2 groups of lysine and cysteine residues (Palczewski et *al*., 1983).

O-phthalaldehyde reacted with TSC yielding a stable fluorescent derivative, which exhibited an absorbance characteristic of isoindole (420 nm) when excited at 338nm with the complete loss of activity of the enzyme to hydrolyze both CMC and xylan (Fig. 1).

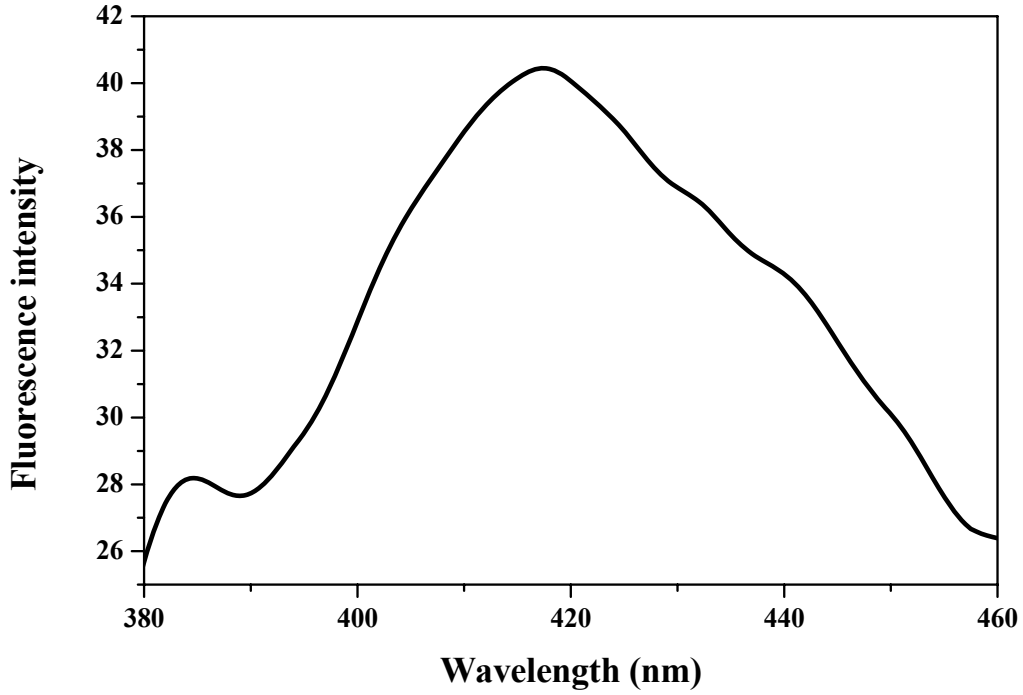


Fig. 1 Isoindole fluorescence of TSC modified by OPTA. Enzyme sample (5µg) was incubated with 1mM of OPTA for 10 min at 25°C. The isoindole fluorescence was monitored at λ_{excit} 338nm and λ_{emiss} 420nm.

Complete inactivation of the by O-phthalaldehyde might be due to the formation of isoindole derivative by cross-linking the proximal thiol and amino groups located at or near the active site of the enzyme (Scheme 1). The substrates (CMC) protect the enzyme

from inactivation by OPTA. Moreover, there was a linear relationship between fluorescence increase and enzyme inactivation.

Inactivation kinetics of TNBS modified enzyme

Incubation of TSC with different concentrations of TNBS resulted in a time and concentration dependant loss of enzyme activity as shown in Fig. 2.

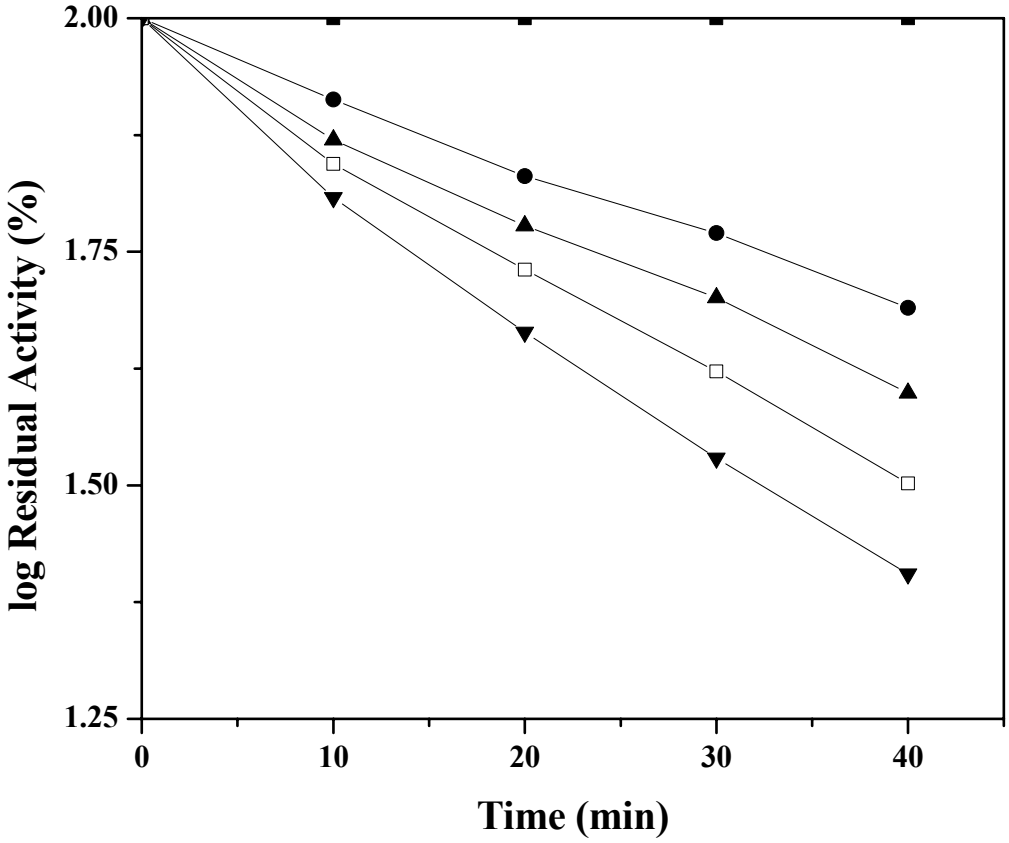


Fig. 2 Kinetics of inactivation of TSC by TNBS. Pseudo first order plots for the inactivation of TSC by TNBS. Enzyme was incubated with 4mM (●), 6mM (▲), 8mM (□) 10mM (▼) and control (■) at 37°C. Aliquots were removed at indicated time intervals.

The reaction followed pseudo first order kinetics. The pseudo first order rate constants (K) were linearly correlated to the concentrations of the reagent, suggesting that no reversible complex was formed during the inactivation process. Furthermore a reaction

order of 0.98 with respect to the modifier was determined from the slope of the double logarithmic plots (Fig.3) indicating that 1 mol of TNBS inactivated 1 mol of enzyme. The complete protection of enzyme against inactivation by OPTA by the substrate, confirms the presence of –SH and $–NH_2$ at the substrate binging site. The complete protection of endoglucanase against inactivation by TNBS by the substrate CMC, confirms the presence of lysine at the substrate binging site.

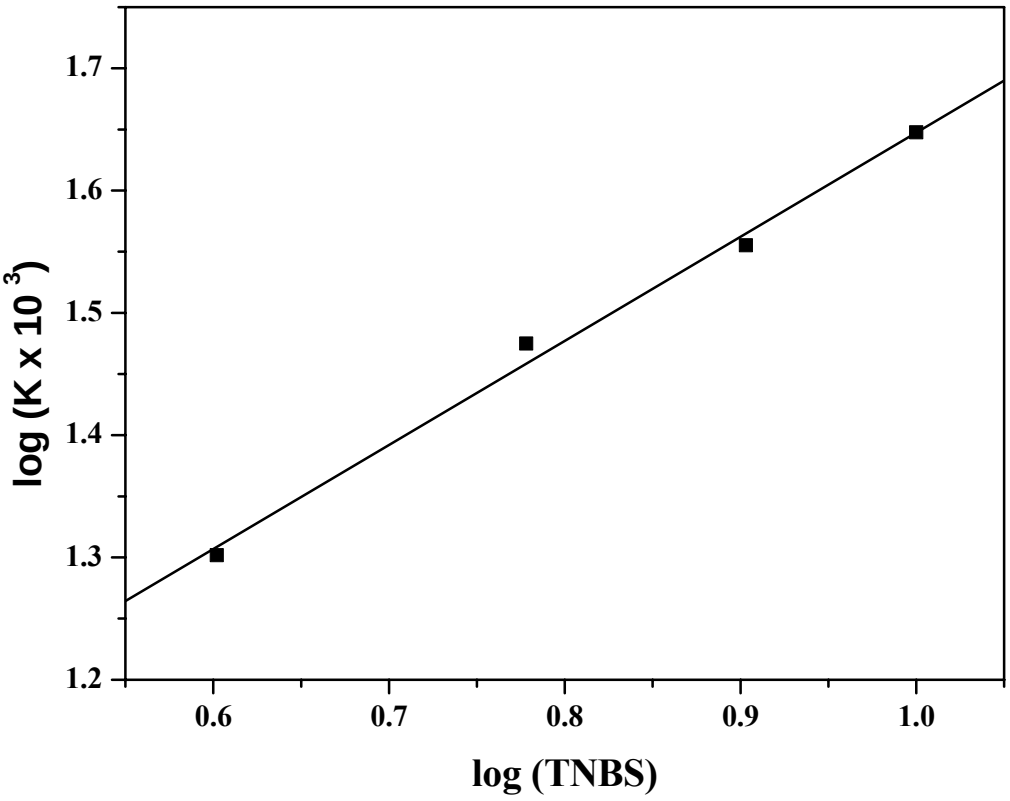


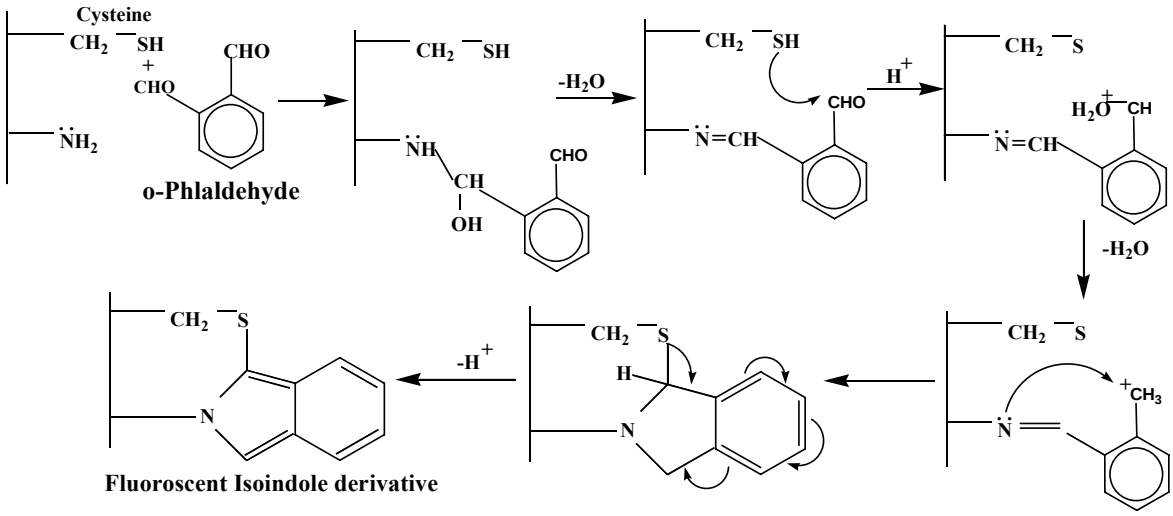
Fig. 3 Double logarithmic plots of pseudo first order rate constants as a function of TNBS concentration

Table-7

Substrate protection of TSC against inhibition by OPTA and TNBS

Reaction	Residual activity (%)
None	100
OPTA (1mM)	0
CMC (1 mg) + OPTA	72.3
CMC (2 mg) + OPTA	86.8
CMC (5 mg) + OPTA	100
TNBS (10 mM)	0
CMC (1 mg) + TNBS	67.6
CMC (2 mg) + TNBS	83.5
CMC (5 mg) + TNBS	98.2

Scheme 1



Modification of cysteine residue by PHMB

The number of cysteine residues in TSC was found to be one from the method of Boyer (Fig.4) The inactivation of the enzyme by PHMB measured in terms of carboxymethyl cellulase activity was found to be dependant on both the time and inhibitor concentration and followed first order kinetics.

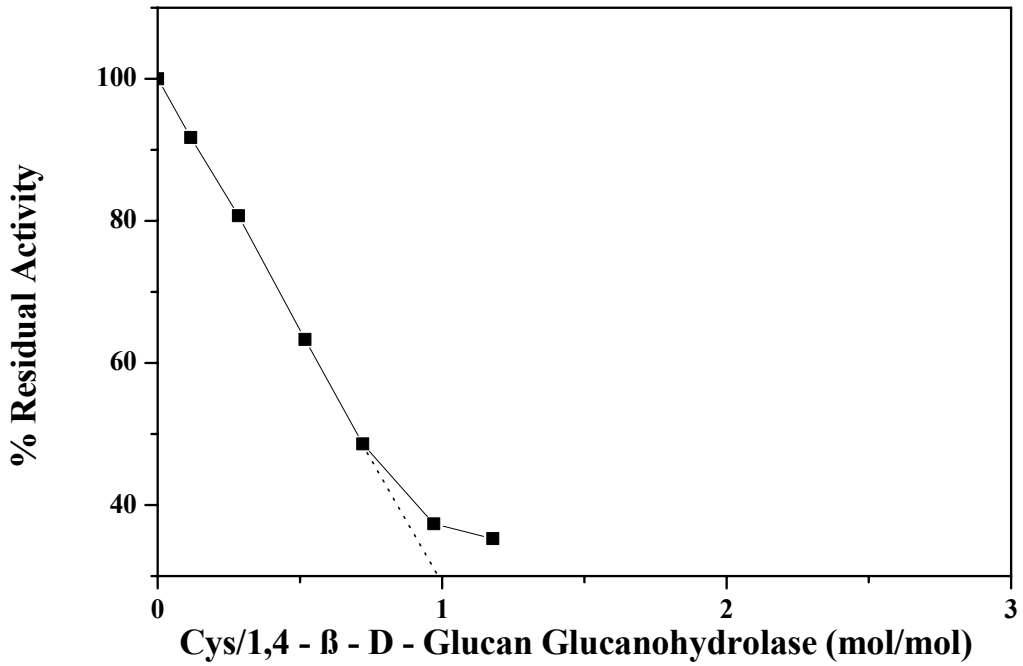


Fig. 4 Plot of percentage residual TSC activity against number of cysteine residues modified.

Monitoring conformational change by circular dichroism measurements

The CD measurements revealed no effect of the modifier on the α- helix and β - sheet content of TSC (Fig. 5). Hence the inactivation of TSC induced by TNBS and PHMB is a result of a direct chemical modification of an essential lysine residue and cannot be attributed to the disruption of the enzyme structure.

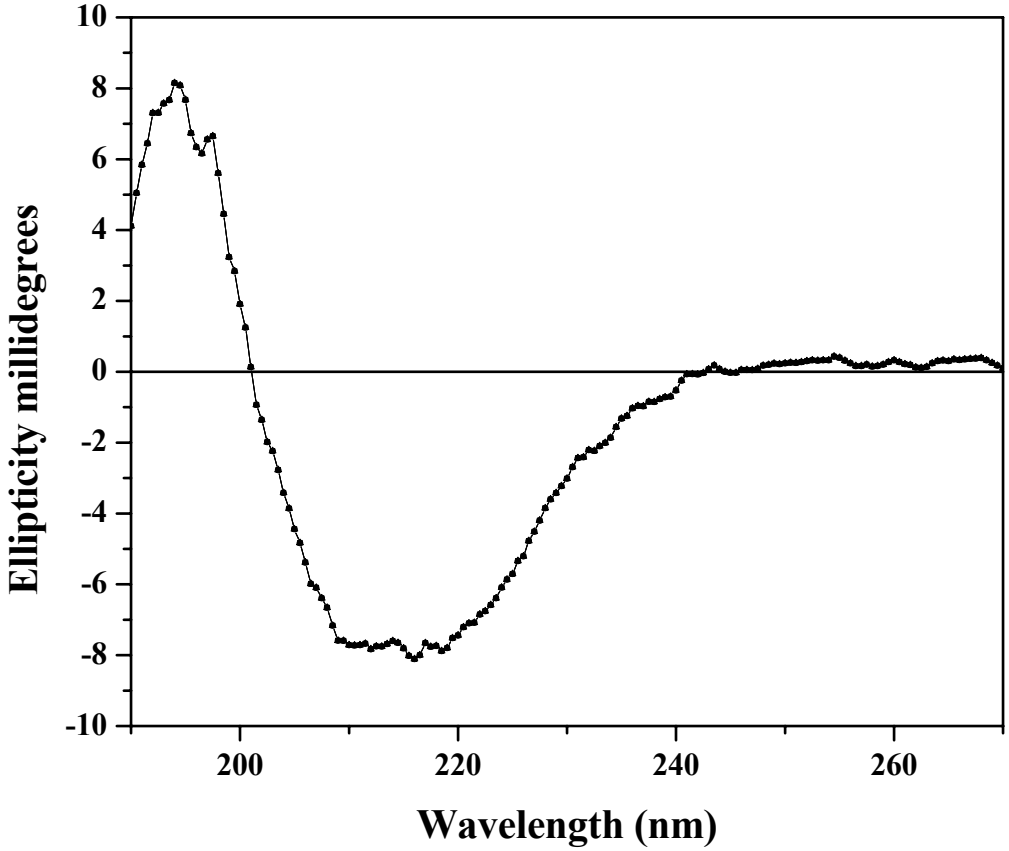


Fig. 5 CD spectra of native (■), TNBS (▲) and PHMB (●) modified TSC. Far-UV CD spectra recorded for native and modified TSC from 280-190 nm at 25°C. Each spectrum represents the average of six scans

Modification of carboxyl group residues by Woodward's Reagent K

The reaction of WRK is initiated by the formation of ketoketenimine which modifies the carboxyl group of the enzyme to give an enol ester with concomitant increase in the absorbance at 340 nm. Fig.6 shows the residual enzymatic activity plotted against the number of carboxyl groups modified. The extrapolation with the Y-intercept passing through 0 % residual activity showed the presence of four carboxyl groups at the catalytic site.

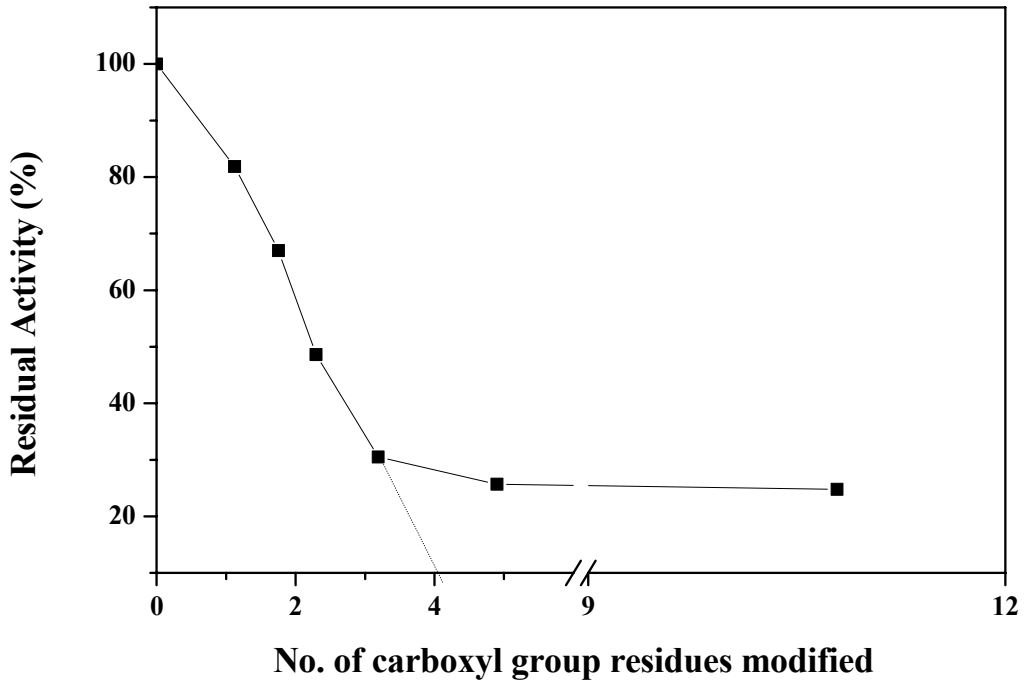


Fig. 6 Plot of percentage residual TSC activity against number of carboxyl group residues modified.

One of the most commonly used procedures for the modification of protein carboxyl groups has been that of coupling with nucleophiles mediated by water-soluble carbodiimide. Experimental evidence from stereochemical studies suggests that the cellulases perform catalysis through a double displacement mechanism similar to hen egg white lysozyme. Such mechanism requires the presence of a nucleophilic residue, which will assist the catalytic acid/base residue during transfer of protons. This acidic residue in order to play its role would have to be in an ionized state under the catalytic conditions.

The reaction of the enzyme with the carboxyl group specific Woodward's Reagent K resulted in progressive loss of activity. When enzyme was incubated with 20 mM WRK for 30 min the extent of inactivation observed was 80%. Inactivation of TSC by WRK was dependent on the time and reagent concentration (Fig. 7). A double logarithmic plot of the observed pseudo-first-order rate constants against reagent concentration yielded an order of 4.0 indicating that modification of four carboxyl residues resulted in the loss of enzyme activity.

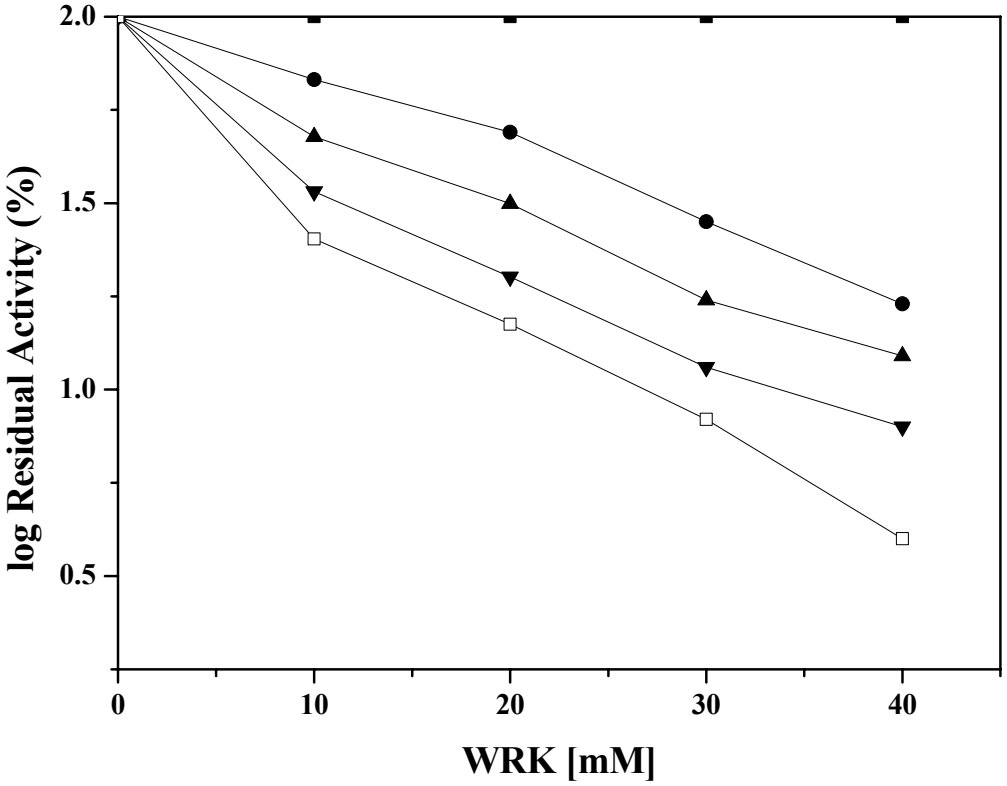


Fig. 7 Inactivation of Xyl II by WRK. Plot of residual activity *versus* time for the inactivation of TSC by WRK. TSC (10 µg) was incubated with WRK 0 mM (■), 5 mM (●), 10 mM (▲), 20 mM (▼), and 30mM (□) at 25 °C.

The three dimensional structures and site directed mutagenesis studies have revealed several highly conserved residues that are in the active site, playing crucial role in the catalysis. As discussed in the introduction of this Chapter, carboxyl acid group residues play important role in positioning of the substrate into the active site.

SUMMARY

Studies involving OPTA as the chemical initiator for fluorescent chemoaffinity labeling and TNBS as chemical modifier have revealed the presence of a single lysine residue in the active site of TSC. The kinetics of inactivation of TSC with OPTA revealed that complete inactivation occurred due to the binding of one molecule of OPTA to the active site of Xyl I. The formation of a single fluorescent isoindole derivative corroborated these findings. The kinetics of inactivation of TSC with TNBS revealed that complete inactivation occurred due to the binding of one molecule of TNBS to the active site of TSC and confirmed the presence of a single lysine residue in the active site of TSC. The substrate protection against OPTA and TNBS, suggested the presence of the lysine residue in the substrate-binding site. OPTA has been known to form a fluorescent isoindole derivative by cross-linking the proximal thiol and amino groups of cysteine and lysine. The involvement of cysteine in the formation of TSC-isoindole derivative has been proved by fluorometric and chemical modification studies on TSC, with group specific reagents and by amino acid analysis. Modification of cysteine and lysine residues by p- hydroxymercury benzoate and 2,4,6-trinitrobenzenesulfonic acid, respectively, abolished the ability of the enzyme to form an isoindole derivative with OPTA, indicating that cysteine and lysine participate in the formation of the isoindole complex. The involvement of essential carboxyl acid group residues has also been discussed

REFERENCES

Alzari, P. M., Hélè ne Souchon, and Roberto Dominguez 1996 The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum* Structure 4: 265-275.

Boyer, P.D. 1954.Spectrophotometric study of the reaction of protein sulfhydryl groups with organic mercurials. J. Am. Chem. Soc. 76:4331- 4337.

Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72:248-254.

Damude, H.G., S. G. Withers, G. K.Douglas ,R.C. Miller and Jr. R. A. J. Warren. 1995. Site-Directed Mutation of the Putative Catalytic Residues of Endoglucanase CenA from *Cellulomonas fimi.* Biochemistry. 34:2220-2224.

Davies, G J, M. Dauter, A. M. Brzozowski, M. E. Bjornvad, K. V. Anderson, M. Schülein. 1998. Structure of the *Bacillus agaradherans* family 5 endoglucanase at 1.6Å resolution and its cellobiose complex at 2.0Å resolution. Biochemistry 37:1926-1932.

Davies, G.J., A.M. Brzozowski, M. Dauter, A. Varrot, M. Ståhlberg. 2000. Structure and function of *Humicola insolens* family 6 cellulases: structure of the endoglucanase, Cel6B, at 1.6 Å resolution Biochem. J. 348:201-207.

Davies, G.J., S.P. Tolley, B. Henrissat, C. Hjort, M. Schülein. 1995. Structures of Oligosaccharide-Bound Forms of the Endoglucanase V from *Humicola. insolens* at 1.9Å resolution Biochemistry 34:16210-16220.

Davies, G.J., V. Ducros, R.J. Lewis, T.V. Borchert, M. Schülein. 1997.Oligosaccharide specificity of a family 7 endoglucanase: insertion of potential sugar-binding subsites J. Biotechnol. 57:91-100.

Davies,G.J., G.G. Dodson, R.E. Hubbard, S.P. Tolley, Z. Dauter, K.S. Wilson, C. Hjort, J.M. Mikkelsen, M. Schülein, Nature 365 (1993) 362-364.

Divne,C., J. Ståhlberg, T. Reinikainen, L. Ruohonen, G. Pettersson, J.K.C. Knowles, T.T. Teeri, T.A.Jones.1994.The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei.* Science 265:524-528.

Divne,C., J. Ståhlberg, T.T. Teeri, T.A. Jones. 1998. High-resolution crystal structures reveal how a cellulose chain is bound in the 50Å long tunnel of cellobiohydrolase 1 from *Trichoderma reesei*. J. Mol. Biol. 275:309-325.

Dominguez,R., H. Souchon, S. Spinelli, Z. Dauter, K.S. Wilson, S. Chauvaux, P. Béguin, P.M. Alzari .1995. A common protein fold and similar active site in two distinct families of β-glycanases Nat. Struct. Biol. 2:569-576.

group modification of proteins using Woodward's Reagent K. Anal. Biochem.151:327-333.

Habeeb, A. F.S.A.1966.Determination of free amino groups in proteins by trinitrobenzenesulfonic acid Anal. Biochem. 14:328-336.

Henrissat,B., P.M. Coutinho, G.J. Davies. 2001. A census of carbohydrate-active enzymes in the genome of *Arabidopsis thaliana* Plant Mol. Biol. 47:55-72.

Hilge, Mark, Sergio M Gloor, Wojciech Rypniewski, Oliver Sauer, Tom D Heightman, Wolfgang Zimmermann, Kaspar Winterhalter, and Klaus Piontek High-resolution native and complex structures of thermostable β-mannanase from *Thermomonospora fusca*–substrate specificity in glycosyl hydrolase family 5 1998 Structure 6: 1433-1444.

Juy, M., A.G. Amit, P.M. Alzari, R.J. Poljak, M. Claeyssens, P. Beüguin, J.P. Aubert.1992. Three-dimensional structure of a thermostable bacterial cellulase. Nature 357:89-91.

K.Palczewski, P. A. Hargrave, M. Kochman. 1983. o- phthalaldehyde, A fluorescent probe of aldolase active site, Eur. J.Biochem.137:429-435.

Kleywegt, Gerard J., Jin-Yu Zou, Christina Divne, Gideon J. Davies, Irmgard Sinning, Jerry Ståhlberg, Tapani Reinikainen, Malee Srisodsuk, Tuula T. Teeri and T. Alwyn Jones 1997,The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes J. Mol. Biol. 272, 383-397.

Kleywegt, Gerard J., Jin-Yu Zou, Christina Divne, Gideon J. Davies, Irmgard Sinning, Jerry Ståhlberg, Tapani Reinikainen, Malee Srisodsuk, Tuula T. Teeri and T. Alwyn Jones 1997,The crystal structure of the catalytic core domain of endoglucanase I from

Trichoderma reesei at 3.6 Å resolution, and a comparison with related enzymes J. Mol. Biol. 272, 383-397

Lynd, L.R., Weimer, P.J., Zyl, W.H., Pretorius I.S., 2002, Microbial Cellulose Utilization: Fundamentals and Biotechnology, Micro. Mol. Biol. Rev., 66:506-577.

M.Sandgren, J.Ståhlberg and C.Mitchinson. 2005. Structural and biochemical studies of GH family 12 cellulases: improved thermal stability, and ligand complexes. Prog. Biophys. Mol. Biol. 89:246-291.

Mackenzie,L.F. Q.P. Wang, R.A.J. Warren and S.G. Withers. 1998. Dynamic enzymatic resolution of thioesters. J. Am. Chem. Soc. 120:5583-5584.

Miller, G.M., 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal. Chem. 31, 426-428.

Ozaki,K. N. Sumitomo, Y. Hayashi, S. Kawai, S. Ito. 1994. Site-directed mutagenesis of the putative active site of endoglucanase K from *Bacillus* sp. KSM-330. Biochim. Biophys. Acta 1207:159-164.

Riedel,K., and K. Bronnenmeier. 1999. Active-site mutations which change the substrate specificity of the *Clostridium stercorarium* cellulase CelZ: Implications for synergism Eur. J. Biochem. 262:218-223.

Rouvinen,J., T. Bergfors, T.T. Teeri, J.K.C. Knowles, T.A. Jones. 1990 Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. Science 249: 380-386.

Sakon, J., D. Irwin, D.B. Wilson, P.A. Karplus. 1997. Crystal structure of endocellulase E4 from *Thermomonospora fusca* at 1.9Å resolution. Nat. Struct. Biol. 4:810-818.

Sakon, J.; Adney, W. S.; Himmel, M. E.; Thomas, S. R.; Karplus, P. A. 1996. Crystal Structure of Thermostable Family 5 Endocellulase E1 from *Acidothermus cellulolyticus* in Complex with Cellotetraose Biochemistry; 35:10648-10660.

Schou, C., G. Rasmussen, M.B. Kaltoft, B. Henrissat and M.Schülein. 1993. Stereochemistry, specificity and kinetics of the hydrolysis of reduced cellodextrins by nine cellulases Eur. J. Biochem. 217:947-953.

Schülein, M. 1997 Enzymatic properties of cellulases from *Humicola insolens*. J. Biotechnol. 57:71-81.

Simons, S. S., and D.F. Johnson. 1978. Reaction of o-Phthalaldehyde and thiols with primary amines: Fluorescence properties of 1- alkyl (and aryl) thio-2 -alkylisoindoles. Anal. Biochem. 90:705-725.

Ståhlberg, J., C. Divne, A. Koivula, K. Piens, M. Claeyssens, T.T. Teeri and T.A. Jones. 1996. Activity Studies and Crystal Structures of Catalytically Deficient Mutants of Cellobiohydrolase I from *Trichoderma reesei***.** J. Mol. Biol. 264:337-349.

Sulzenbacher,G., M. Schülein, G.J.Davies.1997.Structures of the endoglucanse I from *Fusarium oxysporum*: Native, product, and 3, 4-epoxybutyl ß-D-cellobioside inhibited forms at 2.3Å resolution Biochemistry 36:5902-5911.

Sulzenbacher,G., H. Driguez, B. Henrissat, M. Schülein, G.J. Davies. 1996. Structure of the *Fusarium oxysporum* endoglucanase I with a non-hydrolysable substrate analogue: substrate distortion gives rise to a pseudo-axial orientation for the leaving group. Biochemistry 35: 15280-15287.

U. Sinha and J. M. Brewer. 1985 A spectrophotometric method for quantitation of carboxyl

Varrot, A., S. Hastrup, M. Schülein, G.J. Davies. 1999 Crystal structure of the catalytic core domain of the family 6 cellobiohydrolase II, Cel6A, from *Humicola insolens*, at 1.92 Å resolution Biochem. J. 337:297-304.

Wang,Q.P., D. Tull, A. Meinke, N.R. Gilkes, R.A.J. Warren, S.G. Withers. 1993. Glu280 is the nucleophile in the active site of *Clostridium thermocellum* CelC, a family A endo-beta-1,4-glucanase J. Biol. Chem. 268:14096-14102.

Wilson, D.B., 1988. Cellulases of *Thermomonospora fusca*. In: Biomass, Part A: Cellulose and Hemicellulose. In: Wood, W.A., Kellogg, S.T. (Eds.), Met. in Enz., vol. 160. Academic Press, San Diego, CA, pp. 314-323.

Wolfgang, D. E. and D. B. Wilson. 1999. Mechanistic Studies of Active Site Mutants of *Thermomonospora fusca* Endocellulase E2. Biochemistry. 38:9746-9751.

Zhang S., Barr B. K., and Wilson D. B., 2000. Effects of noncatalytic residue mutations on substrate specificity and ligand binding of *Thermobifida fusca* endocellulase Cel6A**,** Eur. J. Biochem., 267, 244 - 252.

Zhang, S.,D.B. Wilson.1997.Surface residue mutations which change the substrate specificity of *Thermomonospora fusca* endoglucanase E2. J. Biotechnol. 57:101-113.

CHAPTER-IV

Unfolding and Chaperone Assisted Folding of 1, 4-β-D-Glucan Glucanohydrolase

ABSTRACT

To gain insight into the molecular aspects of unfolding / refolding of enzymes from extremophilic organisms 1,4-β-D-GlucanGlucanohydrolase (TSC) from an alkalothermophilic *Thermomonospora* sp. was used as the model system. Kinetics of denaturation / renaturation was monitored using intrinsic fluorescence studies. Denaturation studies using the structure - perturbing agent guanidium hydrochloride (GdnCl) indicated that TSC folds through a partially folded state that resembles molten globule. The protein fluorescence measurements suggested a putative intermediate state at 1.8 M guanidine hydrochloride with an emission maximum of 340 nm. The far-uv circular dichroism spectra revealed content of secondary structure similar to the native enzyme. Present investigations were undertaken to understand the folding of TSC using α-crystallin. α-Crystallin is a multimeric protein that has been shown to function as a molecular chaperone and shares extensive structural homology to small heat shock proteins. Here we studied the refolding of GdnCl denatured TSC from its molten globule state (TSC-m complex) in the presence and absence of α-crystallin to elucidate the molecular mechanism of chaperone-mediated in vitro folding. Our results, based on intrinsic tryptophan fluorescence and the fluorescent hydrophobic probe 1-anilinonaphthalene-8- sulfonate (ANS) binding studies, suggest that α-crystallin formed a complex with a putative intermediate molten globule – like intermediate in the refolding pathway of TSC. Reconstitution of the active TSC was observed on cooling the α-crystallin • TSC -m complex to 4°C. Addition of α-crystallin to the molten globule – like intermediate of TSC (TSC-m complex) complex initiated the refolding of TSC with 69 % recovery of the biological activity of the enzyme.

INTRODUCTION

The question of how protein folds into its unique, compact, highly ordered, and functionally active form is one of the most intriguing and perplexing questions of structural and cellular biology. Protein folding is determined by the information encoded by the amino acid sequence and proceeds *in vivo* by the same spontaneous mechanism (Anfinsen, 1973). The folding and assembly of a protein into its biologically active conformation is a complex succession of reactions involving the formation of secondary and tertiary structures and domains, and the oligomerization of folded monomers (Jaenicke, 1987). *In vivo* folding and assembly of proteins occur in a highly complex heterogeneous environment, in which high concentrations of proteins at various stages of folding and with potentially interactive surfaces coexist that may change the folding potentials inherent in the sequence (Jaenicke and Rudolph, 1983). Chemical denaturants like guanidinium chloride (GdnCl) and urea have long been known to unfold proteins by stabilizing the unfolded state compared to the native state. Unfolding by GdnCl and urea has since then become the standard method to determine protein stability and to investigate the kinetics and mechanisms of protein folding reactions. Empirically, a linear correlation between the free energy for unfolding (ΔG^0) and denaturant concentration ([D]) was found in the region of the unfolding transition. The correlation coefficient of this linear free energy relationship is commonly termed m_{eq}-value

$$m_{eq} = \delta \Delta G^0 / \delta [D]$$

It reflects the denaturing strength of the denaturant. Linear relationships were also found between denaturant concentration and the logarithm of the rate constants for protein folding (k_f) and unfolding (k_u) indicating that the activation free energies for folding $\delta \Delta G^0_f$ and unfolding $\delta \Delta G^0_u$) are also linearly dependent on denaturant concentration. Accordingly, kinetic m-values were defined as:

$$m_f = \frac{\partial \Delta G_1^0}{\partial [D]} \text{ and } m_u = \frac{\partial \Delta G_u^0}{\partial [D]}$$

Refolding of the denatured proteins in vitro has been an important issue at the fundamental as well as at the biotechnological level. Denaturation and renaturation are thermodynamic processes, involving a change in free energy and large changes in

conformation between the denatured and the native states (Arai and Kuwajima, 2004). Recently, a number of accessory proteins have been identified that affect the folding and subsequent assembly of proteins. These include the protein isomerases catalyzing cis-trans- isomerization of peptide bonds or disulfide exchange (Lang *et al*., 1987) and the polypeptide binding proteins termed as "molecular chaperones". The molecular chaperones are required for the successful folding, assembly, transport and degradation of proteins within the cell (Hendrick and Hartl, 1993). The molecular mechanism of refolding of proteins by chaperones now appears to be an important consideration for defining how proteins fold in vivo. α-Crystallin, the major protein of the vertebrate eye lens, consists of two types of highly homologous 20-kDa subunits, α- A and α- B. The A and B chains noncovalently self-associate to form a large macromolecular complex of approximately 40 subunits (Siezen and Argos, 1983). Spectroscopic data provided strong evidence that the secondary structure of α-crystallin is dominated by β-sheets (Montfort *et al*., 2001). α-Crystallin shares both sequence and structural homology with small heat shock proteins (sHSPs) that are known to act as molecular chaperones (Klemenz *et al*, 1991). The ability of α-crystallin to suppress aggregation of damaged proteins play a crucial role in maintaining the transparency of the ocular lens, and the failure of this function could contribute to the development of cataracts (Horwitz, 1993; Kelley *et al*., 1993; Gronen *et al*., 1994 and Lorimer and Baldwin, 1998). α-Crystallin has been reported to bind the temperature induced molten globule state of proteins and prevent photo-aggregation of γ- crystallin by providing hydrophobic surfaces (Raman and Rao, 1994). α-Crystallin has been reported to be functionally equivalent to sHSPs namely murine HSP25 and human HSP27 in refolding of α- glucosidase and citrate synthase in vitro (Jakob *et al*.,1993). Despite the growing interest in the chaperone action of α-crystallin, little is known about its mechanism of chaperoning. Elucidating the mechanistic details underlying the efficient refolding of proteins by chaperones now appears to be an important consideration for defining how proteins fold in vivo. Cellulases have raised enormous interest in the past decades in view of their biotechnological applications. However, the reports on unfolding/folding studies of this class of enzyme are scarce and especially of enzymes from extremophiles. For the functional in vitro analysis of α-crystallin, 1,4-β-D-Glucan Glucanohydrolase (TSC) from *Thermomonospora* sp. was used as the model system. Experimental evidence presented in this chapter serves to implicate that the chaperone α-crystallin stabilizes the

molten globule state of TSC and thus restrains the non-native conformer from exploring unproductive pathways. This is the first report on α-crystallin-mediated reconstitution of an active cellulase.

MATERIALS AND METHODS

Materials

1- Anilinonaphthalene -8- sulfonate (ANS), α-crystallin, CMC and GdnCl were obtained from Sigma Chemical Co. USA. All buffer salts were from standard commercial sources and of highest quality available. All other chemicals used in the present work were of analytical grade.

Methods

Purification and assay of TSC:

TSC was produced and purified according to the procedure described in Chapter II. Carboxymethyl cellulase activity was assayed by mixing the suitably diluted enzyme with 0.5 ml of 1% CMC in a final volume of 1ml and incubating the mixture at 50°C for 30 min. The reducing sugar released was determined by the dinitrosalicylic acid method. One unit of carboxymethyl cellulase activity was defined as the amount of enzyme that produced 1μmole of glucose equivalent per minute from CMC, under assay conditions. Protein concentration was determined according to the method of Bradford, using bovine serum albumin as standard.

Denaturation / renaturation studies of TSC:

All denaturation and renaturation experiments were carried out in 50 mM sodium phosphate buffer, pH 7.0. TSC was incubated with varying concentrations of GdnCl for 1.5 h at 28°C, and the accompanying structural changes were investigated using fluorescence and circular dichroism measurements. Renaturation was initiated by rapidly diluting the sample (TSC at varying states of denaturation) in the same buffer with or without α-crystallin, and the carboxymethyl cellulase activity at various times of refolding was measured. Renaturation of TSC was also carried out by substituting α-crystallin with bovine serum albumin (0.6 mg /ml).

Circular dichroism and fluorescence studies

Circular dichroism (CD) spectra were recorded on a JASCO J600 model spectropolarimeter. Changes in the secondary structure of TSC induced by the denaturant GdnCl were monitored in the far UV region (200-250 nm) using a 1 mm path length cell. The tertiary structure was monitored in the near UV region (250-320 nm) using a 10 mm path length cell. The enzyme concentration in these experiments was 0.5 mg/ml. Mean residue ellipticities [θ] (expressed as degree cm^2dmol^{-1}) were determined according to (Labhardt, 1986).

Fluorescence spectra were recorded with a Perkin-Elmer LS 50B spectrofluorimeter equipped with a Julabo F20 water bath. The excitation and emission wavelengths have been mentioned in the legends to figures.

RESULTS AND DISCUSSION

Folding intermediates of TSC

The CD spectrum of TSC in the far – UV region (200 - 260 nm) exhibited a strong negative ellipticities in the region 215-222 nm and a weaker one at 208 nm, characteristic of a protein having an α- helix. TSC was incubated with increasing concentrations of the denaturant GdnCl, and the changes in the negative CD band in the far UV region were monitored. The mean residue ellipticities obtained at 220 nm ($[\theta]_{220}$) were normalized with respect to that in the absence of GdnCl and plotted against the respective GdnCl concentration. A decrease in the negative ellipticity was observed with the addition of GdnCl, and at 1.8 M GdnCl the $[\theta]_{220}$ decreased by almost 51.3% of that in the absence of GdnCl. Further increase in the denaturant concentration resulted in a loss of negative ellipticity until there was a total loss of structure of the CD band in 6 M GdnCl indicating a considerable loss of secondary structure (Fig. 1). 6M GdnCl converted TSC into unfolded polypeptides, and this state has been referred to as TSC-u. A molten globule is an intermediate state in the folding pathway of a protein which has a compact denatured state with native-like secondary structure. Hydrophobic interactions weakly associate side chains, but fixed side-chain packing is lacking. We used the fluorophore ANS to determine the relative amount of exposed hydrophobic surfaces in the folding intermediates of TSC. ANS is not fluorescent in aqueous solutions (λ_{em} 511 nm); however, on addition of proteins containing hydrophobic pockets its emission maximum shifts to shorter wavelengths and the emission intensity is enhanced. As shown in Fig. 2 the binding of ANS to TSC was measured as a function of GdnCl. A maximum increase in the ANS fluorescence (λ_{em} 505 nm) was observed at 1.8 M GdnCl indicating maximum exposure of hydrophobic surfaces in this state of TSC. At higher concentrations of the denaturant, a decrease in the intensity of the dye fluorescence was observed which was accompanied by a shift in the λ_{em} toward red indicating unfolding of TSC.

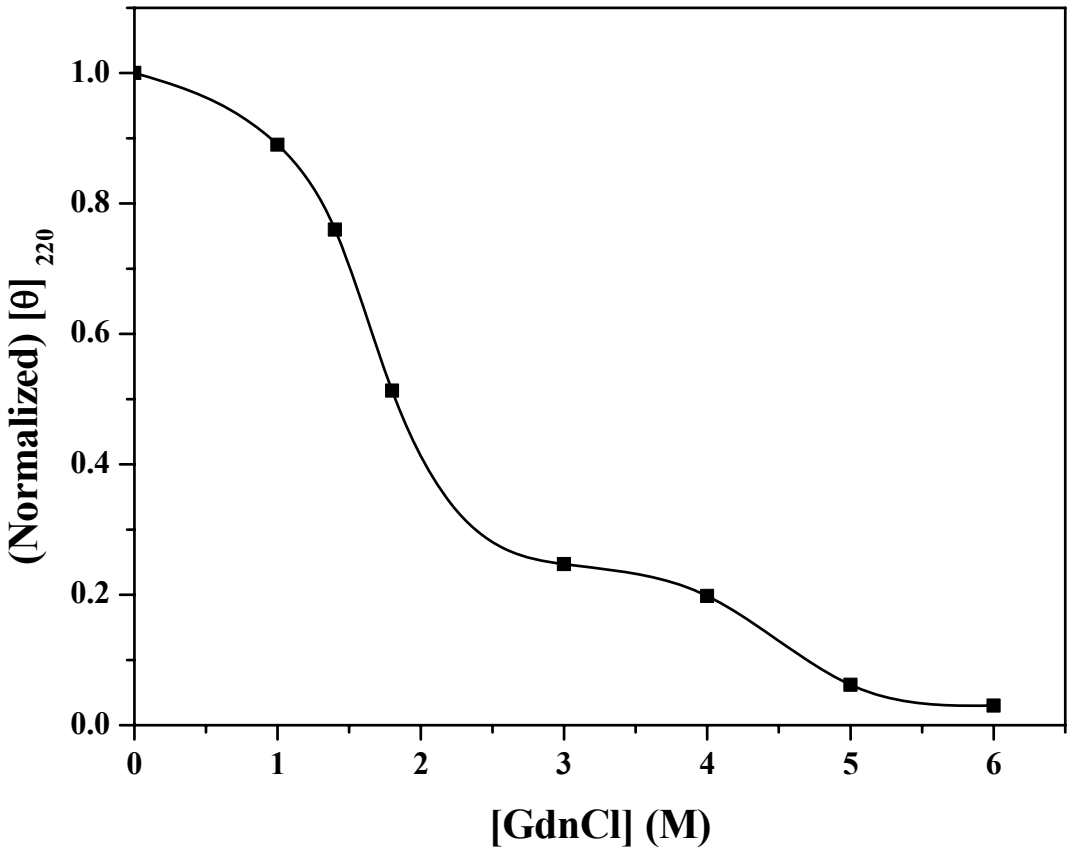


Fig. 1 Dependence of mean residue ellipticity of TSC at 220 nm on GdnCl. TSC was incubated with varying concentrations of GdnCl for 1.5 h at 28°C in 50 mM sodium phosphate buffer, pH 7, and the CD spectra in the far UV region (200–250 nm) were recorded. Protein concentration was 0.5 mg/ml, and a 1 mm path length cell was used. The ellipticity values obtained were normalized with respect to that in the absence of GdnCl.

It has been shown that ANS has a much stronger affinity to the protein molten globule state, with a pronounced secondary structure and compactness, but without a tightly packed tertiary structure as compared with its affinity to the native proteins. The molten globule state is a productive on-pathway intermediate in the real folding reactions in many globular proteins. The accumulation of the molten globule state during folding is a remarkable phenomenon in globular proteins. Precise characterization of molten globules in many globular proteins and their relationship with the folding intermediates has been major subjects of experimental studies of protein folding (Labhardt, 1986). ANS has been widely used to detect the formation of molten globule- like intermediates in the folding pathways of several proteins. This state is characterized to be as compact as the

native protein with solvent - accessible hydrophobic regions and appreciable amount of secondary structure but no rigid structure. It was thus evident from the CD studies that at 1.8 M GdnCl TSC retains substantial amount of secondary structure (Fig.2) but very little tertiary structure.

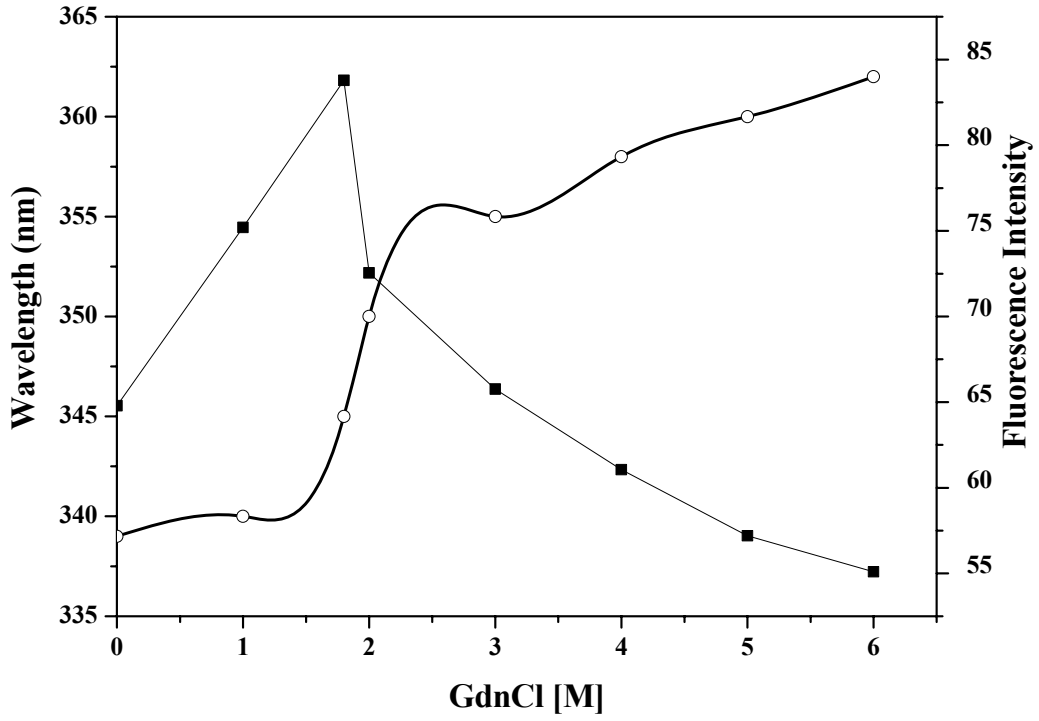


Fig.2 GdnCl-dependent exposure of hydrophobic surfaces of TSC measured by ANS fluorescence. ANS fluorescence intensity at 475 nm (■) and λ_{max} (○) on incubation of 0.25 mM GdnCl-treated TSC (described in legend to Fig. 1) with 10 mM ANS for 15 min. The concentrations indicated are the final concentrations. The samples were excited at 375 nm.

Altogether the CD and ANS binding studies revealed that at 1.8 M concentration of GdnCl, TSC partially unfolded to its molten globule state, which has been referred to as TSC-m.

Chaperone assisted renaturation of TSC

Attempts to refold TSC from the TSC-u state in the absence and presence of α-crystallin were unsuccessful (Fig.3). Similar results were obtained for progressively less denatured states (TSC denatured with 2-4M GdnCl). Further investigations were carried out to study the influence of the α-crystallin on the renaturation of TSC-m. The refolding of TSC-m was initiated at 28℃ in the absence /presence of α-crystallin; after 30min the

samples were shifted to varying temperatures and the CMCase activity recovered at different time intervals was measured.

As shown in Fig.3, TSC-m lacked the ability to spontaneously reconstitute active TSC. However, in the presence of α-crystallin the renaturation process at 4□ followed a sigmoidal time course. As can be observed from the inset of Fig. 3, there was no measurable CMCase activity for the first 15 min (lag phase). Thus similar to renaturation of oligomeric proteins, inactive monomers may be produced in an early folding step which then undergo additional folding and / or association prior to the assembly of TSC into active oligomers. The rate of activation beyond the lag phase was slow, and a maximum of 69.5 % of activity of CMCase activity was recovered in 6h. A value of 177 min was observed for $t_{1/2}$, where the activity recovered was half of the maximal extent. At 28°C, the renaturation process yielded a maximum 14 % of the CMCase activity, whereas the values for the lag phase and $t_{1/2}$ were similar to that observed at 4°C.

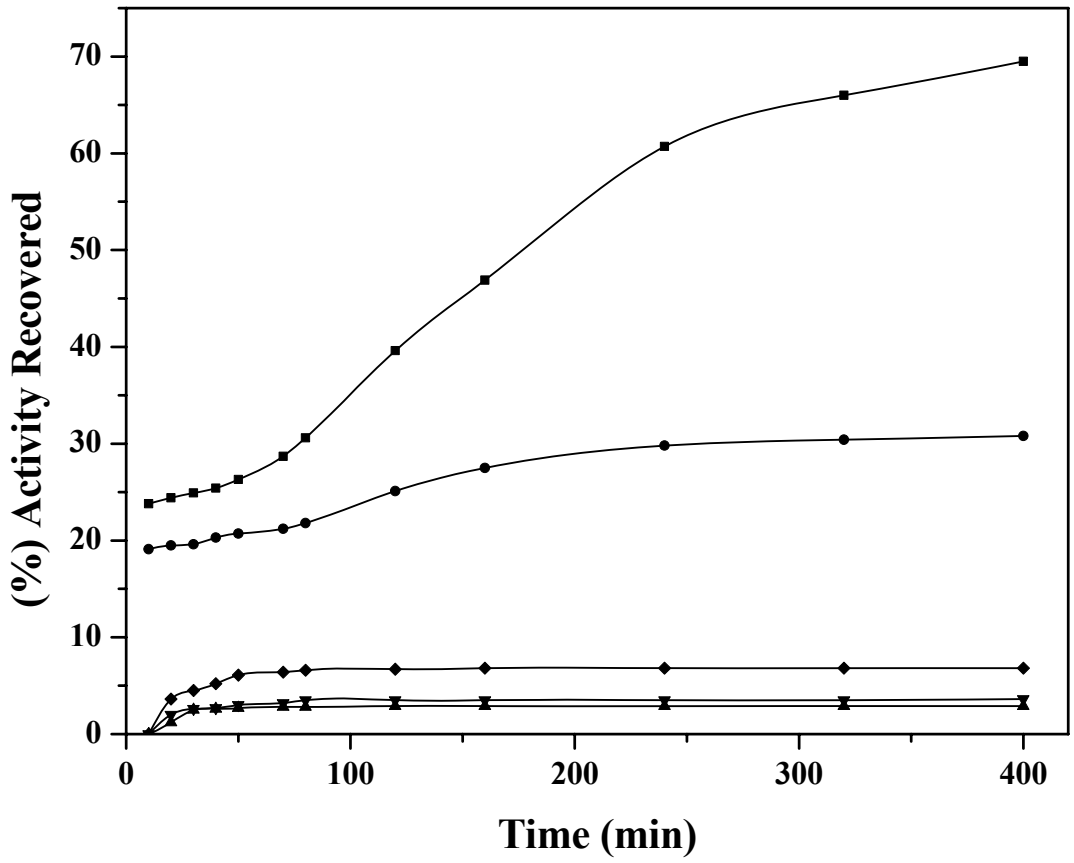


Fig. 3 Time course of renaturation of TSC in the absence or presence of α-crystallin.TSC at a concentration of 25 mM was incubated with 6 or 1.8 M GdnCl for 1.5 h at 28°C to yield TSC -u and TSC -m, respectively. The renaturation process was initiated at the same temperature by diluting 10 ml of the sample into a final volume of 1 ml of 50 mM sodium phosphate buffer, pH 7, with or without α-crystallin (final concentration 0.6 mg/ml). After 30 min the samples were kept at varying temperatures, and 100-ml aliquots of the refolding solution were withdrawn at various times of refolding and assayed for TSC activity. E represents refolding of TSC -m in the absence of α-crystallin and in its presence when the refolding solution was shifted from 28°C to the following temperatures: 4°C (■), 28°C (●), and 37°C (♦). (▲) and (▼) represents the refolding of TSC -u and progressively less denatured states of TSC (TSC denatured with 2–4 M GdnCl) in the absence and presence of α-crystallin under the experimental conditions described above. The percentage activity recovered was determined with reference to native TSC.

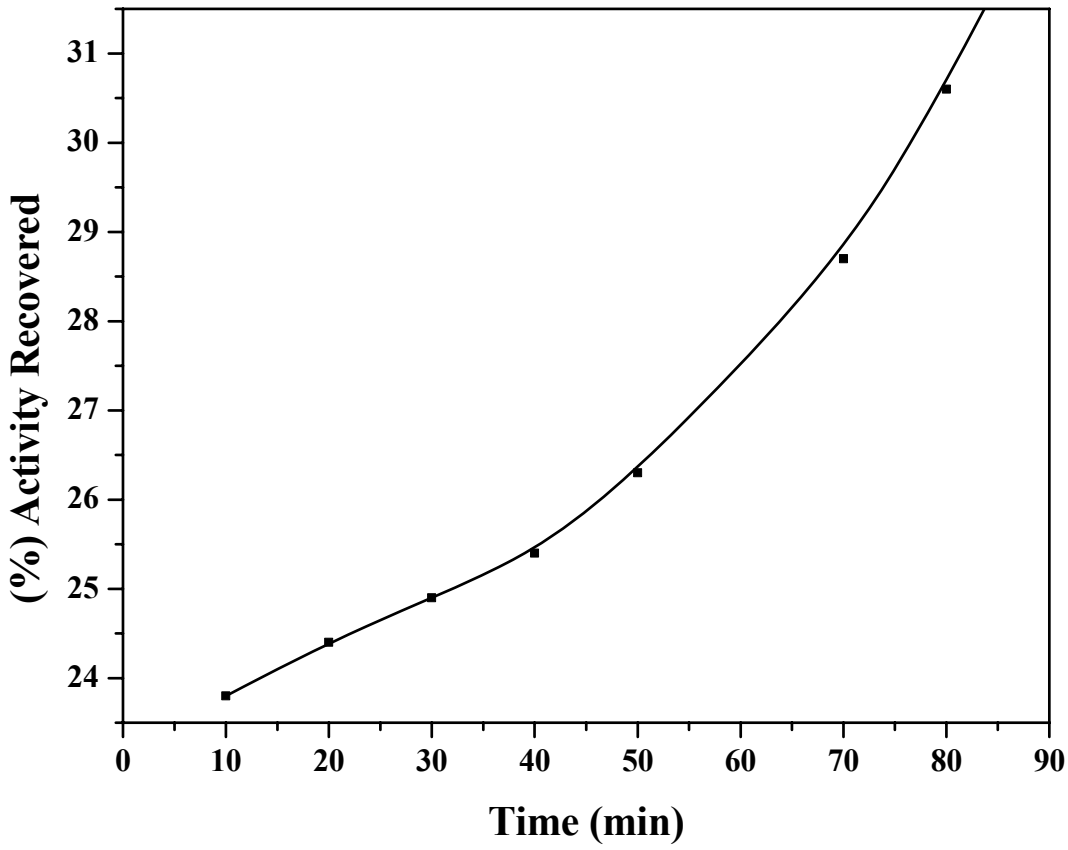


Fig. 4 The early time course of α-crystallin -assisted renaturation demonstrating the lag phase (first 40 min) at 4°C.

The α-crystallin-mediated renaturation of TSC was examined as a function of the chaperone concentration. As shown in Fig. 5, 22.4% of the original CMCase activity was recovered at the lowest concentration of α-crystallin (0.1 mg/ml). The extent of

renaturation increased in concentration - dependent manner, and a maximum 69.5% of the original activity was recovered at α-crystallin concentration of 0.6 – 0.8 mg/ml. The concomitant increase in the extent of renaturation with an increase in α-crystallin can be attributed to simple mass action effects, wherein an increase in the α-crystallin concentration would increase the collisional frequency so as to favor the formation of α-crystallin -TSC-m complex as opposed to forming non-native TSC.

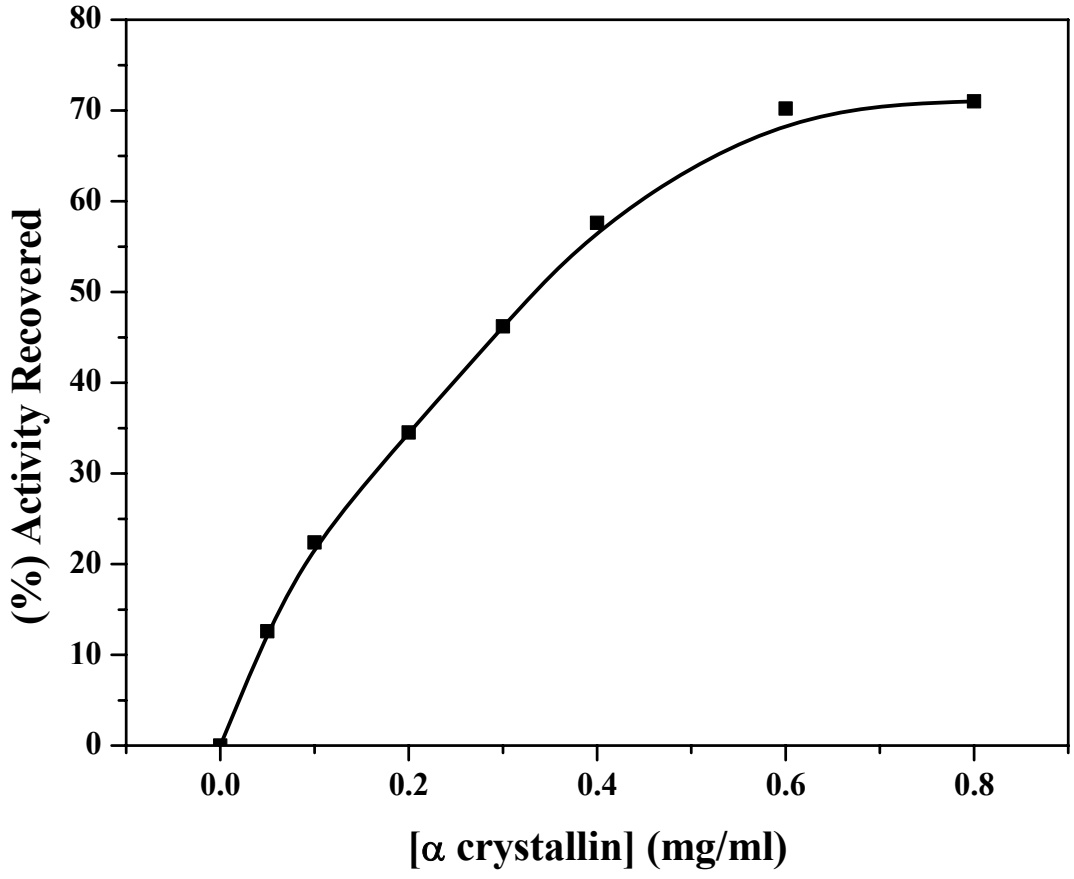


Fig. 5 Reactivation of TSC -m at varying concentrations of α-crystallin. TSC-m was renatured (as described in the legend to Fig. 3) in the presence of varying concentrations of α-crystallin. After 30 min the samples were kept at 4°C, and aliquots withdrawn after 6 h were assayed for TSC activity. The percentage activity recovered is with respect to native TSC.

To test the specificity of α-crystallin, the renaturation of TSC-m was also investigated in the presence of bovine serum albumin alone (0.6 mg/ml), under the conditions described for renaturation with α-crystallin. It was observed that unlike α-crystallin bovine serum

albumin failed to mediate the reconstitution of active TSC. α-Crystallin mediated reconstitution of active TSC was obtained only from its TSC-m state (Fig. 3), indicating that the chaperone probably traps the TSC in a conformation resembling the molten globule. Evidence for this observation was provided by

α-Crystallin forms a complex with folding intermediate

Fluorescence studies were performed to confirm that the TSC bound to α-crystallin exists in the molten globule state. The tryptophanyl fluorescence of native, denatured, and α-crystallin -bound TSC is shown in Fig. 6. Native TSC exhibited an emission maximum at 339 nm, whereas in 6 M GdnCl the emission maximum was shifted to 357 nm which corresponds to the fluorescence maximum of tryptophan in aqueous solution during unfolding process. The TSC bound to α-crystallin exhibited an emission maximum at 344 nm indicating that the tryptophans in the bound form of TSC are more exposed to the solvent than the native enzyme.

The results revealed that the conformation of TSC bound to α-crystallin is neither native-like nor completely unfolded but a partially folded intermediate resembling the molten globule. Thus by confiscating the molten globule state of TSC in the form of a stable binary complex, α-crystallin is able to suppress their interaction that would otherwise lead to aggregation.

Temperature dependence of the exposure of hydrophobic surfaces of α-crystallin* •TSC *-m complex

The thermodynamic properties of the unfolding reactions of globular proteins are now known accurately as a function of temperature through calorimetric studies. The unfolding reactions of different proteins display certain common properties. The enthalpy of unfolding depends on the temperature at which unfolding occurs, which can be varied by adjusting pH or guanidine hydrochloride concentration. "Hydrophobic interaction" refers to the process in which a hydrophobic side chain of an unfolded protein is taken out of H_20 and is buried in the interior of a protein through folding. The temperature dependence of the hydrophobic interactions in protein folding has been studied earlier (Simons and Johnson, 1978; Simons *et al*., 1979; Palczewski *et al*., 1983). Maximum stabilization of these interactions is observed at high temperature where the enthalpy is the dominating factor in determining the stability, and as the temperature is decreased the interactions are weakened.

The temperature-shift experiments (Fig. 3) revealed an increase in the α-crystallin - mediated reconstitution of TSC with the decrease in the temperature of the refolding solution implying that the hydrophobic interactions play a crucial role in the formation of α-crystallin • TSC -m complex.

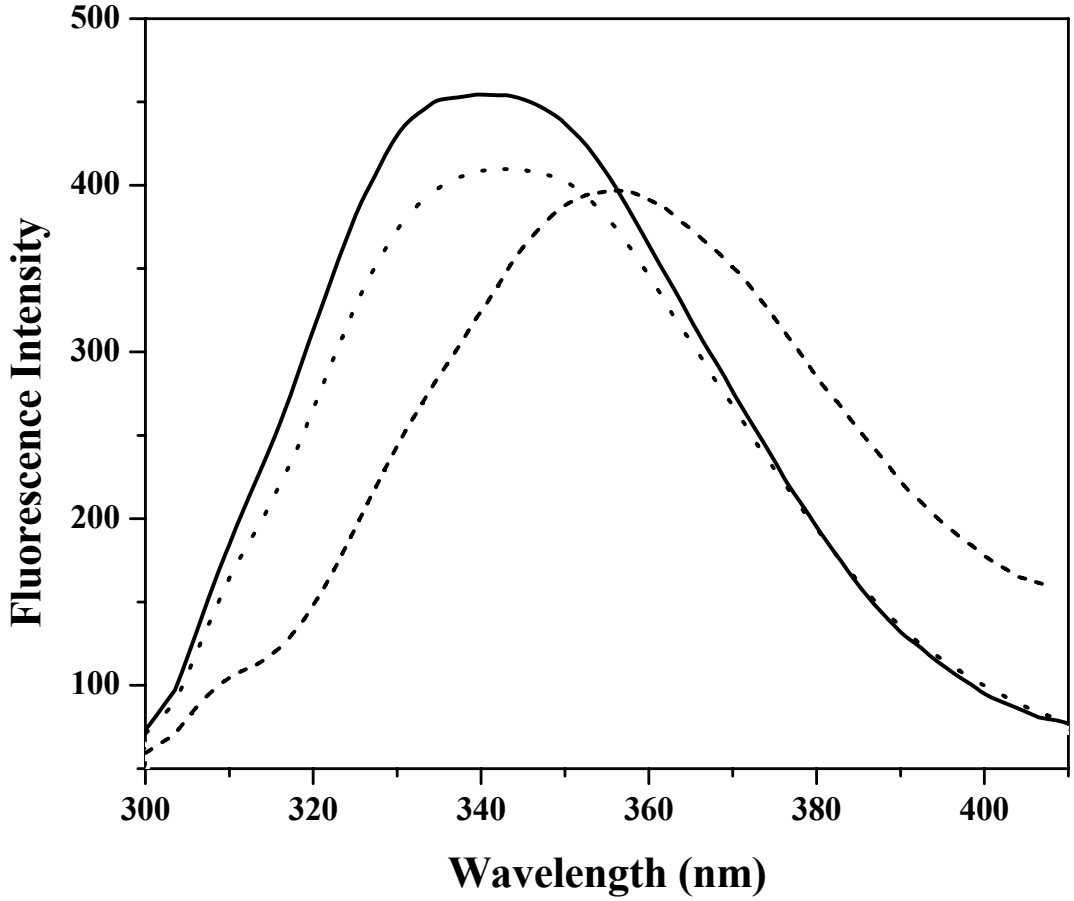


Fig. 7 Tryptophanyl fluorescence of free and α-crystallin bound TSC. TSC at a concentration of 25 μM was incubated with 1.8 M GdnCl for 1.5 h, and a further 10 ml of the sample was diluted into a final volume of 1 ml in 50 mM phosphate buffer, pH 7, containing 0.6 mg/ml α-crystallin at 37°C. After 30 min incubation the refolding solution was diluted to 2 ml, and the tryptophanyl fluorescence was recorded. The fluorescence spectrum of α-crystallin bound TSC (……) was obtained on subtracting the spectrum of α-crystallin • TSC -m complex from that of α-crystallin. _____ and - - - represent the fluorescence spectra of native and denatured TSC in the absence or presence of 6 M GdnCl, respectively. All samples were excited at 295 nm.

Attempts were made to correlate temperature-mediated alterations in the hydrophobic surfaces of the α-crystallin • TSC-m complex to reconstitution of active TSC, using ANS

as a probe for apolar binding sites whose fluorescence is dependent on the hydrophobicity of the environment. As shown in Fig. 8, presence of α-crystallin • TSC -m complex incubated at 37°C resulted in a blue shift in the ANS fluorescence from 511 to 505 nm accompanied by an increase in fluorescence intensity; however, in the presence of the complex incubated at 4°C a 35% decrease in the dye fluorescence was observed compared with that at 37°C (Fig. 8).

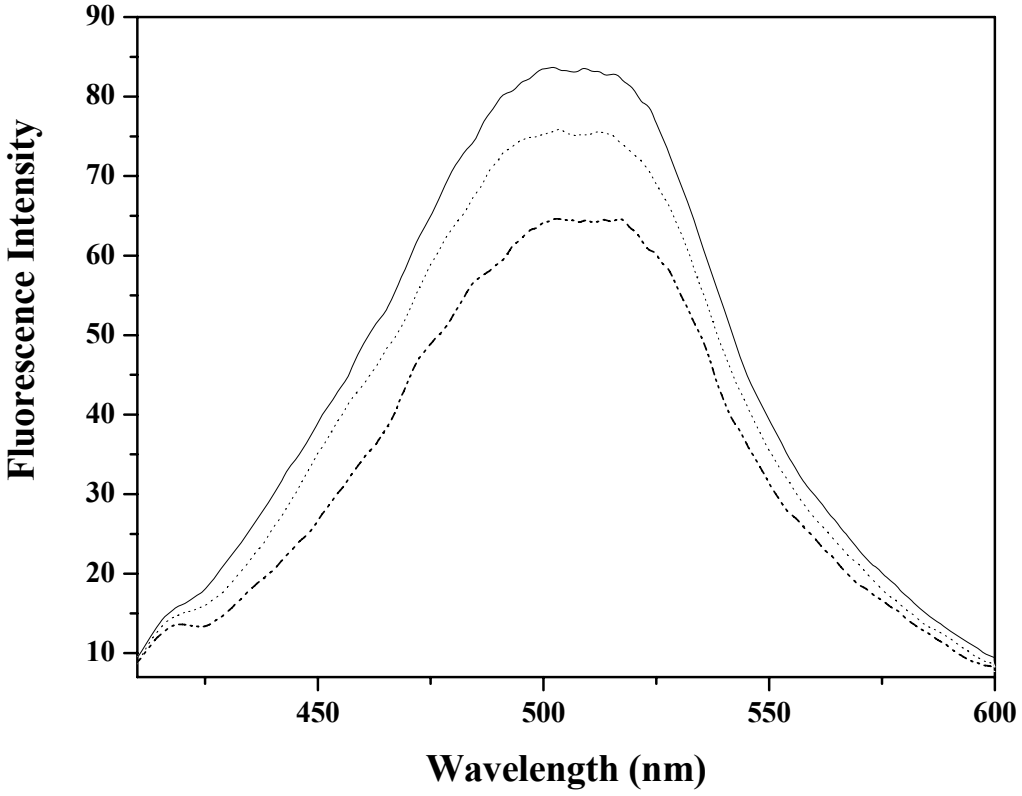


Fig. 8 Temperature-dependent exposure of hydrophobic surfaces of α-crystallin • TSC -m complex measured by ANS binding. Renaturation of 25 μM TSC -m was initiated at 37°C in 50 mM phosphate buffer containing 0.6 mg/ml α-crystallin preincubated at the same temperature for 2 h. After 30 min the samples were incubated at 37°C (__)or shifted to 4°C (----- -) ; furthermore, ANS (final concentration 100 mM) was added after 12 h, and the fluorescence was recorded at the respective temperatures at 1 h incubation, with the excitation wavelength fixed at 375 nm. (-..-..) represents ANS fluorescence at TSC alone.

These results indicate that at 37°C the complex exists in a state with hydrophobic binding sites that are accessible to ANS; however, a decrease in the incubation temperature to 4°C probably mediates a conformational change in the complex that is accompanied by internalization of the hydrophobic surfaces previously exposed. This probably further acts to weaken the hydrophobic interactions holding the α-crystallin • TSC -m complex

and thus reduces the affinity of α-crystallin for the substrate protein further, allowing reconstitution of active TSC (Fig. 3). This observation also explains the cold labiality of α-crystallin • TSC -m complex and the inability of α-crystallin to reconstitute active TSC when refolding was initiated at 4 °C.

Fluorescent chemoaffinity labeling of* TSC *renatured in the presence of α-crystallin

Fluorescent chemoaffinity labeling studies were performed using *o*-phthalaldehyde as the chemical initiator to shed some light on the conformation of TSC renatured in the presence of α-crystallin. Chemoaffinity labeling is a powerful technique and combines some of the advantages associated with the photo-activated and electrophilic affinity labeling. *O-Phthalaldehyde* is a bifunctional agent that cross-links -SH and -NH_2 groups situated in close proximity to form an isoindole derivative that exhibits strong fluorescence (Lang, 1978; Langer *et al*., 1992). TSC reacts with *o*-phthalaldehyde resulting in the formation of fluorescent TSC -isoindole derivative at the active site (λ_{ex} 338 nm; λ_{em} 410 nm).

As shown in Fig. 8, incubation of the α-crystallin renatured TSC with *o*-phthalaldehyde resulted in the formation of TSC -isoindole derivative as observed with the native enzyme. However, TSC when renatured in the absence of α-crystallin failed to form the derivative. These results suggest that α-crystallin mediates refolding of TSC to a conformation similar to that of the native enzyme.

α-Crystallin operates by providing hydrophobic surfaces that interact with the molten globule state of TSC. The hydrophobic interactions play an important role in the formation of α-crystallin • TSC -m complex. Conformational changes have been proposed to play a major role in the binding of folding intermediates and in the discharge of polypeptides from molecular chaperones. The present investigation was carried out to gain some insight into the conformation of TSC interacting with the chaperone α-crystallin and to understand the mechanism underlying the reconstitution of active enzyme. The conditions for the unfolding of native TSC were chosen so that the unfolded enzyme or its folding intermediates would serve as a substrate for the α-crystallin-mediated reconstitution of active TSC. The work on denaturation studies using the structure-perturbing agent GdnCl revealed that the folding of TSC involves an intermediate that resembles the molten globule.

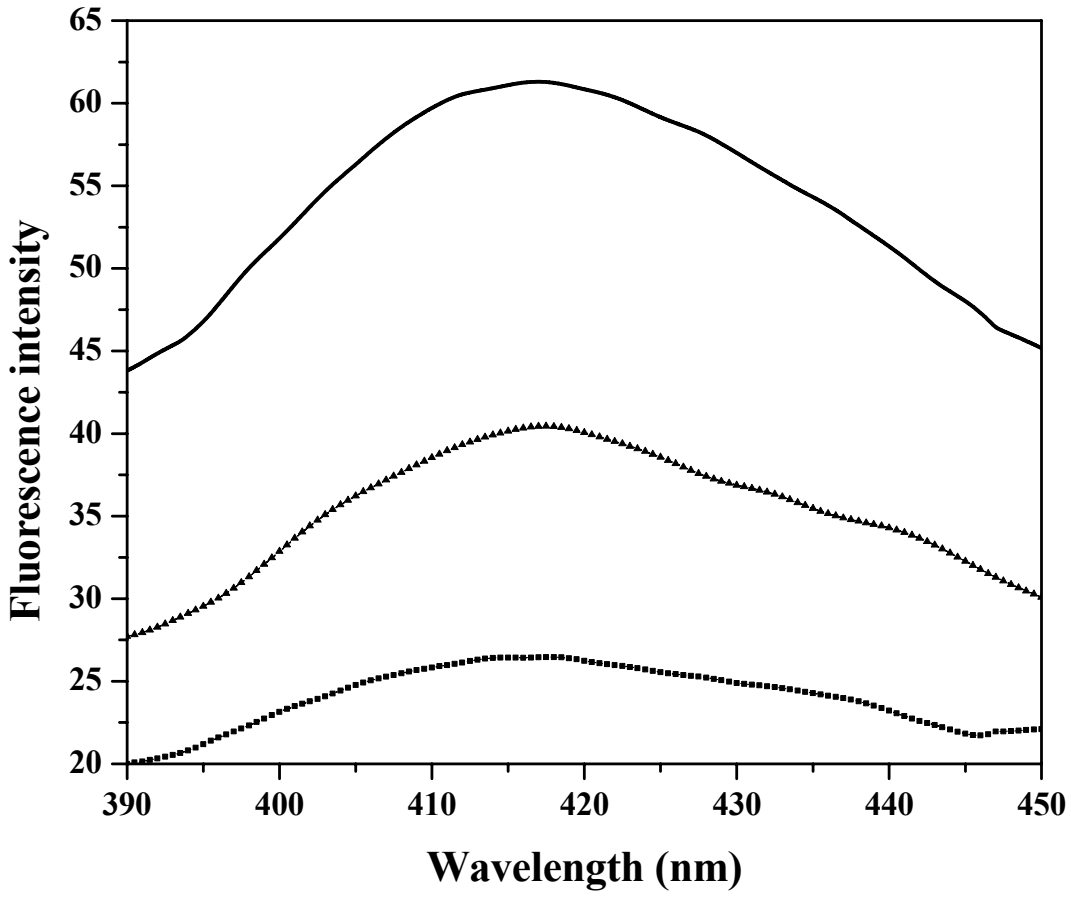


Fig. 9 Isoindole fluorescence of α-crystallin-renatured TSC on reaction with *o*-phthalaldehyde. TSC -m was renatured as described in legend to Fig. 3. After 30 min the refolding solution was further incubated at 4°C for 6 h, when maximum TSC activity was recovered. Furthermore, 10 ml of 3 mM *o*-phthalaldehyde was added to 2 ml of the renatured TSC, and the spectra were recorded after 30 min, with the excitation wavelength fixed at 338 nm. Similar experiments were repeated for renaturation of TSC -m in the absence of α-crystallin. ▲ and ■ represent the isoindole spectra of TSC renatured from TSC -m state in the presence and absence of α-crystallin, respectively. _____ represents the Isoindole spectra of native TSC.

The existence of molten globule like intermediates has been demonstrated with several proteins, and these intermediates are known to be involved in various cellular functions such as membrane translocation of proteins (Bychkova *et al*., 1988), chaperone-assisted protein folding (van det Goot *et al*.,1991), and also in various genetic diseases (Pace,1986). Interest in such intermediates is strong since they have been proposed to be an obligatory intermediate formed early in the folding pathway (Kim and Baldwin, 1990). Raman *et al*. (1997) showed that α-crystallin inhibits the aggregation of lysozyme and binds to aggregation- prone, molten-globule-like intermediate in the refolding

pathway. A common feature of the molten globule state is the exposure of hydrophobic surfaces that lead to aggregation of proteins during folding. The work present studies using TSC revealed that the chaperone α-crystallin operates by interacting with the hydrophobic regions that appear on the surface of molten globule state of TSC. This probably reduces the concentration of the free partially folded TSC (TSC -m) during renaturation and thus prevents loss of enzyme activity due to their hydrophobic aggregation. Lowering the temperature to 4°C induces a conformational change in the α-crystallin • TSC-m complex that is accompanied by a concomitant internalization of previously exposed hydrophobic surfaces. This acts to reduce the hydrophobic interactions involved in the formation of the complex and thus the affinity of the chaperone for the substrate protein further allowing reconstitution of the active TSC. Earlier it has been reported that α-crystallin does not prevent the photo-aggregation of γ-crystallin at low temperatures. However, it can do so at temperatures above 30°C. Our present investigation also supports this view, since α-crystallin -mediated reconstitution of TSC was observed when the refolding process was initiated at 28 and 37°C and not when initiated at 4°C which is attributed to the inability of the chaperone to prevent aggregation of TSC -m at low temperature. Delay experiments revealed the inability of α-crystallin to dissolve TSC aggregates formed in its absence implying that the chaperone α-crystallin should be present during stress conditions. The dependence of protein aggregation reactions on temperature and concentration is known. The inability of TSC -u and TSC -m to spontaneously reconstitute active TSC under the conditions used in the present investigation is due to the fact that aggregation competes with the correct folding pathway. The kinetic competition between refolding and aggregation has been reported to be a major determinant for lower yields or irreversibility in refolding of proteins *in vitro* (Langer *et al*., 1992). However, refolding of TSC may be possible under different experimental conditions, the presence of α-crystallin resulted in a substantial amount of reconstitution of active TSC from the TSC -m state. It has been reported that in *E. coli* a cascade of molecular chaperones mediate folding of proteins. This chapter reveals that α-crystallin is able to reconstitute TSC via interaction with its non-native conformer characterized by an increased surface hydrophobicity but a remarkably low degree of unfolding. The inability of α-crystallin to reconstitute TSC from its extended conformation implies that *in vivo* other chaperones may be involved in binding to the unfolded polypeptides and prevent premature misfolding and aggregation, whereas the proper folding and assembly may depend on the subsequent transfer of the partially

folded polypeptide to α-crystallin. Further evidence is provided by the observation that α-crystallin prevents aggregation of lens proteins induced by oxidative stress and UV radiation. These conditions are not likely to unfold protein molecules completely but induce formation of partially folded state with hydrophobic surfaces that result in its aggregation (Das *et al*., 1996).The most direct approach to determine the structures of folding intermediates, as well as the interactions that stabilize them, is to study populated intermediates (Douglas and Baldwin, 1998). Unfolding of cellulases from *Trichoderma reesei* were studied using Dithiothreitol, Brefeldin A and ionophore A23187 (Pakula *et al*., 2003). But the reports on chaperone mediated refolding of cellulase are scarce. The present investigation on α-crystallin adds to the information available on its chaperone function and to our knowledge, the present study is the first demonstration of α-crystallin mediated refolding of a cellulase.

SUMMARY

In the present investigation the conformation and microenvironment of the active site of 1,4-β-D-Glucan Glucanohydrolase has been probed. The enzyme was completely inactivated by low concentrations of GdnCl which was used as an active site perturbant. Fluorescence studies revealed the inactivation of the enzyme by GdnCl precedes gross conformational change. Denaturation studies using GdnCl indicated that the enzyme folds through a partially folded state that resembles the molten globule. The protein fluorescence measurements suggested a putative intermediate state at 1.8 M guanidine hydrochloride with an emission maximum of 340 nm. The Far-UV circular dichroism spectra revealed content of secondary structure similar to the native enzyme. Studies with the fluorescent apolar probe 1- anilinonaphthalene-8-sulfonate (1,8ANS) were consistent with the native or fully unfolded protein. Fluorescence experiments revealed that α crystallin interacts with the molten globule state of the enzyme and prevents its aggregation.

REFERENCES

Anfinsen, C.B. 1973. Principles that govern the folding of protein chains. Science 181:223−230.

Arai M. and K. Kuwajima. 2000. Role of the molten globule state in protein folding. In Adv. Pro. Chem. (Protein folding mechanisms) Academic press Edited by F. M.Richards, D.S. Eisenberg, P. S.Kim, C.R. Matthews. 53:209-271.

Baldwin R. L. 1986. Temperature Dependence of the Hydrophobic Interaction in Protein Folding Proc. Natl. Acad. Sci. U.S.A. 83:8069-8072.

Bychkova, V. E., R.H. Pain and O. B. Ptitsyn. 1988. The 'molten globule' state is involved in the translocation of proteins across membranes? FEBS Lett., 238: 231-234.

Das, K. P., J. M. Petrash, and W. K. Surewicz. 1996. Conformational Properties of Substrate Proteins Bound to a Molecular Chaperone α-crystallin, J. Biol. Chem. 271:10449-10452.

Douglas, V. L., and R. L. Baldwin. 1998. Protein Folding: Matching Theory and Experiment. Biophys. J. 75: 428-434.

Ellis, R.J. 1990. Molecular chaperones: The plant connection, Science 250:954-959.

Gloubinoff, P., J.T.Christeller, A.A. Gatenby and G.H.Lorimer. 1989. Reconstitution of active dimeric ribulose bisphosphate carboxylase from an unfolded state depends on two chaperonin proteins and Mg-ATP. Nature. 342: 884–888.

Groenen, P.J., K.B. Merck, W.W. de Jong, H.Bloemendal. 1994 Structure and modifications of the junior chaperone alpha-crystallin. From lens transparency to molecular pathology. Eur. J. Biochem. 225:1-19.

Hendrick, J.P., F.U.Hartl. 1993. Molecular chaperone functions of heat-shock proteins. Annu Rev Biochem. 62:349-84.

Horwitz, J. 1993. Proctor Lecture. The function of alpha-crystallin. Invest. Ophthalmol. Vis. Sci. 34, 10–21.

Jaenicke, R. 1987. Folding and association of proteins. Prog. Biophys. Mol. Biol. 49:117−237.

Jaenicke, R. and R. Rudolph.1986. Refolding and association of oligomeric proteins Met. Enzymol. (Enzyme Structure part L) Academic press edited by C.H.W.Hirs and S.N.Timasheff. 131: 218-250

Jakob, U., M. Gaestel, K. Engel, and J. Buchner. 1993. Small heat shock proteins are molecular chaperones. J. Biol. Chem. 268: 1517-1520.

Kassenbrock, C.K. and Kelly, R.B. 1989. Interaction of heavy chain binding protein (BiP/GRP78) with adenine nucleotides. EMBO J. 8: 1461–1467.

Kelley, M.J., I.David, N.Iwasaki, J.Wright, and T.R.Shearer, 1993. α-crystallin chaperone activity is reduced by calpain II in vitro and in selenite cataract. J. Biol. Chem. 268: 18844–18849.

Kim, P.S. and R.L.Baldwin. 1990. Intermediates in the Folding Reactions of Small Proteins, Annu. Rev. Biochem.59: 631-660.

Klemenz, R., E. Frohli, R.H. Steiger, R. Schafer, and A.Aoyama.1991. α B-Crystallin is a Small Heat Shock Protein. Proc. Natl. Acad. Sci. U.S.A 88:3652-3656

Labhardt, A. M. 1986. Folding intermediates studied by circular dichroism Met.Enzymol. (Enzyme Structure part L) Academic press edited by C.H.W.Hirs and S.N.Timasheff. 131: 126-135.

Lang, K., F.X.Schmid, G.Fisher, 1987. Catalysis of protein folding by prolyl isomerase. Nature 329:268-70.

Langer, T., C.Lu, H. Echols, J. Flanagan, M. K. Hayer, and F. U.Hartl. 1992. Successive action of DnaK, DnaJ and GroEL along the pathway of chaperone-mediated protein folding. Nature 356**:**683-689.

Lorimer, G. H. and T. O. Baldwin.1998. Lens crystallin: Chaperone like properties In Met.Enzymol. (Molecular chaperones) Academic press edited by J. Horwitz, Q. ling Huang, L. Ding, and M. Bova. 290:365-384.

Montfort, R.V., C. Slingsby and E. Vierlingt. 2001. Structure and function of the small heat shock protein/α-crystallin family of molecular chaperones. In Adv. Pro. Chem. (Protein folding in the cell) Academic press Edited by A. Horwich. 59: 105-156.

Pace, C. N. 1986. Determination and analysis of urea and guanidine hydrochloride denaturation curves Met. Enzymol. (Enzyme Structure part L) Academic press edited by C.H.W.Hirs and S.N.Timasheff. 131:266-280.

Pakula, T. M., M. Laxell, A. Huuskonen, J. Uusitalo, M. Saloheimo, and M. Penttila¨. 2003. The Effects of Drugs Inhibiting Protein Secretion in the Filamentous Fungus *Trichoderma reesei:* evidence for down-regulation of genes that encode secreted proteins in the stressed cells. J. Biol. Chem. 278, 45011-45020.

Palczewski, K., P. A. Hargrave, M. Kochman. 1983. o- phthalaldehyde, A fluorescent probe of aldolase active site, Eur.J.Biochem. 137:429-435.

Raman B., and C.M. Rao. 1994. Chaperone-like activity and quaternary structure of alpha-crystallin J. Biol. Chem. 269:27264-27268.

Raman, B., T. Ramakrishna and C. M. Rao. 1997. Effect of the chaperone-like alpha-crystallin on the refolding of lysozyme and ribonuclease A FEBS Letters, 416, 369-372 .

Siezen, R.J., and P.Argos. 1983. Structural homology of lens crystallins. III. Secondary structure estimation from circular dichroism and prediction from amino acid sequences. Biochim. Biophys. Acta. 748:56–67.

Simons S. S. and D.F. Johnson. 1978. Reaction of o-Phthalaldehyde and thiols with primary amines: Fluorescence properties of 1- alky (and aryl) thio – 2 – alkylisoindoles. Anal. Biochem. 90:705 – 725.

Simons, S. S., Jr. E. B. Thompson, D. F. Johnson. 1979. Fluorescent chemoaffinity labeling. Potential application of a new affinity labeling technique to glucocorticoid receptors, Biochemistry 18:4915-4922.

Surewicz, W. K., and P. R.Olesen. 1995 On the thermal stability of alpha-crystallin: a new insight from infrared spectroscopy. Biochemistry. 34:9655–9660.

van det Goot, F. G., J. M. Gonzales-Manas, J. H.Lakey, and F. A Pattus. 1991. 'Molten-globule' membrane-insertion intermediate of the pore-forming domain of colicin A. Nature 354: 408–410.

Zettlmeissl, G., R. Rudolph and R. Jaenicke. 1979. Reconstitution of lactic dehydrogenase. Noncovalent aggregation vs. reactivation. 1. Physical properties and kinetics of aggregation. Biochemistry. 18: 5567-5571.

CHAPTER-V

Crystallization and Preliminary X-ray Characterization of 1,4-β-D-Glucan Glucanohydrolase and its Amino Acid Sequence Analysis

ABSTRACT

The cellulase from Alkalothermophilic *Thermomonospora sp.* which depolymerizes cellulose has been crystallized by hanging drop vapour diffusion method. It has a molecular weight of ~ 14 kDa. Single crystals were grown in a week in the crystallization condition: 0.2 M Na Acetate + 0.1M Na Cacodylate pH 6.5 + 25% PEG 8000. Similar crystals were grown with Mg / Ca acetate in place of Na acetate salt. Crystals belong to primitive orthorhombic lattice with unit cell parameters a = 49.9 Å, b = 79.5 Å and c = 99.7 Å, and diffract to better than 2.3 Å resolution. Matthew's coefficient of 2.25 $Å^3$/Da indicates the presence of four monomers in the crystal asymmetric unit.

Amino acid sequence analysis

The research interest of this part of the thesis was to sequence the purified 1,4 – β – D – Glucan Glucanohydrolase. The amino acid sequence analysis of the peptides obtained after tryptic digestion was carried out. The N- terminal sequence is **GSVINPATRNY.** The C-terminal sequence is **ASHMDQTYYLCSDVNFV**

The other five peptides sequenced are **NYGCWLR, SMDAVGPWK, HVVFTIWK, LPDGQLCSGGLAEGGR, PLTWGSLDLVHR**

Of these **ASHMDQTYYLCSDVNFV** and **LPDGQLCSGGLAEGGR** showed sequence homology with the cellulase binding domain of other enzymes from *Themobifida fusca* and *Streptomyces coelicolor*. The conserved residues found in the multiple sequence alignment of these peptides, were cysteine, serine, phenylalanine, and glycine.

INTRODUCTION

Cellulases have a large number of biotechnological applications. They have been intensively studied at the biochemical and structural level in recent years. The knowledge of high resolution structures of cellulases is one of the ways to improve the catalytic efficiency of cellulases. An increase in the cellulase activity can also be achieved through protein engineering. Both the basic and applied aspects of cellulase research require knowledge of the atomic structure of cellulases. The three dimensional structure of TSC, therefore, would be a representative for such low molecular weight cellulases and should decipher the active site residues which are responsible for the bifunctional activity of the enzyme.

In general, cellulases are modular enzymes with molecular mass varying from approximately 14 kDa to 90 kDa, consisting of a catalytic domain connected to a cellulose binding domain by flexible linker region. Amino acid sequence comparisons, supplemented by hydrophobic cluster analysis, have allowed classification of glycoside hydrolases into more than 60 families, which, on the basis of crystallographic investigations, can now be grouped into "clans" of related structures. (Henrissat, 1991, 1997) Crystal structures for representative of the cellulase – containing GH families 5-9, 12, 45, 48 have been elucidated and revealed a remarkable variety of folding topologies. (Bourne and Henrissat, 2001).

3-D structures of cellulases

Family 6 cellulases

The *T. reesei* Cel6A core (formerly CBHII) was the first cellulase for which a 3-D structure was solved (Tomme *et al*., 1995). Since 1990, 3-D structures have also been reported for the endoglucanase Cel6A (formerly E2) from *T. fusca* as well as both the cellobiohydrolase Cel6A and the endoglucanase Cel6B from *Humicola insolens* (Tormo, *et al*., 1996, Szabo *et al*., 2001). The topology of these proteins is a variant of the β/α-barrel with a central β -barrel made up of seven parallel strands.

Family 7 cellulases

The 3-D structures of a number of cellobiohydrolases and endoglucanases from this family are known (Nagy *et al*., 1998, McLean *et al*., 2000; Boraston *et al*., 2001; Xie *et al*., 2001a; Xie *et al*., 2001b; Notenboom *et al*.,2002). The 3-D structure is a double β -sandwich, similar in topology to those found in the legume lectin family such as concanavalin A and the lichenases from family 16(Boraston *et al*., 2001). As with family 6, the structural difference between the cellobiohydrolases and the endoglucanases is revealed in the extent and nature of the loops flanking the substrate binding subsites: in the cellobiohydrolases these loops are extended and form an enclosed tunnel. The 3-D structures of both cellobiohydrolases and endoglucanases from family 7 revealed a `trio' of carboxylates in the active centre (Nagy *et al*., 1998; McLean *et al*., 2000; Boraston *et al*., 2001; Xie *et al*., 2001; Xie *et al*., 2001; Notenboom *et al*., 2002).

Family 8 cellulases

The structure of this family is an $(\alpha/\alpha)_6$-barrel formed by six inner, and six outer, α -helices. The architecture of the active site cleft gives room for at least five glucosyl binding subsites.

Family 9 cellulases

The structure of cellulases in this family consists of an $(\alpha/\alpha)_6$-helical barrel. Some structures show a tight association to an immunoglobulin-like domain such as the *C. thermocellum* Cel9A (formerly CelD) (Riedel et al, 1999). Other structures appear tightly bound to CBM from family 3 as witnessed for the *T. fusca* endoglucanase Cel9A (formerly E4) (Simpson and Barras,1999).

Family 45 cellulases

The native structure of *H. insolens* Cel45A (formerly EG V) was solved in 1993 and, later, a sited-directed inactive variant Asp10Asn structure including oligosaccharide was solved (Davies *et al*., 1993; Davies *et al*., 1995). The three structures, native, wild type with cellobiose and inactive soaked with cellohexose revealed a structure consisting of a six stranded β-barrel domain with long interconnecting loops. A 40Å groove exists along the

surface of the enzyme, and this contains the catalytic residues Asp10 (probably catalytic base) and Asp121 the catalytic acid.

Fig.1 shows the ribbon representation of the main fold of the catalytic domain in various glycosyl hydrolase families.

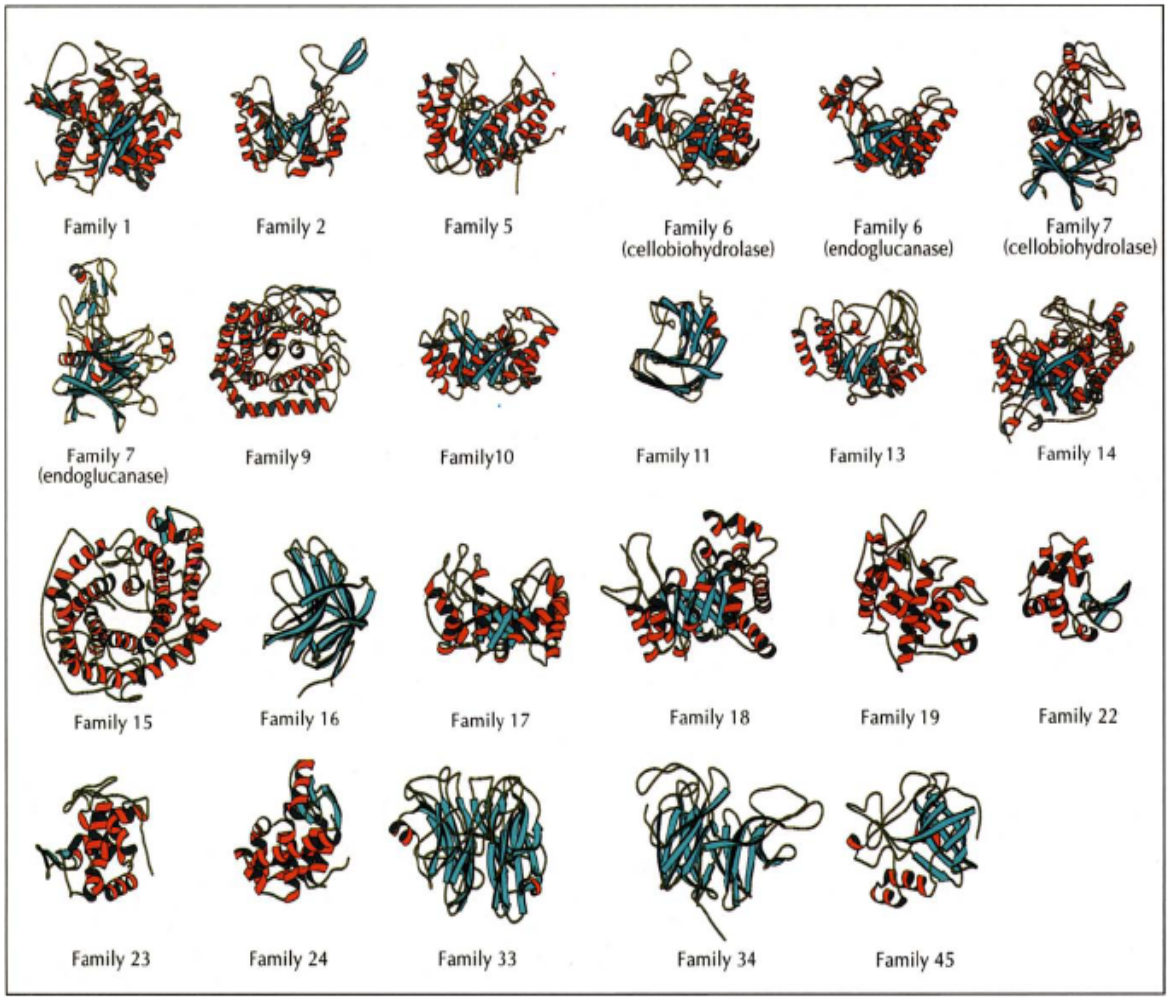


Fig. 1 Ribbon representation of the main fold of the catalytic domain in various glycosyl hydrolase families. β strands are shown in cyan and a helices in red.

Table 1 summarizes the representative crystal structures of cellulases and xylanases

Table-1

Crystal structures of some cellulases and xylanases

Enzyme	Source organism	Resolution	Reference
Family 5 endocellulase	*Acidothermus cellulolyticus*	2.4 Å	Sakon *et al*,1996
endoglucanase	*Aspergillus niger*	2.1 Å	Khademi *et al*, 2002
Endoglucanase Cel 5A	*Bacillus agaradharens*	1.75 Å	Varrot *et al*,2000
family 5 endoglucanase	*Bacillus agaradherans*	1.6 Å	Davies *et al*,1998
Alkaline endoglucanase	*Bacillus circulans*	2.2 Å	Hakamada et al, 2002
Alkaline cellulase K	*Bacillus licheniformis*	1.9 Å	Shirai *et al*, 2001
Cellulase K	*Bacillus licheniformis*	1.9 Å	Shirai *et al* 2001
Xylooligosaccharide deacytelase	*Bacillus subtilis*	1.9 Å	Vincent *et al*,.2003
Exo- 1-3-β-Glucanase Glycosyl hydrolase family 5	*Candida albicans*	1.9 Å	Cutfield *et al*,1999
exo-β-1,4-glucanase	*Cellulomonas fimi*	2.0 Å	Bedarkar *et al*,1992
β-1,4-xylanase	*Cellulomonas fimi*	2.0 Å	Bedarkar *et al*,1992
Xylanase 10C	*Cellvibrio japonicus*	1.9 Å	Pell *et al*,2004
Endoglucanase 5A	*Cellvibrio mixtus*	1.4 Å	Pires *et al*,2004
CeIF	*Clostridium cellulolyticum*	2.0 Å	Reverbel-Leroy *et al*,1998
IIIa scaffoldin CBD	*Clostridium cellulolyticum*	2.2 Å	Shimon *et al*, 2000
Family 5 endoglucanase	*Clostridium thermocellum*	2.3 Å	Domínguez *et al*, 1996
Cohesin domain	*Clostridium thermocellum*	2.45 Å	Shimon *et al*, 1997
Cel A	*Clostridium thermocellum*	1.65 Å	Alzari *et al*,1996
*ct*Cel9D-Cel44A	*Clostridium thermocellum*	2.1 Å	Najmudin *et al*, 2005
endoglucanase V	*Humicola insolens*	1.5Å	Davies *et al*,1996
endoglucanase I (EG1)	*Humicola insolens*	2·5 Å	Davies *et al*,1992
Family 45 endoglucanase	*Melanocarpus albomyces*	2.0 Å	Hirvonen *et al*, 2003
endoglucanase	*Nasutitermes takasagoensis*	1.4 Å	Khademi *et al*, 2002
CBH58	*Phanerochaete chrysosporium*	1.32 Å	Muñoz *et al*,2001
psychrophilic	*Pseudoalteromonas*	1.8 Å	Violot *et al*,2003

cellulase	*haloplanktis*		
xylanase A	*Pseudomonas fluorescens*	1·8 Å	Pickersgill *et al*, 1993
Cel2A family 12 endoglucanase	*Rhodothermus marinus*	1.8 Å	Crennell *et al*,2002
GH10 xylanase	*Streptomyces olivaceoviridis*	1.8 Å	Charnock *et al*,1998
Family 10 xylanase	*Streptomyces olivaceoviridis E-86*	2.0 Å	Fujimoto *et al*,2004
Xylanase	*Thermoascus aurantiacus*	1.8 Å	Natesh *et al*,1999
Xylanase	*Thermoascus aurantiacus*	0.89 Å 1.11 Å	Natesh *et al*,2003
β-mannanase	*Thermomonospora fusca*	1.5Å	Hilge *et al*,1998
Cel6A	*Thermomonospora fusca*	1.5Å	Larsson *et al*,2005
Cellobiohydrolase I	*Trichoderma reesei*	1.8 Å	Divne *et al*, 1994
Cellobiohydrolase II	*Trichoderma reesei*	2.0 Å	Rouvinen *et al*, 1990
CBHI	*Trichoderma reesei*	2·0 Å	Divne *et al*,1993
EGI	*Trichoderma reesei*	2·5 Å	Divne *et al*, 1993
endoglucanase I	*Trichoderma reesei*	3.6 Å	Kleywegt *et al*, 1997
Cel 12A	*Trichoderma reesei*	1.9 Å	Sandgren *et al*, 2001
CBHI	*Trichoderma reesei*	1.7 Å	Divne *et al*, 1998
Cellobiohydrolase Cel 7A	*Trichoderma reesei*	1.9 Å	Ståhlberg *et al*,2001

Crystallization

The essential feature of a crystal is its ordered and three dimensionally periodic internal structure. The crystallization of molecules from solution is a reversible equilibrium phenomenon, and the specific kinetic and thermodynamic parameters will depend on the chemical and physical properties of the solvent and the solute involved. The impetus for molecules to crystallize is the minimization of free energy which can be accomplished only when the molecules are arranged in asymmetrical and periodically repetitive fashion in the solid state. Because the lack of crystal growth could be the consequence of either the nature of the protein construct or its level of purity, it is desirable to make the effort to obtain highly purified and homogeneous protein to use in crystallization trials.

Initial protein crystallization trial experiments routinely involve hanging-drop and sitting-drop vapor-diffusion methods. The vapor-diffusion technique uses the evaporation and diffusion of volatile species (including water) between solutions of different concentrations as a means of achieving supersaturation. The sitting-drop vapor-diffusion experiment has advantages over the hanging- drop experiments, for the purposes of robotic image capture

and analysis, because all crystallization drops are in a precision molded plasticware device and are, therefore, all in the same focal plane. One other method for crystallizing proteins is known as the microbatch method. A major distinction between the microbatch method and vapor-diffusion methods is that the microbatch trial is typically initiated with crystallization solutions with higher concentrations of precipitating agent, and tends to promote crystallization or precipitation on shorter time scales than vapor-diffusion methods, which are set up to allow a slow equilibration. Vapor diffusion drops eventually reach equilibrium and the microbatch drops are allowed to continually lose vapor (slowly or quickly, depending on the nature of the oil overlay). Both techniques are amenable to high-speed automation.

Improving crystal quality by site-directed mutagenesis:

It is quite typical of crystallographic investigations to focus on the first crystal form of reasonable quality obtained from screens. For soluble proteins, the norm is to use crystals diffracting to better than ≈ 3.0Å. If the only available crystals diffract poorly, an investigator can pursue one of two possibilities: either attempt to improve the diffraction quality of the existing crystals or—if this fails—to make modifications to the protein. The former can frequently be accomplished by special treatment of crystals, such as thermal annealing (Kriminski *et al*., 2002) or dehydration (Esnouf *et al*., 1998; Tahirov *et al*., 1998). Modifying the protein is very likely to change its physicochemical properties in a way sufficiently significant to alter the crystallization process and to warrant re-initiating the screens from scratch. This has both drawbacks and benefits: the drawback is that the investigator should no longer expect the crystallization observed for the previous samples to be reproducible. On the other hand, new crystal forms may appear with superior diffraction qualities. The modifications are essentially the same as those typically considered if a protein does not crystallize at all, e.g., truncations, changing the construct boundaries or surface mutagenesis. As indicated above, the rational surface mutagenesis strategy can yield crystals of superior quality as compared to wild-type protein (Tahirov *et al*., 1998). This can be particularly important when an accurate structure of a potential drug target is sought.

MATERIALS AND METHODS

Materials

Crystallization screening kits (Crystal Screen 1 and Crystal Screen 2) were obtained from Hampton Research, USA. PEG was obtained from Sigma Chemical Co. USA. All buffer salts were from standard commercial sources and of highest quality available. All other chemicals used in the present work were of analytical grade.

Methods

Crystallization

The enzyme TSC was purified from source as described in Chapter II. Initial crystallization conditions were screened by the hanging drop vapour diffusion method using the commercial crystallization screening kits (Crystal screen 1 and Crystal screen 2) from Hampton Research, USA. Protein concentration of 20 mg/ml and a drop size of 4 µl (1:1 ratio of protein and reservoir solution) were used for initial screening. About 10 mg pure protein was obtained from 4 l culture filtrate. The native enzyme was used for all crystallization setups. A protein concentration of 20 mg ml^{-1} (in triple-distilled water), a reservoir volume of 300 µl and a drop size of 4 µl (1:1 ratio of protein and reservoir solution) were used for initial screening. Long irregular rod shaped crystals appeared in a week in the following crystallization conditions from Crystal Screen 1 (Figs.2-7): condition no. 6 (0.2 M MgCl2, 0.1 M Tris–HCl pH 8.5, 30% PEG 4000), condition No. 18 (0.2 M magnesium acetate, 0.1 M sodium cacodylate pH 6.5, 20% PEG 8000), condition No. 28 (0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG 8000) and condition No. 46 (0.2 M calcium acetate, 0.1 M sodium cacodylate pH 6.5, 18% PEG 8000). The crystal obtained in condition No. 28 was found to be of good quality and size to collect the x-ray diffraction data compared with the crystals grown under the other conditions, which were not suitable for diffraction experiments. The crystal grown in the condition 28 (0.2 M Na (Ca or Mg) Acetate, 0.1 M Na cacodylate pH 6.5 and 20% PEG 8000) was used for the data collection.

Data collection

A single crystal of approximately 1.0 mm x 0.5 mm x 0.2 mm in size was used for full diffraction data collection. The diffraction intensity data were measured at 100 K using an imaging-plate detector mounted on a rotating-anode X-ray generator (Cu K_α radiation; λ = 1.5418 Å). Crystal freezing included a 3 minutes soaking of the crystal in a cryoprotectant solution containing the original crystallization solution and 20% Glycerol. The pre-soaked crystal was then submitted to immediate flash-freezing directly within a cold nitrogen-gas stream (Oxford Cryosystems). TSC crystal diffracted beyond 2.3 Å and data were processed using the programs *DENZO* and *SCALEPACK* from the *HKL2000* package (Otwinowski and Minor, 1997).

Amino acid sequencing of TSC

The amino acid sequencing of the peptides obtained by tryptic digestion was carried out by automated Edman degradation using a Procise ^{TM} protein sequencer (Applied Biosystems). The sample (500 pmol) was applied to a glass filter and then cycled through a conditioning process of repetitions of Edman degradation.

Multiple sequence alignment of TSC

The peptides of 15 or more amino acid residues and the N-terminal sequence were used to check the homology with other cellulases using BLAST program at NCBI. These peptides were used for sequence alignment with the members of conserved domain family of cellulose binding domains of different enzymes and also with cellulases having high homology. The multiple sequence alignments were constructed using CLUSTAL W software.

RESULTS AND DISCUSSION

Crystallization and data collection

TSC which depolymerizes cellulose has been crystallized by hanging drop vapour diffusion method. It has a molecular weight of ~ 14 kDa. The diffraction quality crystals of TSC were obtained in various crystallization conditions. The crystal grown in a reservoir solution of 0.2 M Na Acetate, 0.1 M Na Cacodylate pH 6.5 and 20% PEG 8000 was used for the diffraction data collection. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit cell parameters a = 49.9 Å, b = 79.5 Å, c = 99.7 Å. Similar crystals were grown with Mg / Ca acetate in place of Na acetate salt. Statistics for data collection and processing to 2.3 Å resolution are summarized in Table 2. The Matthews coefficient V_M (Matthews, 1968) was calculated to be 2.7 $Å^3$ Da^{-1} suggesting three molecules in the crystal asymmetric unit. This V_M value corresponds to a solvent content of approximately 55%. The crystal structure solution is under progress. The structure solution is not feasible by the molecular replacement method as the primary structure of the TSC is not known. Hence, attempts are being made for growing better quality crystals to get an atomic resolution data set which would help in solving the structure by *ab initio* method. We are also exploring the possibility of heavy atom derivatives to determine the structure of TSC enzyme using the method of multiple isomorphous replacement (MIR).

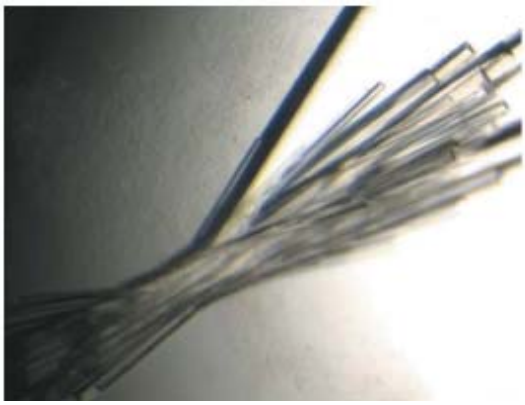


Fig. 2 Crystals of native TSC grown in crystallization condition no. 6 (0.2 M MgCl2, 0.1 M Tris–HCl pH 8.5, 30% PEG 4000) of Crystal screen 1 from Hampton Research.

Fig.3 Crystals of native TSC grown in crystallization condition no. 18 (0.2 M magnesium acetate, 0.1 M sodium cacodylate pH 6.5, 20% PEG 8000) of Crystal screen 1 from Hampton Research.

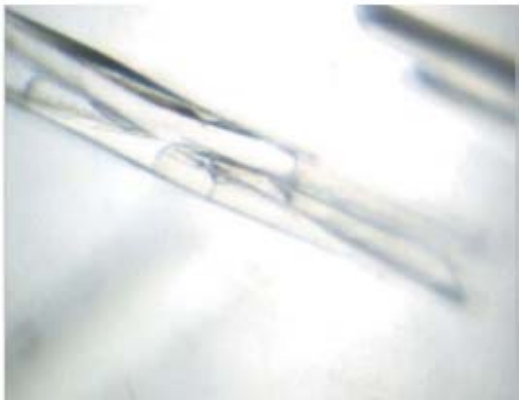


Fig. 4 Crystals of native TSC grown in crystallization condition no. 18 (0.2 M magnesium acetate, 0.1 M sodium cacodylate pH 6.5, 20% PEG 8000) of Crystal screen 1 from Hampton Research.

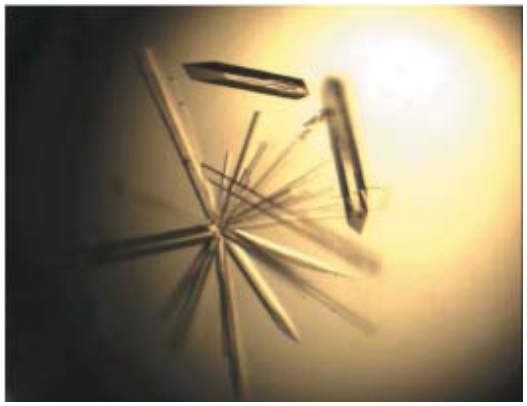


Fig. 5 Crystals of native TSC grown in crystallization condition no. 28 (0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG 8000) of Crystal screen 1 from Hampton Research.

Fig. 6 A single crystal of native TSC grown in crystallization condition no. 28 (0.2 M sodium acetate, 0.1 M sodium cacodylate pH 6.5, 30% PEG 8000) of Crystal screen 1 from Hampton Research.

Fig. 7 Crystals of native TSC grown in crystallization condition no. 46 (0.2 M calcium acetate, 0.1 M sodium cacodylate pH 6.5, 18% PEG 8000) of Crystal screen 1 from Hampton Research.

The crystal used for data collection (Fig. 6), with a maximum dimension of 1 mm, was about ten times the linear dimensions of the crystals in Figs. 2, 3, 4, 5 and 7

Table-2

Crystallographic data collection statistics

Values in parentheses refer to the highest resolution shell.

Wavelength (Å)	1.5418
Temperature (K)	100
Space Group	$P2_12_12_1$
Cell parameters (Å, °)	a = 49.9, b = 79.5, c = 99.7, $\alpha = \beta = \gamma = 90$
Unit-cell volume ($Å^3$)	391552
Matthews coefficient ($Å^3\ Da^{-1}$)	2.7
Solvent content (%)	55
Number of molecules in AU	3
Resolution (Å)	20.00 – 2.30 (2.38 – 2.30)
Observed reflections	146987
Unique reflections	18357 (1806)
Redundancy	8.0
Completeness (%)	100.0 (100.0)
Rmerge* (%)	10.6 (43.9)
$\langle I/\sigma(I)\rangle$	18.1 (4.9)

* Rmerge = $\Sigma \mid I - \langle I \rangle \mid / \Sigma I$

Amino acid sequencing and homology studies of TSC

Seven peptides were obtained after tryptic digestion of TSC and sequenced using Procise ™ protein sequencer (Applied Biosystems). The sequences of peptides obtained by tryptic digestion are as follows:

N-terminal sequence of TSC is:

GSVINPATRNY

The sequences other six peptides:

NYGCWLR

SMDAVGPWK

HVVFTIWK

ASHMDQTYYLCSDVNFV

LPDGQLCSGGLAEGGR

PLTWGSLDLVHR

The N-terminal sequence **GSVINPATRNY ASHMDQTYYLCSDVNFV, LPDGQLCSGGLAEGGR** and were used for homology and multiple sequence alignment studies.

Sequence homology of peptides of TSC

Sequence homology of **GSVINPATRNY:**
BLAST search showed 93 % homology of the peptide with putative secreted cellulose-binding protein of *Thermobifida fusca* YX (Fig. 8)

Identities = 11/11 (100%), Positives = 11/11 (100%), Gaps = 0/11 (0%)

1 **GSVINPATRNY** 11 peptide

GSVINPATRNY ⟶ **100 % homology**

38 **GSVINPATRNY** 48 *Thermobifida fusca* YX

The peptide has 72 % homology with secreted cellulose-binding protein [*Streptomyces coelicolor* A3 (2)

Identities = 8/11 (72%), Positives = 11/11 (100%), Gaps = 0/11 (0%)

1 **GSVINPATRNY** 11 peptide

GSV++PA+RNY ⟶ **72 % homology**

44 **GSVVDPASRNY** 54 *Streptomyces coelicolor*

Sequence homology of **LPDGQLCSGGLAEGGR**

BLAST search showed 93 % homology of the peptide with putative secreted cellulose-binding protein of *Thermobifida fusca* YX

Identities = 15/16 (93%), Positives = 15/16 (93%), Gaps = 0/16 (0%)

1 **LPDGQLCSGGLAEGGR** 16 peptide

LPDGQLCSGGL EGGR ⟶ **93 % homology**

100 **LPDGQLCSGGLTEGGR** 115 *Thermobifida fusca* YX

The peptide has 80 % homology with secreted cellulose-binding protein [*Streptomyces coelicolor* A3 (2)

Identities = 12/15 (80%), Positives = 12/15 (80%), Gaps = 0/15 (0%)

2	**PDGQLCSGGLAEGGR**	16	peptide
	PDGQLCSGG E GR	⟶	**80 % homology**
107	**PDGQLCSGGRTESGR**	121	*Streptomyces coelicolor*

75% homology was observed with secreted cellulose-binding protein *Streptomyces avermitilis* MA-4680

Identities = 9/10 (90%), Positives = 10/10 (100%), Gaps = 0/10 (0%)

2	**PDGQLCSGGL**	11	peptide
	PDG+LCSGGL	⟶	**90 % homology**
88	**PDGKLCSGGL**	97	*Streptomyces avermitilis*

Sequence homology of **ASHMDQTYYLCSDVNFV**

BLAST search showed 100 % homology of the peptide with putative secreted cellulose-binding protein of *Thermobifida fusca* YX

Identities = 17/17 (100%), Positives = 17/17 (100%), Gaps = 0/17 (0%)

1	**ASHMDQTYYLCSDVNFV**	17	peptide
	ASHMDQTYYLCSDVNFV	⟶	**100 % homology**
206	**ASHMDQTYYLCSDVNFV**	222	*Thermobifida fusca* YX

The peptide has 87 % homology with secreted cellulose-binding protein [*Streptomyces coelicolor* A3 (2)

Identities = 14/16 (87%), Positives = 16/16 (100%), Gaps = 0/16 (0%)

1 **ASHMDQTYYLCSDVNF** 16 peptide

ASHMDQTY+LCSDV+F ⟶ 87 % homology

212 **ASHMDQTYFLCSDVDF** 227 *Streptomyces coelicolor*

75% homology was observed with putative endoxyloglucan transferase [*Oryza sativa* (japonica cultivar-group)

Identities = 6/8 (75%), Positives = 8/8 (100%), Gaps = 0/8 (0%)

8 **YYLCSDVN** 15 peptide

YY+CSDV+ ⟶ 75 % homology

103 **YYMCSDVD** 110 endoxyloglucan transferase

Multiple Sequence alignment of TSC

Trypsin- digestion of TSC led to the recovery by reverse – phase HPLC of a series of six peptide fragments. The C- terminal sequence ASHMDQTYYLCSDVNFV showed high homology to cellulose binding domain of *Thermobifida fusca* and *Streptomyces coelicolor* with identity of 100% and 87% respectively. Multiple sequence alignment of the C- terminal domain of the enzyme showed the presence of conserved cysteine, phenylalanine and serine residues.

The peptide LPDGQLCSGGLAEGGR showed high homology to cellulose binding domain of *Thermobifida fusca* and *Streptomyces coelicolor* with identity of 93% and 80% respectively. The multiple sequence alignment of LPDGQLCSGGLAEGGR showed the presence of conserved glycine.

The N-terminal sequence GSVINPATRNY showed high homology to cellulose binding domain of *Thermobifida fusca* and *Streptomyces coelicolor* with identity of 100% and 72% respectively. The multiple sequence alignment of GSVINPATRNY showed the presence of conserved glycine, alanine, arginine, tryptophan and asparagine.

(Please find more information about sequences aligned in Appendix.)

An alignment will display by default the following symbols denoting the degree of conservation observed in each column:

"*" means that the residues or nucleotides in that column are identical in all sequences in the alignment.

":" means that conserved substitutions have been observed, according to the COLOUR table above.

"." means that semi-conserved substitutions are observed.

```
gi|72161672     MHRYSRT-----GKHR-WTVRALAVLFTALLGLTQWTAPASAHGSVINPATRNYGCWLRW 54
gi|71915404     MHRYSRT-----GKHR-WTVRALAVLFTALLGLTQWTAPASAHGSVINPATRNYGCWLRW 54
peptide         -------------------------------------------GSVINPATRNY------ 11
gi|6434713      MTCHDRAKIQLAGRARRATTLVLSTLAAVLLTLIPWSGTAVAHGSVVDPASRNYGCWERW 60
gi|34498778     ----------------------MKPHFSIVMMLAGLSTSSWAHGTMEVPINRTYSCSKEG 38
gi|50937973     ----------MGSLGRRPWVGGLTAAMIFAVAVCGFCFSGASAAAAAPTFGDNFEITGAE 50
gi|50937975     ----------MGSLGRRPWVGGLTAAMIFAVAVCGFCFSGASAAAAAPTFGDNFEITGAE 50
                                                          .:   .   .:
```

Fig. 8 Sequence alignment of GSVINPATRNY to that of other cellulases with high homology using CLUSTAL W software. Residues identical in all sequences are indicated below the alignment by asterisk.

Multiple sequence alignment of **GSVINPATRNY** to that of other cellulases with high homology showed the presence of conserved Glycine, Serine, Proline, Asparagine, and Tryptophan.

```
gi|46015849    MGYAYDFTG-------DSKYLDGMFDGISYLLGRNAMDQSYVTGYGERPLQNPHDRFWTP 536
gi|6272570     MGYAYDFTKGDADKKKSNSYFNGLTSAMDYLLGRNPMVQSYVTGYGSNPLENPHHRFWAY 749
gi|37703323    LAYAYDVDHS------GTKYLNGAAEALDYIYGRNGLGFSYVSGYGDKTMEWPHHRFWAQ 828
gi|13423734    LGLAHDLTG-------EIAYRYGAVDAMDYVLGRNPLDQSYVSGFGARPMRHPHHRFWAK 547
gi|6468241     LATAYDITG-------GSKYRDGAVQSMDYILGRNALNMSYVTGYGEVNAHNQHSRWYAH 676
gi|48837427    LAVAFDLTG-------DTKYRDGVLEGMDYIFGRNALNQSYVTGYGDKDSRNQHSRWYAH 679
gi|67987581    LGTAFDLTG-------DDAYRDGVVRGMDYLLGRNALNNSYITGYGDVFSKNQHSRMYAN 702
gi|4104166     LALAYDFTG-------QTKYRTAVYGALDYLQGRNPLNQSYITGYGENPVRNVHHRFWAH 532
gi|48861645    LGLAYDFTG-------DDVYAKTVSKAVNYLFGNNTLSFSYISGHGENALQQPHHRFWAG 527
peptide        ----------------------------------------G--SVINPATRNY-- 11
gi|1064979     LQVANKISP-------NNDYVNAALDAISHVFGRNYYNRSYVTGLGINPPMNPHDRRSGA 512
                                                        *         *
```

Fig. 9 Multiple sequence alignment of GSVINPATRNY to that of members of cd02850 (Cellulase_N_term: Cellulase N-terminus domain.) of Conserved Domain Database. Residues identical in all sequences are indicated below the alignment by asterisk.

The N-terminus of cellulase may be related to the immunoglobulin and/or fibronectin type III superfamilies. These domains are associated with different types of catalytic domains at either the N-terminal or C-terminal end and may be involved in homodimeric/ tetrameric /dodecameric interactions. Members of this family include members of the alpha amylase family, sialidase, galactose oxidase, cellulase, cellulose, hyaluronate lyase, chitobiase, and chitinase.

Multiple sequence alignment of **GSVINPATRNY** to that of members of cd02850 (Cellulase_N_term: Cellulase N-terminus domain.) of Conserved Domain Database showed the presence of conserved Glycine and Arginine. These two residues are identical in all sequences aligned as indicated by asterisk.

```
gi|1200196     RVFKPFDATY---ATQCLEAAKKAYAWALQNP-KVAFY--NPMDVATGSYSDGEFSDEFA 337
gi|606791      RIYKHASTPY---ASKCLEAAKKAYAWGQQNPSRNYLA--NPSDVSTGAYENDNPNDEKV 338
gi|6572464     RVYAPFDKQYPGASARMLKAARSAWAWAQQHP-NVIYR--QPDDVRTGGYDDATVDDEFA 362
gi|4104166     PVG-AVDAGF---AAGPLSAARTAYAAAKANPNRIAS---DADGTGGG-YGDPSVTDEFY 372
gi|6468241     RLYRPYDREF---AAKALDAARTAWTAALAHPDLLAD---ESDGIGGGAYPDNEVADEFY 517
gi|530014      RLWKDYDPTF---AADCLEKAEIAWQAALKHPDIYAEYTPGSGGPGGGPYNDDYVGDEFY 565
gi|121806      RIWKDIDAGF---AALCLTAAERAWNAAQANPNDIYS---GNYDNGGGGYGDRFVADEFY 376
gi|1769559     TVFKNIDKEY---AEKLLKAAKFGYQYLEKNSNLIIS-------MPNEPYYDGNDTDNRL 442
gi|1663519     LIYKDIDPDF---AQECLDAAINAWKFLEKNPENIVS-------PPG-PYNVYDDSGDRL 530
peptide        -------------GSVINPATRNY---------------------------------- 11
gi|1064979     RIFRPYDPQY---AEKCINAAKVSYEFLKNNPANVFAN---QSGFSTGEYATVSDADDRL 354
                              :  *   :
```

Fig. 10 Multiple sequence alignment of GSVINPATRNY to that of members of pfam02927.12p3 (CelD_N: N-terminal ig-like domain of cellulase, Putative carbohydrate binding domain) of Conserved Domain Database. Residues identical in all sequences are indicated below the alignment by asterisk.

Alanine is the identical residue in all sequences in the multiple sequence alignment of **GSVINPATRNY** to that of members of **pfam02927.12p3 (**CelD_N: N-terminal ig-like

domain of cellulase, Putative carbohydrate binding domain) and Isoleucine and Tryptophan are two other conserved residues.

```
gi|1346278      T------------------------------------------------------------ 773
gi|1170252      PKLDKAGITYRLPVLGAVIKNNILDVVTEFHGVAIQYSLDG-KTWHKYDDTKKPQVSTKA 827
gi|10835638     AKLDKGGVAYRLPVPGARVAGGKLEANIALPGLGIEYSTDGGKQWQRYDAKAKPAVSGEV 840
gi|3978256      AKLAKSGVKANLDVPGGEVNKGTLTTSVATAYLDVEYSVDQGKSWQTYQQPVNVGDTQVW 843
peptide         ------------LPDGQLCSG---------GLA--------------EG---------- 14
gi|5881581      QIPQQVGVNYRIPLPGGAIEGGKLIVNTRFVGMATEYSVDQGVTWHRYEGPVTLTASSKV 896


gi|1346278        --------------------
gi|1170252        LVRSVSTNGKTGRAVEVLAK 847
gi|10835638       QVRSVSPDGKRYSRAEKV-- 858
gi|3978256        LRSSLNDQVKSRITTIN--- 860
peptide           --------GK---------- 16
gi|5881581        QLRTVSDSGKVSRVATVN-- 914
```

Fig. 11 Multiple sequence alignment of LPDGQLCSGGLAEGGR to that of members of CHB_HEX (pfam 0373.11 Putative carbohydrate binding domain) of Conserved Domain Database.

Multiple sequence alignment of **LPDGQLCSGGLAEGGR** to that of members of CHB_HEX (pfam 0373.11 Putative carbohydrate binding domain) showed the presence of conserved Glycine, Proline and Leucine residues.

```
peptide         ----------------------------------------------LP------------ 2
gi|4099045      NHGITPTFDDQNPKVDLALERLIANNNGFKAGDKVQFTTFLRNSGDLPIGKDSKVVVKFY 418
gi|3090434      SNGIKATS------------------------------------------------------ 325
gi|2148986      DCGLQNGSASEESG------------------------------------------------ 350
gi|121845       G------------------------------------------------------------ 301


peptide         -DG-QL-----CSGGLAEGGR--------------------------------------- 16
gi|4099045      VDGVQLGDPVEISGGIAVGQRIPVTSPELTVSKTDFT---VKAVIDSSSTYADGAEDAVI 475
gi|3090434      -----------VLSGIFGG-VTPTSSPTSTPTSTPTS---TPTPTPSPTPSPGNNGTILY 370
gi|2148986      -----------FSTGVANGATTSILDDMEDGDRYGYLGGAWAAVEDQENGGASSISNEKI 399
gi|121845       ------------PSGVRTVRAAPIQPSWVAWN-------LGAQLLTKPWVVRSVWLPVIQ 342
                           *:
```

Fig. 12 Multiple sequence alignment of LPDGQLCSGGLAEGGR to that of members of pfam00150 Cellulase: Cellulase (glycosyl hydrolase family 5) of Conserved Domain Database. Residues identical in all sequences are indicated below the alignment by asterisk.

Glycine was identical residue in all sequences aligned in the multiple sequence alignment of **LPDGQLCSGGLAEGGR** to that of members of pfam00150 Cellulase: Cellulase (glycosyl hydrolase family 5)

```
gi|12225048     FADGKGIATQSGGANLEKVGNKVFEGNDDGAALYVSNRINNWDAADFKFEDIG------- 91
gi|5360744      FAEGKGAAVQSGGATITHVTGKFFDGNGDGAALYISNRVNNWDAADFRFSDIG------- 90
gi|27227837     FQNGKGKATQSGGAQLEAVTGKSFEGNDDGAALKVSNRSNNYDGVDIAFTDLG------- 98
peptide         LPDGQ---LCSGG---------LAEG---GR----------------------------- 16
gi|2760908      DFENGTTQGWSPRGNSTTVETVYRIAYEGDYSLKVSGRSAAWDGAIVDVTSSVSVNTMYT 105
gi|3810965      ------------------------------------------------------------
gi|3201481      SFEEGTTQGWTARGGVK-VDVTAEQAYQGKQSLQTTGRTEAWNGPSLSLTDVVHKNEVVE 100
gi|974180       DED-------------------------------------------------------- 44
gi|1103639      DVSALTFDDGPNGATTTALLDFLAERDIRATFCVIGQNIQAPGGADVLRRIVADG----- 107
gi|5880612      VVSAVDFEDGTTGTWTQSGSPTLAVVES----------PDGADDGQVLSITRAAD----- 105
```

Fig. 13 Multiple sequence alignment of LPDGQLCSGGLAEGGR to that of members of Domain of unknown function (DUF1083) of xylanases.

Multiple sequence alignment of **LPDGQLCSGGLAEGGR** to that of members of Domain of unknown function (DUF1083) of xylanases showed the presence of conserved Glutamate and Glycine residues.

```
gi|72161672     QPLTWDSLELVHQTGSYPPA----QNIQFTVHAPN-RSGRHVVFTIWKASHMDQTYYLCS 217
gi|71915404     QPLTWDSLELVHQTGSYPPA----QNIQFTVHAPN-RSGRHVVFTIWKASHMDQTYYLCS 217
peptide         ----------------------------------------------ASHMDQTYYLCS 12
gi|6434713      QALTWGELQQVARTGSYGPS----QNYEIPVSTSG-LTGRHVVYTIWQASHMDQTYFLCS 223
gi|34498778     QPLKWSDLEPFGTYNGNPPLDDK-QRYHMTMKLPTGKTGRHIIYNVWKRSDSEEAFYSCS 200
gi|50937973     IVFFVDKVPIREYRNSDKPNTFFPIMKPMYVFSSIWNADDWATRGGLEKTDWTKAPFISS 230
                                                          :.  :: : .*

gi|72161672     DVNFV---------------------------------------------------- 222
gi|71915404     DVNFV---------------------------------------------------- 222
peptide         DVNFV---------------------------------------------------- 17
gi|6434713      DVDFG---------------------------------------------------- 228
gi|34498778     DVNFTNDGKPEPPPISNPWKEAGSVTAHENLPDKSSVTLRIFDSHGRDVESHKVDLSASS 260
gi|50937973     YRDFTADACSWGTAAASPPSCAASTG------------------------------NSWW 260
                  :*
```

Fig. 14 Sequence alignment of ASHMDQTYYLCSDVNFV to that of other cellulases with high homology using CLUSTAL W software. Residues identical in all sequences are indicated below the alignment by asterisk.

When the peptide **ASHMDQTYYLCSDVNFV** was aligned with other cellulases with high homology, Serine and Phenylalanine were the identical residues in all sequences while Glutamine, Threonine and Cysteine were among the conserved residues.

```
gi|1346278      RAAIIAMEARYKKLMAQGKPEDAQKYRLVETADKTRYSSIQHYNDN-TLNVCIANTYTFI 504
gi|1170252      LAAVKSMEARYRKFMAEGDVVKAEMYLLSDPNDTTQYYSIQHYQDN-TINPCMESSFVFM 492
gi|10835638     RAAVVSMEARYKKLHAAGKEQEANEFRLVDQTDTSNTTSVQFFNRQSYLNPCLDSSQRFV 486
gi|3978256      RASIKAMEARYRKLMKAGDEVGAKAYLLSDPEDKSKYITVQNYTDN-SINACMDSSYNFM 516
gi|5881581      RAAIKAMEHRYNKYKDS-DPVEANRFRLIDPLDQSQYYTPQFYTDN-FVNPALESTFTFL 566
peptide         ------------------------------------ASHMDQTYY-------LCSDVN--FV 17
                                                     :      :        .  .  *:
```

Fig. 15 Multiple sequence alignment of ASHMDQTYYLCSDVNFV to that of members of CHB_HEX (pfam 0373.11 Putative carbohydrate binding domain) of Conserved Domain Database. Residues identical in all sequences are indicated below the alignment by asterisk.

Multiple sequence alignment of **ASHMDQTYYLCSDVNFV** to that of members of CHB_HEX (pfam 0373.11 Putative carbohydrate binding domain) of Conserved Domain Database showed the presence of identical Phenylalanine residue while Alanine, Tryptophan, Cysteine, Asparagine and Valine were the conserved residues

```
gi|2760908      DTINLFSTIPGLQIQITELDISVYTSSTQQYDTLPQDIMIKQALKFKELFEMLKRHSDR- 811
gi|3810965      EAFEKVIALG-VNVQVTELDVSIYSSNSELEMPVTDELMLQQAYRYRELFDLFDDLGKRG 506
gi|3201481      EAFNRAAALG-VHIQVTELDMSIYSGNSEQEKPVTDEMMLEQAYRYRALFDLFKEFDDRG 662
gi|12225048     LSLERFISLG-VEVSITELDVQAGSD-----FKLTDEIAEAQGYLYAQLFDIYKKHAAN- 652
gi|5360744      LSLKRFTELG-VEVSISELDIRAGSN-----YQLTEKEANAQAYLYAQLFKIFREYSDS- 650
gi|27227837     MSLERFISLG-VEVSVTELDIMAGTN-----SAITEKEAKQQAYLYASLMDLYKKNSEH- 657
gi|974180       EAIQDYSQVV-DEVMITEFDVKAGLGYDGTDETKDKEYTKQAHYFRKIYEAAKALNAEGV 576
gi|1103639      ASIDAFATTG-LLQAVTELDAPIDGT-------VSQEKLVAQGYYYADVFDMLR--QYP- 691
gi|5880612      AALTRFAALP-VVQAVTELDNTVGTP-------VTEANLIKQGHWYQDAFNVFR--SHAD 681
peptide         ------------------------------------ASHMDQTYYLCSDVNFV------- 17
```

Fig.16 Multiple sequence alignment of ASHMDQTYYLCSDVNFV to that of members of Domain of unknown function (DUF1083). This family consists of several domains of unknown function exclusively found in bacterial xylanase proteins (usually at the C-terminus) although it is tandemly repeated in a number of family members. This family is always found in conjunction with pfam00331 and usually with either pfam02018 or pfam00395. The function of this family is unknown.

Multiple sequence alignment of **ASHMDQTYYLCSDVNFV** to that of members of Domain of unknown function (DUF1083) showed the presence of conserved Glutamine and Tryptophan.

Trypsin cleaves the peptide bond at C- terminal of basic Arginine and Lysine. As there is neither Arginine nor Lysine at C- terminal of **ASHMDQTYYLCSDVNFV**, it was predicted to be C- terminal of TSC. It was confirmed by 100% homology with C- terminal of cellulose binding domain of other enzymes. Tryptophan and carboxylic acids have been shown to be involved in the catalytic mechanism of xylanases, cellulases and lysozyme. GH clan C consists of 4 major group of sequences of glycoside hydrolases. Conserved cysteines present in GH clan C do not play any role in the catalytic activity of the enzyme however they are needed for a local stabilization of GH-C enzyme structures (Kleywegt et al, 1997). Conserved cysteine residues can be found in family 12 glycosyl hydrolases. Cysteine is one of the conserved residues in the active site sequences of all family 7 cellulases (Sandgren et al, 2005). However, information about the involvement of lysine residues in the active site of cellulases is not available. As a step towards the understanding of the structure–function relationships of TSC from *Thermomonspora* sp., the primary objective of the present work was to decipher the structural environment of the active site residues and compare it with that of other cellulases. Our studies demonstrated the presence of essential lysine and cysteine residue in the active site of TSC by fluorescent chemoaffinity labeling and group specific modifications. This is the first report of presence of lysine in the active site of cellulase. The multiple sequence alignment of the C - terminal peptide showed the presence of conserved cysteine residue corroborating the chemical modification studies suggesting its involvement in the substrate binding of the enzyme. Chemical modification data confirms the findings.

SUMMARY

Crystallization and preliminary X- ray characterization

The 1,4-β-D-Glucan Glucanohydrolase was crystallized by hanging – drop vapour-diffusion method using commercial crystallization kits. Good quality crystals, suitable for crystallization were obtained using 0.2M sodium acetate, 0.1M sodium cacodylate pH 6.5, 18% PEG 8000.

The crystals were diffracted at 2.3 Å resolution. The crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit cell parameters a = 49.9, b = 79.5,

c = 99.7Å. The Matthew's coefficient VM was calculated to be 2.7 $Å^3Da^{-1}$

Amino acid sequence analysis

The research interest of this part of the thesis was to sequence the purified 1,4 – β – D – Glucan Glucanohydrolase.The amino acid sequence analysis of the peptides obtained after tryptic digestion was carried out. The N- terminal sequence is **GSVINPATRNY.** The C-terminal sequence is **ASHMDQTYYLCSDVNFV**

The other five peptides sequenced are **NYGCWLR, SMDAVGPWK, HVVFTIWK, LPDGQLCSGGLAEGGR, PLTWGSLDLVHR**

Of these **ASHMDQTYYLCSDVNFV** and **LPDGQLCSGGLAEGGR** showed sequence homology with the cellulase binding domain of other enzymes from *Themobifida fusca* and *Streptomyces coelicolor*. The conserved residues found in the multiple sequence alignment of these peptides, were cysteine, serine, phenylalanine, and glycine.

REFERENCES

Alzari, Pedro M, Hélè ne Souchon, and Roberto Dominguez 1996 The crystal structure of endoglucanase CelA, a family 8 glycosyl hydrolase from *Clostridium thermocellum* Structure 4: 265-275.

Bedarkar, S., Gilkes, N.R., Kilburn, D.G., Kwan, E., Rose, D.R., Miller, R.C. Jr., Warren, R.A.J. and Withers, S.G. (1992) Crystallization and preliminary X-ray diffraction analysis of the catalytic domain of Cex, an exo-1,4- 1,4-glucanase and/β-1,4-xylanase from the bacterium *Cellulomonas fimi.* J. Mol. Biol. 228, 693 - 695.

Biol. 228, 970-972.

Boraston, A. B., Creagh, A. L., Alam, M. M., Kormos, J. M., Tomme, P., Haynes, C. A. et al. (2001). Binding specificity and thermodynamics of a family 9 carbohydrate-binding module from *Thermotoga maritima* xylanase 10A. Biochemistry, 40, 6240–6247.

Bourne Y., Henrissat B. (2001) Glycoside hydrolases and glycosyltransferases: families and functional modules. Curr. Opin. Struct. Biol. 11, 593-600

Coughlan, M. (1985). Cellulases: production, properties and applications. *Biochem. Soc. Trans.* 13, 405 – 406

Crennell, Susan J., Gudmundur O. Hreggvidsson and Eva Nordberg Karlsson The Structure of *Rhodothermus marinus* Cel12A, A Highly Thermostable Family 12 Endoglucanase,at 1.8 Å Resolution 2002 J. Mol. Biol. 320, 883-897

Cutfield, Sue M.,Gideon J. Davies, Garib Murshudov, Bryan F. Anderson, Peter C. E. Moody, Patrick A. Sullivan and John F. Cutfield The structure of the exo-β-(1,3)-glucanase from *Candida albicans* in native and bound forms: relationship between a pocket and groove in family 5 glycosyl hydrolases J. Mol. Biol.294, 1999, 771-783

David J. Scott, Wendy A. Offen, Shirley Roberts, Gavin Pell, Harry J. Gilbert, Gideon J. Davies and James A. Brannigan 2003 Multifunctional xylooligosaccharide/Cephalosporin C Deacetylase Revealed by the Hexameric Structure of the *Bacillus subtilis* Enzyme at 1.9 Å Resolution J. Mol. Biol. 330, 593-606

Davies, G. J., Dauter, M.; Brzozowski, A. M.; Bjornvad, M. E.; Andersen, K. V.; Schulein, M. 1998 Structure of the *Bacillus agaradherans* Family 5 Endoglucanase at 1.6 Å and Its Cellobiose Complex at 2.0 Å Resolution Biochemistry; 37; 1926-1932.

Davies, G., Tolley, S., Wilson, K., Schiilein, M., Wodlike, H.F. and Dodson, G. (1992) Crystallization and preliminary X-ray analysis of a fungal endoglucanase J. Mol.

Davies,G.J., G. Dodson, M. H. Moore, S. P. Tolley, Z. Dauter, K. S. Wilson, G. Rasmussen and M. Schülein 1996 Structure determination and refinement of the *Humicola insolens* endoglucanase V at 1.5 Å Resolution Acta Cryst. D52, 7-17

Davies,G.J., G.G. Dodson, R.E. Hubbard, S.P. Tolley, Z. Dauter, K.S. Wilson, C. Hjort, J.M. Mikkelsen, M. Schülein, Nature 365 (1993) 362-364.

Davies,G.J., S.P. Tolley, B. Henrissat, C. Hjort, M. Schülein, Structures of Oligosaccharide-Bound Forms of the Endoglucanase V from *Humicola. insolens* at 1.9Å resolution Biochemistry Biochemistry 34 (1995) 16210-16220.

Divne, C., Sinning, I., Stahlberg, J., Pettersson, G., Bailey, M., Siika-Aho, M., Margolles-Clark, E., Teeri, T. and Jones, A.T. (1993) Crystallization and preliminary X-ray studies on the core proteins of cellobiohydrolase I and endoglucanase 1 from *Trichoderma reesei.* J. Mol. Biol., 1993, 905-907

Divne, C.,J Ståhlberg, T Reinikainen, L Ruohonen, G Pettersson, JK Knowles, TT Teeri, and TA Jones 1994The three-dimensional crystal structure of the catalytic core of cellobiohydrolase I from *Trichoderma reesei* Science 265: 524-528.

Divne,Christina, Jerry Ståhlberg , Tuula T. Teeri and T. Alwyn Jones 1998 High-resolution crystal structures reveal how a cellulose chain is bound in the 50 Å long tunnel of cellobiohydrolase I from *Trichoderma reesei* J. Mol. Biol. 275, 309-325

Domínguez, Roberto, Hélène Souchon, Marie-Bernard Lascombe and Pedro M. Alzari

Esnouf R.M., Ren J., Garman E.F., Somers D.O., Ross C.K., Jones E.Y., D.K., Stammers, Stuart D.I., Continuous and Discontinuous Changes in the Unit Cell of HIV-1 Reverse Transcriptase Crystals on Dehydration *Acta Cryst.* (1998). D54, 938-953

Fujimoto, Zui,Satoshi Kaneko, Atsushi Kuno, Hideyuki Kobayashi, Isao Kusakabe, and Hiroshi Mizuno 2004 Crystal Structures of Decorated Xylooligosaccharides Bound to a Family 10 Xylanase from *Streptomyces olivaceoviridis* E-86 J. Biol. Chem., 279: 9606 – 9614

Hakamada Y, Endo K, Takizawa S, Kobayashi T, Shirai T, Yamane T, Ito S (2002) Enzymatic properties, crystallization, and deduced amino acid sequence of an alkaline endoglucanase from *Bacillus circulans*. Biochim Biophys Acta 1570:174–180

Henrissat, B. and Bairoch, A.(1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316, 695 – 696.

Henrissat, B. and Davies, G.(1997). Structural and sequence based classification of glycosyl hydrolases. *Curr. Opin. Struct. Biol.* 7, 637 – 644.

Hilge, Mark, Sergio M Gloor, Wojciech Rypniewski, Oliver Sauer, Tom D Heightman, Wolfgang Zimmermann, Kaspar Winterhalter, and Klaus Piontek High-resolution native and complex structures of thermostable β-mannanase from *Thermomonospora fusca* – substrate specificity in glycosyl hydrolase family 5 1998 Structure 6: 1433-1444

Hirvonen ,Mika and Anastassios C. Papageorgiou 2003 Crystal Structure of a Family 45 Endoglucanase from *Melanocarpus albomyces*: Mechanistic Implications Based on the Free and Cellobiose-bound Forms J. Mol. Biol. 329,403-410

Khademi, S., D. Zhang, S. M. Swanson, A. Wartenberg, K. Witte and E. F. Meyer Determination of the structure of an endoglucanase from *Aspergillus niger* and its mode of inhibition by palladium chloride *Acta Cryst.* (2002). D58, 660-667

Khademi, S.,L. A. Guarino, H. Watanabe, G. Tokuda and E. F. Meyer. Structure of an endoglucanase from termite, *Nasutitermes takasagoensis Acta Cryst.* (2002). D58, 653-659

Kleywegt, Gerard J., Jin-Yu Zou, Christina Divne, Gideon J. Davies, Irmgard Sinning, Jerry Ståhlberg, Tapani Reinikainen, Malee Srisodsuk, Tuula T. Teeri and T. Alwyn Jones 1997,The crystal structure of the catalytic core domain of endoglucanase I from *Trichoderma reesei* at 3.6 Å resolution, and a comparison with related enzymes J. Mol. Biol. 272, 383-397

Kriminski S., Caylor C.L., Nonato M.C., Finkelstein K.D., Thorne R.E., Flash-cooling and annealing of protein crystals *Acta Cryst.* (2002). D58, 459-471

Larsson, Anna M., Terese Bergfors, Elisa Dultz, Diana C. Irwin, Annette Roos, Hugues Driguez, David B. Wilson, and T. Alwyn Jones Crystal Structure of *Thermobifida fusca* Endoglucanase Cel6A in Complex with Substrate and Inhibitor: The Role of Tyrosine Y73 in Substrate Ring Distortion Biochemistry 2005

Mateja, A., Devedjiev,Y., Krowarsch, D., Longenecker, K., Dauter, Z., Otlewski, J. and Derewenda, Z.S. (2002).The impact of Glu-->Ala and Glu-->Asp mutations on the

crystallization properties of RhoGDI: the structure of RhoGDI at 1.3 A resolution. Acta Crystallogr. D Biol. Crystallogr., 58, 1983-1991

Matthews, B. W. (1968). Solvent Content of protein crystals. *J. Mol. Biol.* 33, 491-497.

McCarter, J. D. & Withers, S. G. (1994). Mechanisms of enzymatic glycoside hydrolysis. Curr. Opin. Struct. Biol. 4, 885-892

McLean, B. W., Bray, M. R., Boraston, A. B., Gilkes, N. R., Haynes, C. A. & Kilburn, D. G. (2000). Analysis of binding of the family 2a carbohydrate-binding module from *Cellulomonas fimi* xylanase 10A to cellulose: specificity and identification of functionally important amino acid residues. Protein Eng. 13, 801–809.

Muñoz, Inés G., Wimal Ubhayasekera, Hongbin Henriksson, Istvan Szabó, Göran Pettersson, Gunnar Johansson, Sherry L. Mowbray and Jerry Ståhlberg 2001 Family 7 cellobiohydrolases from *Phanerochaete chrysosporium*: crystal structure of the catalytic module of Cel7D (CBH58) at 1.32 Å resolution and homology models of the isozymes J. Mol. Biol. 314, 1097-1111

Nagy, T., Simpson, P., Williamson, M. P., Hazlewood, G. P., Gilbert, H. J. & Orosz, L. (1998). All three surface tryptophans in type IIa cellulose binding domains play a pivotal role in binding both soluble and insoluble ligands. FEBS Letters, 429, 312–316.

Najmudin, S., C. I. P. D. Guerreiro, L. M. A. Ferreira, M. J. C. Romão, C. M. G. A. Fontes and J. A. M. Prates. Overexpression, purification and crystallization of the two C-terminal domains of the bifunctional cellulase *ct*Cel9D-Cel44A from *Clostridium thermocellum. Acta Cryst.* (2005). F61, 1043-1045

Natesh, R.; Bhanumoorthy, P.; Vithayathil, P. J.; Sekar, K.; Ramakumar, S.; Viswamitra, M. A. Crystal structure at 1.8 Å resolution and proposed amino acid sequence of thermostable xylanase from *Thermoascus aurantiacus*. *J. Mol. Biol.* 1999, *288*, 999-1012.

Natesh,R., K. Manikandan, P. Bhanumoorthy, M. A. Viswamitra and S. Ramakumar Thermostable xylanase from *Thermoascus aurantiacus* at ultrahigh resolution (0.89 Å) at 100 K and atomic resolution (1.11 Å) at 293 K refined anisotropically to small-molecule accuracy Acta Cryst. (2003). D59, 105-117

Notenboom, V., Boraston, A. B.,Williams, S. J., Kilburn, D. G. & Rose, D. R. (2002). High-resolution crystal structures of the lectin-like xylan binding domain from *Streptomyces*

lividans xylanase 10A with bound substrates reveal a novel mode of xylan binding. Biochemistry, 41, 4246–4254.

Otwinowski, Z. and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymology,* 276**,** 307 - 326.

Pell,Gavin Lóránd Szabo, Simon J. Charnock, Hefang Xie, Tracey M. Gloster, Gideon J. Davies, and Harry J. Gilbert 2004 Structural and Biochemical Analysis of *Cellvibrio japonicus* Xylanase 10C: how variation in substrate-binding cleft influences the catalytic profile of family gh-10 xylanases J. Biol. Chem., 279: 11777 – 11788

Pickersgill, R.W., Jenkins, J.A., Scott, M., Connerton, Hazlewood, G.P. and Gilbert, H.J, (1993) Crystallization and preliminary X-ray analysis of the catalytic domain of xylanase A from *Pseudomonas fluorescens* var. *cellulosa.* J. Mol. Biol. 229, 246-248.

Pires, Virgínia M. R.,Joanna L. Henshaw, José A. M. Prates, David N. Bolam, Luís M. A. Ferreira, Carlos M. G. A. Fontes, Bernard Henrissat, Antoni Planas, Harry J. Gilbert, and Mirjam Czjzek 2004 The Crystal Structure of the Family 6 Carbohydrate Binding Module from *Cellvibrio mixtus* Endoglucanase 5A in Complex with Oligosaccharides Reveals Two Distinct Binding Sites with Different Ligand Specificities J. Biol. Chem., 279: 21560 – 21568

Reverbel-Leroy, C., G.Parsiegla,V.Moreau, M. Juy, C. Tardif, H. Driguez, J.-P. Bélaich and R. Haser. Crystallization of the catalytic domain of *Clostridium cellulolyticum* CeIF cellulase in the presence of a newly synthesized cellulase inhibitor *Acta Cryst.* (1998). D54, 114-118

Riedel,K., K. Bronnenmeier, Active-site mutations which change the substrate specificity of the *Clostridium stercorarium* cellulase CelZ: Implications for synergism Eur. J. Biochem. 262 (1999) 218-223.

Rouvinen, J., Bergfors, T., Teeri, T., Knowles, J.K.C. and Jones, T.A. (1990)) Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei.* Science 249, 380-386.

Sakon, J.; Adney, W. S.; Himmel, M. E.; Thomas, S. R.; Karplus, P. A. 1996. Crystal Structure of Thermostable Family 5 Endocellulase E1 from *Acidothermus cellulolyticus* in Complex with Cellotetraose Biochemistry; 35:10648-10660.

Sandgren, M., Ståhlberg, J.and Mitchinson, C. (2005). Structural and biochemical studies of GH family 12 cellulases: improved thermal stability and ligand complexes. *Prog. Biophy. Mol. Bio.* 89, 246 – 291.

Sandgren, Mats, Andrew Shaw, Traci H. Ropp, Shan Wu, Richard Bott, Alexander D. Cameron, Jerry Ståhlberg, Colin Mitchinson and T.Alwyn Jones 2001 The X-ray crystal structure of the *Trichoderma reesei* family 12 endoglucanase 3, Cel12A, at 1.9 Å resolution J. Mol. Biol. 308, 295-310

Shimon, L. J. W., S. Pagès, A. Belaich, J.-P. Belaich, E. A. Bayer, R. Lamed, Y. Shoham and F. Frolow. Structure of a family IIIa scaffoldin CBD from the cellulosome of *Clostridium cellulolyticum* at 2.2 Å resolution *Acta Cryst.* (2000). D56, 1560-1568

Shimon, Linda JW,Edward A Bayer, Ely Morag, Raphael Lamed, Sima Yaron, Yuval Shoham, and Felix Frolow A cohesin domain from *Clostridium thermocellum*: the crystal structure provides new insights into cellulosome assembly 1997 Structure 5: 381-390.

Shirai, Tsuyoshi, Hirokazu Ishida, Jun-ichi Noda, Takashi Yamane, Katsuya Ozaki, Yoshihiro Hakamada and Susumu Ito 2001 Crystal structure of alkaline cellulase K: insight into the alkaline adaptation of an industrial enzyme J. Mol. Biol. 310, 1079-1087

Shirai,Tsuyoshi, Hirokazu Ishida, Jun-ichi Noda, Takashi Yamane, Katsuya Ozaki, Yoshihiro Hakamada and Susumu Ito 2001 Crystal structure of alkaline cellulase K: insight into the alkaline adaptation of an industrial enzyme J. Mol. Biol.310,1079-1087

Simon, J., Charnock, Tracey D. Spurway, Hefang Xie, Marie-Hélène Beylot, Richard Virden, R. Antony J. Warren, Geoffrey P. Hazlewood, and Harry J. Gilbert 1998 The Topology of the Substrate Binding Clefts of Glycosyl Hydrolase Family 10 Xylanases Are Not Conserved J. Biol. Chem., 273**:** 32187 – 32199

Sinnott, M. L. (1990). Catalytic mechanism of enzymic glycosyl transfer. Chem. Rev. 90, 1171-1202.

Ståhlberg,Jerry, Hongbin Henriksson, Christina Divne, Roland Isaksson, Göran Pettersson, Gunnar Johansson and T. Alwyn Jones 2001 Structural basis for enantiomer binding and separation of a common β-blocker: crystal structure of cellobiohydrolase Cel7A with bound (*S*)-propranolol at 1.9 Å resolution J. Mol. Biol. 305,79-93

Szabo, L., Jamal, S., Xie, H., Charnock, S. J., Bolam, D. N., Gilbert, H. J. & Davies, G. J. (2001). Structure of a family 15 carbohydrate-binding module in complex with

xylopentaose. evidence that xylan binds in an approximate 3-fold helical conformation. J. Biol. Chem. 276, 49061–49065.

Tahirov T.H., Oki H., Tsukihara T., Ogasahara K., Yutani K., Libeu C.P., Izu Y., Tsunasawa S., Kato I., High-Resolution Crystal of Methionine Aminopeptidase from Pyrococcus furiosus Obtained by Water-Mediated Transformation J. Struct. Biol. 121 (1998) 68–72.

The Crystal Structure of a Family 5 Endoglucanase Mutant in Complexed and Uncomplexed Forms Reveals an Induced Fit Activation Mechanism J. Mol. Bio., 1996 257, 1042-1051

Tomme, P., Warren, R. A. J., Miller, R. C. J., Kilburn, D. G., and Gilkes, N. R. (1995) in *Enzymatic Degradation of Insoluble Polysaccharides* (Saddler, J. N. , and Penner, M., eds), Vol. 618 , pp. 142-163, American Chemical Society, Washington, D. C.

Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y. & Steitz, T. A. (1996). Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. 15, 5739–5751.

Varrot, Annabelle, Martin Schülein and Gideon J. 2000 Davies Insights into ligand-induced conformational change in Cel5A from *Bacillus agaradhaerens* revealed by a catalytically active crystal form J. Mol. Biol.297, 819-828

Vincent, Florence, Simon J. Charnock, Koen H. G. Verschueren, Johan P. Turkenburg,

Violot,S., R. Haser, G. Sonan, D. Georlette, G. Feller and N. AghajariExpression, purification, crystallization and preliminary X-ray crystallographic studies of a psychrophilic cellulase from *Pseudoalteromonas haloplanktis. Acta Cryst.* (2003). D59, 1256-1258

Xie, H., Bolam, D. N., Nagy, T., Szabo, L., Cooper, A., Simpson, P. J. et al. (2001). Role of hydrogen bonding in the interaction between a xylan binding module and xylan. Biochemistry, 40, 5700–5707.

Xie, H., Gilbert, H. J., Charnock, S. J., Davies, G. J., Williamson, M. P., Simpson, P. J. et al. (2001). *Clostridium thermocellum* Xyn10B carbohydrate-binding module 22-2: the role of conserved amino acids in ligand binding. Biochemistry, 40, 9167–9176.

APPENDIX

gi|1346278 Beta-hexosaminidase B precursor (N-acetyl-beta glucosaminidase)(Beta-GlcNAcase) (Beta-N-acetylhexosaminidase) (Beta-NAHase)

gi|1170252 Beta-hexosaminidase (N-acetyl-beta-glucosaminidase) (Beta-N-acetylhexosaminidase) (Chitobiase)

gi|10835638 Chain A, Beta-N-Acetylhexosaminidase Mutant D539a Complexed With Di- N-Acetyl-Beta-D-Glucosamine (Chitobiase)

gi|3978256 beta-hexosaminidase [*Pseudoalteromonas* sp. S9]

gi|5881581 transglycosylative enzyme [*Alteromonas* sp. O-7]

gi|2760908 family 10 xylanase [*Caldicellulosiruptor* sp. Rt69B.1]

gi|3810965 110kDa xylanase (XynE) [*Aeromonas punctata*]

gi|3201481 endo-1,4-beta-xylanase [*Bacillus* sp. BP-23]

gi|12225048 xylanase A [*Clostridium josui*]

gi|5360744 xylanase C [*Clostridium stercorarium*]

gi|27227837 xylanase 5 [*Paenibacillus* sp. W-61]

gi|974180 xylanase [*Eubacterium ruminantium*]

gi|1103639 endoxylanase [*Cellulomonas fimi*]

gi|5880612 endo-1,4-beta-xylanase [*Cellulomonas pachnodae*]

gi|72161672 putative secreted cellulose-binding protein [*Thermobifida fusca* YX]

gi|71915404 putative secreted cellulose-binding protein [*Thermobifida fusca* YX]

gi|6434713 putative secreted cellulose-binding protein. [*Streptomyces coelicolor* A3(2)]

gi|34498778 carbohydrate-binding protein [*Chromobacterium violaceum* ATCC 12472]

gi|50937973 putative endoxyloglucan transferase [*Oryza sativa* (japonica cultivar-group)]

gi|50937975 xyloglucan endotransglycosylase-like protein [*Oryza sativa* (japonica cultivar-group)

cd02850

gi|46015849 Chain A, Structural Basis For The Exocellulase Activity Of The Cellobiohydrolase Cbha From *C. Thermocellum*

gi|6272570 endoglucanase K [*Clostridium cellulovorans*]
gi|37703323 Cel9B [*Ruminococcus albus*]
gi|13423734 glycosyl hydrolase, family 9 [*Caulobacter crescentus* CB15]
gi|6468241 secreted endoglucanase [*Streptomyces coelicolor* A3(2)]
gi|48837427 Uncharacterized protein contain chitin-binding domain type 3 [*Thermobifida fusca*].

gi|67987581 Cellulase [*Kineococcus radiotolerans* SRS30216]
gi|4104166 endoglucanase [Pseudomonas sp. YD-15]
gi|48861645 Cellobiohydrolase A (1,4-beta-cellobiosidase A) [*Microbulbifer degradans* 2-40].
gi|1064979 Endoglucanase Celd; Ec: 3.2.1.4; Synonyms: 1,4-Beta-D-Glucan-Glucanohydrolase, Cellulase Endo-1,4-Beta-Glucanase D. SOURCE *Clostridium thermocellum*

pfam02927.12p3

gi|1200196 endo-1,4-beta-D-glucanase SOURCE *Fibrobacter succinogenes*
gi|606791 endoglucanase D SOURCE *Fibrobacter succinogenes*
gi|6572464 endo-1,4-beta-D-glucanase [*Xanthomonas campestris* pv. campestris]
gi|4104166 endoglucanase [*Pseudomonas* sp. YD-15]
gi|6468241 secreted endoglucanase [*Streptomyces coelicolor* A3(2)]
gi|530014 cellulose 1,4-beta-cellobiosidase [*Clostridium thermocellum*]
gi|121806 Endoglucanase A precursor (Endo-1,4-beta-glucanase) (Cellulase) (EGA)
gi|1769559 PbgB [*Fusobacterium mortiferum*]
gi|1663519 endoglucanase [*Clostridium thermocellum*]
gi|1064979 Endoglucanase Celd; Ec: 3.2.1.4; Synonyms:1,4-Beta-D-Glucan

Glucanohydrolase, Cellulase Endo-1,4-Beta-Glucanase D.
SOURCE: *Clostridium thermocellum*

gi|72161672 putative secreted cellulose-binding protein [*Thermobifida fusca* YX]

gi|21219173 secreted cellulose-binding protein [*Streptomyces coelicolor* A3(2)]

gi|6434713 putative secreted cellulose-binding protein. [*Streptomyces coelicolor* A3(2)]

gi|27261026 putative endoxyloglucan transferase [*Oryza sativa* (japonica cultivar group)]

gi|71915404 putative secreted cellulose-binding protein [*Thermobifida fusca* YX]

gi|29828710 secreted cellulose-binding protein [*Streptomyces avermitilis* MA-4680]

gi|29605814 putative secreted cellulose-binding protein [*Streptomyces avermitilis* MA-4680]