

Structural-Biochemical and Mechanistic Studies of Two Novel and Versatile Archaeal Glucanotransferase-cum-Exoamylase Enzymes

A Thesis

Submitted
By

Pallavi

PH12104

For the award of the degree of

Doctor of Philosophy



Department of Biological Sciences
Indian Institute of Science Education and Research (IISER) Mohali
Sector-81, Mohali-140306, Punjab, India

Dedicated to

Maa

Declaration

The work presented in this thesis entitled “Structural-Biochemical and Mechanistic Studies of Two Novel and Versatile Archaeal Glucanotransferase-cum-Exoamylase Enzymes” has been carried out by me under the supervision of Professor Purnananda Guptasarma at the Indian Institute of Science Education and Research (IISER), Mohali. This work has not been submitted in part or full for a degree, a diploma, or a fellowship to any university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgment of collaborative research and discussions. This thesis is a bona fide record of original work done by me and all sources listed within have been detailed in the references.

Date:

Pallavi

Place:

(Candidate)

In my capacity as the supervisor of the candidate’s thesis work, I certify that the above statements by the candidate are true to the best of my knowledge.

Date:

Prof. Purnananda Guptasarma

Place:

(Supervisor)

Acknowledgments

The journey of my Ph.D. has been guided and accompanied by many people who I can never thank enough in words.

I would still start by thanking my supervisor, Professor Purnananda Guptasarma, for his constant inspiration, guidance and thought-provoking discussions. His ability to analyze things (both scientific data and life situations) is admirable and unparalleled. He gives a new perspective to everything which is a quality worth learning from him. I will cherish the time spent in his lab and lab meetings for years to come and he will always be remembered as a source of great inspiration and strength. I thank him for providing me with the opportunity to work in his lab and helping me learn about the ways and means of science as well as life.

I extend my thanks to my doctoral committee members: Dr. Samrat Mukhopadhyay and Dr. Shashi Bhushan Pandit for their suggestions which helped improve the work.

I would also like to thank all the senior lab members Dr. Satya Prakash, Dr. Purna Sharma, Dr. Kanika Arora, Dr. Sukhdeep Kumar, Dr. Nitin Kishor, Dr. Javed Khan, Dr. Neeraj Dhaunta and Dr. Rajendra for helping me familiarize with the lab set up, basic experimental techniques and their timely suggestions. Special thanks to Dr. Purna Sharma for her timely support and inspiration and to Dr. Arpana Kumari for lending her support and loving company at all times.

I thank the present lab members Bhisem Thakur, Arpita Mrigwani, Arpita Sarkar and Archit Gupta for maintaining a friendly and cheerful atmosphere in the lab. A special mention is required for Prince Tiwari for his encouragement and support during the inevitable rough times that one goes through during a Ph.D. He is also thanked for his suggestions that helped me in troubleshooting various problems and company at all times.

I would also like to thank the master thesis students of the lab Manmeet, Prachi, Anu, Abhishek, Shahnaz, Santosh, Sreelakshami, Anupama and Harpreet for their loving company.

I would like to extend my thanks to all my batchmates and especially to Shivani, Rohan, and Manisha for helping me in every possible way throughout the course of my Ph.D.

I would also like to extend my gratitude towards Mr. Amoldeep Singh Kainth, my mentor during my Masters' course and the one who truly inspired and encouraged me in the direction of research.

I wish to thank all the staff members of IISER Mohali for their timely support.

A special thanks to all my family members (especially to my Mamaji and Mamiji), my sister Bhavya and brother Daksh as well as my friends Dr. Priyasha, Sheveta, Preetika, and Asmita for their unwavering love and support and their ability to keep me high in spirits.

I would like to express my deepest gratitude to the Almighty without whose blessings nothing is possible.

Financial support from IISER Mohali and CPSDE are also acknowledged.

Abstract

Carbohydrates constitute a large share of all the macromolecules found in nature. The study of their structure, function, and synthesis has increased in importance greatly in recent times, in the fields of both food and health. In the food industry, modified carbohydrates are in the limelight. In the health sector, the purpose, function, and patterns of glycosylation of proteins are gaining importance. Modifications of carbohydrates are mainly carried out by carbohydrate-modifying enzymes. Such enzymes have recently been classified and categorized on the basis of their similarities of sequence, in the CAZy database, and on the basis of their functions in glycosidic bond cleavage, bond formation, debranching functions, isomerization functions etc. The product of each of these enzyme functions is essential at some stage of carbohydrate metabolism and, consequently, in the industry related to carbohydrates. This had led to the search for enzymes with new specificities, and features, involving the formation of unique products, or versatile enzymes featuring multiple specificities or catalytic functions. In particular, there is a focus on enzymatic functions that include both the breaking as well as synthesizing of glycosidic bonds, as glycosyl hydrolases and as glycosyltransferases, respectively. Enzymes that perform these functions take part in various metabolic pathways, and are useful in the industry: in the generation of sweeteners, novel dietary carbohydrates, etc., cosmetic industry: as thickening agents, detergent industry, textile industry and with the recent emphasis on generation and use of clean energy such enzymes find a huge role in production of biofuels.

Glycosyl Hydrolases are enzymes that hydrolyze/break a glycosidic bond to generate products smaller in size than the initial substrate. They include broadly endo-acting enzymes as well as exo-acting enzymes to create a variety of oligosaccharide species. Glycosyltransferases are enzymes that transfer sugars/glucans from donor to acceptor molecules to produce oligosaccharides of varying lengths. The choice of donor and acceptor molecules and the degree of polymerization of the products formed depends upon the source of the enzyme i.e. the organism from which it is derived and the role evolution has played in its maturation.

The glycosyltransferases of higher organisms are broadly classified into Leloir and non-Leloir enzymes. The Leloir enzymes use glycosyl esters of nucleoside

mono or diphosphates as donor molecules. Whereas, non-Leloir enzymes use glycosyl phosphates as donors molecules. Glycosyltransferases from lower order organisms are however classified on the basis of the smallest donor and acceptor molecules they utilize (since they mainly utilize glucose based saccharides).

The Archaea was categorized as the third domain of life in the late 1970s. The use of archaeal enzymes in different processes has many advantages when used in large-scale industrial processes. Since most archaea are found in habitats characterized by harsh physical and chemical conditions, enzymes isolated from them are found to be able to function at high temperatures and highly acidic/alkaline pH conditions, and they are found to contain stable secondary, tertiary and quaternary structures. These properties make archaeal enzymes highly suitable candidates for various industrial processes. Most enzymes of archaeal origin are also promiscuous in nature, i.e., they are less specific and can work with a wider range of chemically-similar substrates. This property can be utilized by incorporating protein engineering techniques and creating modifications in order to exploit their full potential.

Of the carbohydrate-modifying enzymes, glucanotransferases are enzymes which cleave a glycosidic bond from a donor molecule to release a small sugar (which can either be a monosaccharide or a disaccharide) which is then transferred to another carbohydrate or non-carbohydrate entity, to yield products of varying lengths and chemical structures.

The present thesis deals with an enzyme from *Pyrococcus furiosus* which was characterized as an α -amylase in 1993, but which was re-classified as a 4- α -Glucanotransferase in 2005. Since then, the enzyme has been referred to in the literature as a 4- α -Glucanotransferase, and there is no discussion of its possible amylase function. We refer to this enzyme as PfuAmyGT, keeping in mind its organism of origin (Pfu), and its possible functions as an amylase (Amy) and a glucanotransferase (GT). To better understand the action of hyperthermophilic 4- α -Glucanotransferases, we also cloned a related enzyme from another organism *Thermococcus onnurineus*, which we refer to as TonAmyGT. Both of these enzymes were studied biophysically to assess their: (1) Secondary structure, using Circular Dichroism (CD) Spectroscopy, (2) Tertiary structure, using Fluorescence

Spectroscopy, (3) Quaternary structure, using Gel filtration chromatography and Dynamic Light Scattering, (4) Thermal stability, using CD Spectroscopy and Differential Scanning Calorimetry, and (5) Chemical stability, using CD Spectroscopy. They were also studied biochemically and in terms of their activity as an amylase and as a glucanotransferase by (6) the Starch-iodide method to establish amylase activity, (7) Zymography to study amylase activity, and (8) Thin Layer Chromatography (TLC) to study both amylase and transferase activity. In addition, these proteins were also studied: (9) Bioinformatically, for comparison with 4- α -Glucanotransferases whose crystal structures are available, and to build reasonable hypotheses regarding their mechanisms of function, in keeping with all the experimental data from the biochemical and other studies.

With PfuAmyGT, we found it to possess an exo-amylase activity when starch was used as a substrate which appeared to be enhanced in the presence of maltose. The exo-amylase activity which is observed at 90 °C with starch, could be observed at ambient temperatures too when maltose was present along with starch. These results were further confirmed by zymography. In TLC experiments, we observed the formation of varying lengths of oligosaccharides by the action of PfuAmyGT on starch. The intensities of these oligosaccharides increased when maltose was added to the reaction mixture. The formation of oligosaccharides can only be explained when PfuAmyGT is considered to be an amylase-cum-glucanotransferase, which uses maltose as a sink for the addition of glucose units derived through exo-amylase action upon starch. Thus we established that PfuAmyGT contains both exo-amylase and transferase activities. As a glucanotransferase, we established the smallest donor to be maltotriose, the smallest acceptor to be glucose, and the smallest transferred unit to be glucose. These features cause us to categorize PfuAmyGT as a new type of glucanotransferase, rather than as a known type of glucanotransferase.

Bioinformatic analyses based on a sequence-homology-based prediction of the structure revealed that the structure of PfuAmyGT contains three domains:

Domain 1 (D1): a (β/α)₇ barrel domain housing the catalytic site for glycosidic bond cleavage.

Domain 2 (D2): a small segment made of α helix and loops belonging to the Domain of Unknown Function (DUF) 1925.

Domain 3 (D3): a β -sandwich fold belonging to the Domain of Unknown Function (DUF) 1925.

In order to understand the function of each domain in PfuAmyGT and also gain insights into its mode of action, we individually cloned each domain as well as in combinations, but could not observe any activity. On comparison of the predicted structure of PfuAmyGT with a 4- α -Glucanotransferase from *Thermococcus litoralis* (whose crystal structure is available), we identified the putative donor and acceptor sites in PfuAmyGT and also a tryptophan bearing loop which could possibly act to transfer the excised glucose unit from donor to acceptor molecule. To test this hypothesis, we created three point mutations in the loop region, one residue is likely to participate in a catalytic action upon the donor glucan, and other two are likely to be involved in the transfer of the excised glucan. We observed loss of transferase activity on TLC in case of the mutation (to alanine) of a loop residue suspected to participate in glucoside bond hydrolysis (involving an aspartate residue), and also in one case of mutation of a residue suspected to transfer the glucan (a tryptophan). Therefore, we established that the tryptophan is an essential residue taking part in the transferase action of PfuAmyGT.

With TonAmyGT, we cloned, expressed and purified the protein to determine its structure, oligomeric status, thermal and chemical stability. Its activity was also examined using TLC. Briefly, in comparison with PfuAmyGT, we found that it uses maltose as the smallest donor as well as the smallest acceptor glucan, although this enzyme also uses glucose as the smallest transferred glucan.

We believe that these studies of two amylases-cum-glucanotransferases provide new mechanistic insights, and applications, to the field of carbohydrate modifying enzymes.

Table of Contents

1. Introduction	1
1.1 Exo and Endoamylases.....	7
1.2 Different categories of Glucanotransferases.....	10
1.3 Single active site v/s Split active sites.....	13
1.4 Archaeal enzymes.....	14
1.5 Glucanotransferase from <i>Pyrococcus furiosus</i> originally annotated as an amylase (PfuAmyGT).....	15
1.6 4- α -Glucanotransferase from <i>Pyrococcus furiosus</i>	18
1.7 The multi-domain architecture of PfuAmyGT.....	19
1.8 The multi-domain architecture of an enzyme from <i>Thermococcus litoralis</i> which is homologous to PfuAmyGT.....	19
1.9 The issues addressed in this thesis with PfuAmyGT and Ton AmyGT.....	20
2. Materials and Methods	25
2.1 Cloning, Expression and Purification.....	25
2.1.1 Materials.....	25
2.2 Methods.....	35
2.2.1 Cloning.....	35
2.2.2 Expression of recombinant proteins in <i>E coli</i>	42
2.2.3 Glycerol stock preparation.....	43
2.2.4 Nondenaturing/ Native protein purification.....	43
2.2.5 Refolding of proteins.....	43
2.2.6 Mass Spectrometry.....	44
2.2.7 Protein concentration estimation.....	45
2.2.8 Gel filtration chromatography.....	46
2.2.9 Fluorescence spectroscopy.....	46
2.2.10 Circular Dichroism (CD) spectroscopy.....	46
2.2.11 Dynamic Light Scattering (DLS).....	47
2.2.12 Differential Scanning Calorimetry (DSC).....	47
2.2.13 Activity Assays.....	48
2.2.14 Structure analysis tools.....	49
3. Results and Discussion	51
3 Physico-chemical characteristics of PfuAmyGT and its domains.....	51
3.1 Cloning, expression and purification.....	51
3.2 Confirmation of the chemical identities of all domains constructs by mass spectrometry.....	56
3.3 Secondary, tertiary and quaternary structures of PfuAmyGT and its domains.....	60

3.3.1 Determining the oligomeric status of the proteins.....	60
3.3.2 Determination of the secondary structure of PfuAmyGT.....	63
3.3.3 Determination of the presence of tertiary structure.....	64
3.3.4 Stability of PfuAmyGT and its domains.....	67
3.4 Amylase activity in PfuAmyGT and its domains.....	70
3.4.1 Starch-iodine and TLC based assays of amylase activity.....	70
3.4.2 Plate-based assay.....	71
3.4.3 Zymogram.....	72
3.4.4 Thin Layer Chromatography (TLC).....	73
3.4.5 Amylase activity is significant at higher temperatures.....	74
3.4.6 Amylase activity is aided by the addition of maltose.....	74
3.4.7 Amylase activity co-operates with a glucanotransferase activity to create multiple oligosaccharides from a single substrate.....	77
3.5 Characterization of Glucanotransferase activity of PfuAmyGT and its domains.....	78
3.5.1 Conversion of starch to multiple oligosaccharides is temperature dependent.....	78
3.5.2 Maltose accelerates the reaction by acting as an acceptor.....	79
3.5.3 The transferred unit is glucose.....	80
3.5.4 The smallest donor capable of transferring glucose is maltotriose.....	81
3.5.5 The smallest acceptor capable of accepting the transferred glucose is glucose itself.....	82
3.5.6 Time-dependence of PfuAmyGT activity.....	85
3.5.7 Temperature-dependence of PfuAmyGT activity.....	85
3.5.8 pH-dependence of PfuAmyGT activity.....	86
3.5.9 Effect of organic solvents.....	88
3.5.10 Activity of PfuAmyGT in the presence of denaturants.....	93
3.5.11 Glucanotransferase activity of the domains of PfuAmyGT.....	94
3.6 Identification of putative donor and acceptor sites through structural bio-informatic analysis.....	96
3.7 Glucose is transferred from the head of the donor saccharide to the tail of the acceptor saccharide.....	100
3.8 The acceptor site is an open groove while the donor site is a buried tunnel.....	101
3.9 Acceptor site lies in domain 2 and donor site is in domain 1.....	103
3.10 A tryptophan bearing loop might transfer the glucose.....	103
3.11 Mutational analysis.....	104
3.11.1 Gel filtration chromatography of mutant forms of PfuAmyGT.....	105

3.11.2 Determination of the secondary and quaternary structure of mutants...	106
3.11.3 Thermal stability of the mutants.....	108
3.11.4 Activity/behavior of the mutants.....	109
3.12 Acarbose inhibits the glucanotransferase activity of PfuAmyGT.....	111
3.13 Versatility of full-length PfuAmyGT's glucanotransferase activity.....	112
3.13.1 PfuAmyGT works on sucrose as an acceptor.....	112
3.13.2 Cellulose, Pectin, and Xylan also act as glucose/xylose donors.....	114
3.13.3 Mass spectrometric analysis of products formed.....	116
3.14 Complementation studies.....	120
3.14.1 Co-expression.....	121
3.14.2 Co-refolding.....	122
3.15 Cloning, Expression and Purification of a glucanotransferase from <i>Thermococcus onnurineus</i> (TonAmyGT/TonGT).....	123
3.16 Confirmation of identity by mass spectrometry.....	125
3.17 Oligomeric status of TonAmyGT.....	126
3.17.1 Gel filtration chromatography.....	126
3.17.2 Dynamic Light Scattering (DLS).....	127
3.18 Secondary and tertiary structure of TonAmyGT.....	127
3.19 Thermal and chemical stability of TonAmyGT.....	128
3.20 Activity of TonAmyGT.....	130
3.20.1 Determining the smallest donor and acceptor molecule.....	130
3.20.2 Temperature-dependent activity of TonAmyGT.....	133
3.20.3 Time-dependent activity of TonAmyGT.....	134
3.20.4 pH-dependent activity of TonAmyGT.....	135
4. Conclusion	138

1. Introduction

Introduction

Proteins, the workhorses of life, were first described by the Dutch chemist Gerardus Johannes Mulder. They were given their name by the Swedish chemist Jons Jacob in 1838. In 1933, Linus Pauling postulated the presence of secondary structure in proteins. One of the major (and most exploited) roles of proteins relates to their function as enzymes. The term enzyme was first used by Wilhelm Friedrich Kühne, a professor of physiology at the University of Heidelberg in 1877 (Gurung, 2013). Enzymes are biomolecules that catalyze a wide range of chemical reactions that take place during normal metabolism in living systems, lowering the activation energy barrier that separates reactants and products. The use of enzymes outside cells and living systems dates back to at least ~400 BC when they were used in cheese making and for the preservation of food and beverages by ancient Egyptians. In the modern era, ever since the discovery of the first enzyme, Diastase, in 1833, as a new chemical or biochemical entity, by Anselme Payen, the use of enzymes and their products in different fields has become widely prevalent (source: paraphrased from Wikipedia).

Enzymes which are found naturally in preparations like tissue samples or plant extracts have found uses in various industries relating to the processing of food, textile, leather etc. Fermentation technologies have revolutionized the production of enzymes from different sources and enhanced their usage. Recombinant DNA technology further revolutionized the fermentation-based sourcing of enzymes from microbial secretions or cytoplasmic sources, in terms of large-scale production, as well as in terms of enzyme engineering through the construction/formation of enzymes that are not found to occur in nature (Hans E. Schoemaker, 2010).

The current industrial applications of enzymes exploit mainly those that perform hydrolysis reactions, i.e., enzymes used in hydrolysis or modification of carbohydrates, proteins, lipids or DNA, as well as other substrates remain in great demand. With the advent of green technologies and the emphasis currently laid on protecting the environment from the use and generation of hazardous chemicals in everyday processes, the focus is upon enzymes that can degrade waste material derived from plant tissue, which dominates the planet's biomass, and other sources (Pedro Fernandes, 2010).

Introduction

Since polysaccharides are the most abundant component of biomass, the demand for carbohydrate-active or carbohydrate-processing enzymes has increased tremendously in recent years, especially in relation to the production of biofuels. Polysaccharides are also present around us in the form of food as well as in items that we use on a daily basis, like paper, cosmetics, flocculants, encapsulants, adhesives thickening agents, pharmaceuticals, etc. (Bemiller, 2009). In the case of the food industry, the use of enzymes dates back to the production of vinegar. The use of enzymes was then extended to the production of beer, bread, cheese, and wine. Carbohydrates require modifications in the form of their hydrolysis, oxidation, amidation, ester formation, glycosylation, etc. for optimum utilization in terms of food products as well as other resources.

A large repertoire of enzymes is now available to us, and a common classification system is used worldwide to refer to enzymes performing different functions. For this purpose, the International Union of Biochemistry (IUB) set up the Enzyme Commission in 1956 (IUB, 1961), and this commission classified enzymes into six major categories (Cornish-Bowden, 2014):

EC 1- Oxidoreductases: They catalyze the reactions in which a substrate donates one or more electrons to an electron acceptor, itself becoming oxidized.

EC 2- Transferases: They catalyze reactions in which a chemical group is transferred from a donor substrate to an acceptor substrate.

EC 3- Hydrolases: They catalyze reactions in which a bond in a substrate is hydrolyzed (with the help of water) to produce two products.

EC 4- Lyases: They catalyze non-hydrolytic reactions in which a chemical group is removed from a substrate leaving a double bond.

EC 5- Isomerases: They catalyze one-substrate one-product reactions that can be regarded as isomerization reactions.

EC 6- Ligases: They catalyze the joining of two or more molecules coupled to the hydrolysis of ATP or an analogous molecule. These enzymes are also sometimes called synthetases.

Many kinds of such enzymes from different sources are currently employed to modify different types of carbohydrates (Cheng and Qu-Ming, 2012). Examples

of long-chain carbohydrates commonly processed by enzymes today are cellulose, pectin, and starch.

Cellulose: Cellulose is a linear polysaccharide made up of glucose units joined by glycosidic bonds (Figure 1) found as a major cell wall component in plants. It contains a beta acetal linkage which renders it indigestible by humans who do not possess cellulase enzymes capable of the hydrolysis of such bonds (Zhang *et al*, 2014). Various enzymes have been employed to degrade cellulose and/or modify its properties, to render it more suitable for different industrial processes. These include lipases, proteases, beta-galactosidases, and cellulases or endoglucanases. Lipases are generally employed to convert naturally occurring cellulose into hydroxyethylcellulose and carboxymethylcellulose, both of which tend to be more soluble in water than cellulose. Lipases also cause hydrophobic modifications required to synthesize surfactants from cellulose. Proteases can help graft vinyl groups on hydroxyethylcellulose and β -galactosidases can also be used to graft galactose. Finally, cellulases from diverse sources are already well established as reagents for creating hydrolytic products from cellulose, hydrolyzing it down to cellobiose (consisting of two glucose units) as a substrate for the enzyme cellobiase or beta-glucosidase, which then hydrolyzes cellobiose into glucose.

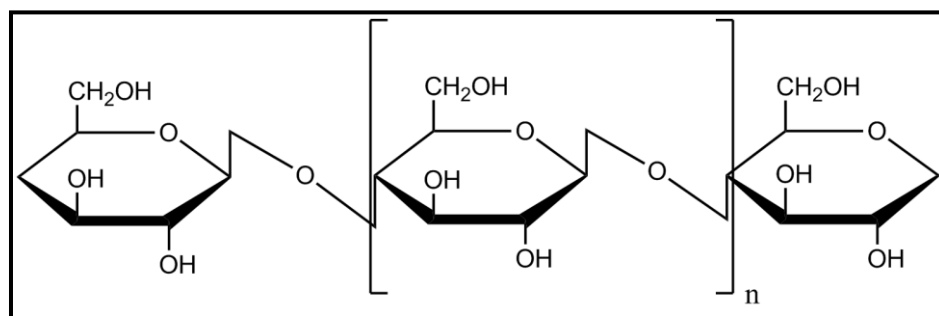


Figure 1: Schematic diagram representing the structure of cellulose.

Pectin: Pectin is another heterogeneous compound whose modified products can be used in various processes. Enzymes like polygalacturonase, pectin lyase, laccase and peroxidase, etc. help to develop processes which make pectins more soluble and degradable into simpler forms.

Starch: Starch is the second most abundant biomolecule which is found in plants. It is present as a reserve, in the form of starch granules within tubers and seed endosperm (source: starch.eu). Starch is mainly composed of 20-30 % amylose which consists of glucose units joined by α -1,4 glycosidic bonds and 70-80 % amylopectin which consists of α -1,4 linkages, as well as α -1,6 linkages that serve to help chains form branches (Figure 2). The ratio of amylose to amylopectin may vary depending upon the source from which starch is isolated. Apart from being used as a food material, starch also finds uses in many other industries because of its ability to bind, thicken, provide texture, stabilize and form gels.

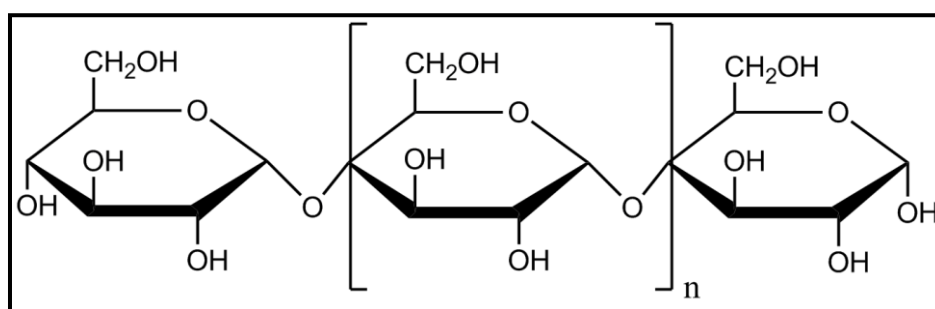


Figure 2: Schematic representation of starch.

The use of starch and its modified, as well as hydrolytic products as food and in other everyday life processes, dates back thousands of years. Thus, starch modifying enzymes are of great interest. Since these enzymes are involved in both hydrolysis and formation of glycosidic bonds, recent re-visitations of the field of glycobiology have highlighted their importance in major areas such as biomass degradation, production of alternative biofuels, human health (glycosylation of proteins and lipids, cell surface glycans as biomarkers of various diseases, etc.) and industrial applications such as the production of sweeteners, use in the baking industry (as anti-staling agent) and production of starch thermo-reversible gels (Chung-wai Chiu and Daniel Solarek, 2009).

The large repertoire of glycosyl hydrolases and glycosyltransferases have now been further sub-classified in many ways. One of these classifies such enzymes on the basis of their substrate specificity and reaction mechanism (Ahemad *et al*, 2015). The enzyme classifications included are:

Endoamylases: These randomly hydrolyze the inner α -1,4 glucosidic linkages of a starch-type polysaccharide (i.e., amylose or amylopectin) (Figure 3). The hydrolysis products may include linear oligosaccharides (such as maltose, which is a disaccharide, as well as longer oligosaccharides known as maltooligosaccharides) as well as branched oligosaccharides. The most common enzyme known (and widely described) under this category is the α -amylase.

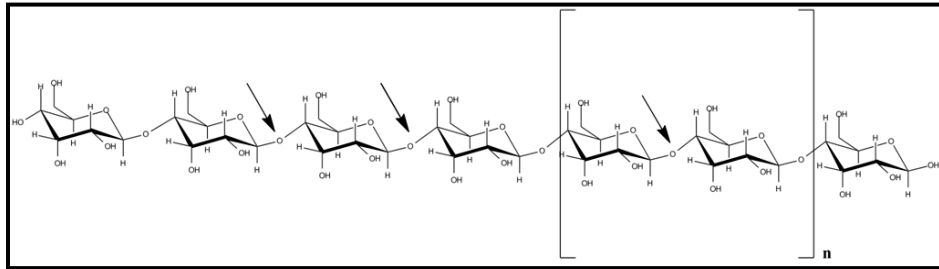


Figure 3: Representation of the action of endoamylases.

Exoamylases: These act on the terminal non-reducing ends of polysaccharides to yield glucose, maltose, and limit dextrins (Figure 4). β -amylases are well known under this category. Glucoamylases also fall in this category but produce β glucose units as products. Similarly, α -Glucosidases hydrolyze α -1,4 glucosidic linkages in disaccharides and oligosaccharides but these are inactive upon large polysaccharides like starch (unlike glucoamylases). The product released in this reaction is α -glucose.

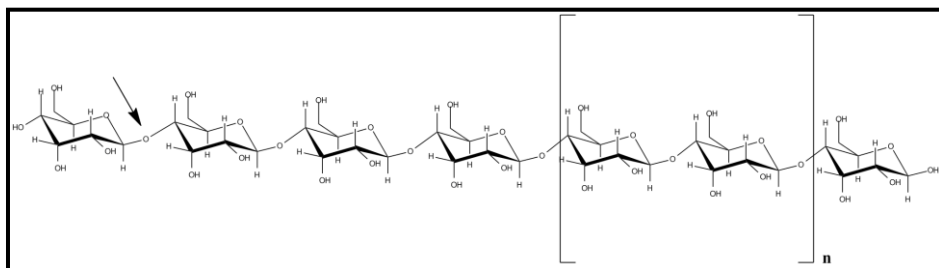


Figure 4: Representation of the action of exoamylases.

Debranching enzyme: These enzymes hydrolyze α -1,6-glycosidic linkages at branch points of polysaccharides, and include pullulanases and isoamylases (Figure 5).

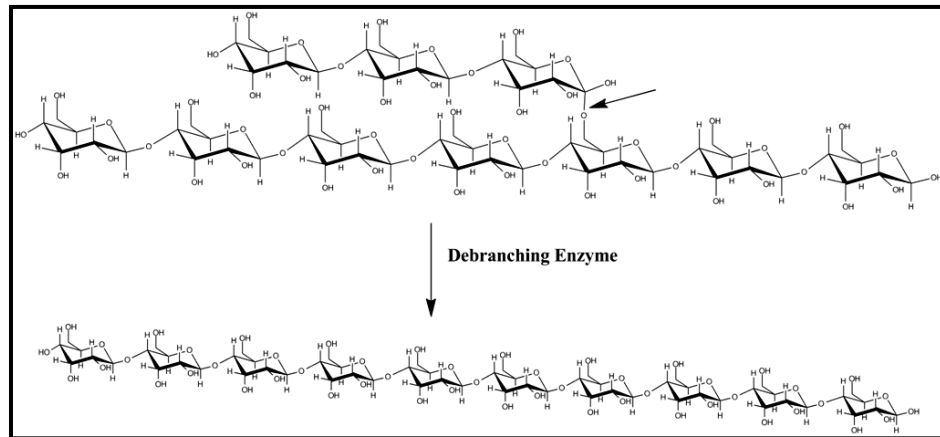


Figure 5: Representation of the action of debranching enzymes.

4- α -Glucanotransferases: These hydrolyze the α -1,4-glycosidic linkages of donor substrate molecules and transfer the resulting product to an acceptor molecule. This is why they are called 'glucanotransferases'. The donor molecules are generally long polysaccharides like starch which are broken down into smaller oligosaccharides. The acceptor molecules are generally small monosaccharide units. The lengths of the donor and acceptor molecules can vary, depending upon the type of enzyme. This aspect has been discussed later.

The introduction of the CAZy database for enzymes in 1998 has re-classified the above-mentioned enzymes on the basis of their amino acid sequences and structural similarities. The database currently covers enzymes catalyzing hydrolysis, synthesis or modification of carbohydrates and glycol-conjugates. They are classified under different categories in different families:

Glycoside Hydrolases (GHs): These are responsible for the hydrolysis and/or rearrangement of glycosidic bonds between two or more carbohydrate molecules, or between a carbohydrate molecule and a non-carbohydrate molecule.

Glycosyl Transferases (GTs): These enable the formation of glycosidic bonds between different activated or non-activated donor molecules and acceptor molecules, to yield a wide variety of products ranging from cyclic molecules to linear chains of varying lengths.

Polysaccharide Lyases (PLs): These perform non-hydrolytic cleavage of glycosidic bonds. They cleave uronic acid-containing polysaccharide chains to generate an unsaturated hexenuronic acid residue and a new reducing end containing the saccharide moiety.

Carbohydrate Esterases (CEs): These facilitate the hydrolysis of carbohydrate esters. The substrates can be the ones in which the sugar of the carbohydrate behaves as an acid, and also the ones in which it behaves as an alcohol.

Auxiliary Activities (AAs): These include other enzymes, such as redox enzymes that act in conjunction with CAZymes.

There are also additional non-catalytic protein units associated with the above-mentioned enzymes called the Carbohydrate-Binding Modules (CBMs) which help in the binding of the enzyme to the substrate carbohydrates included in the database. All enzymes that process carbohydrates do not have such modules. Some possess the carbohydrate-binding and catalytic properties within the same domain that binds to the carbohydrate as substrate. However, others possess an additional module as a CBM domain, which presumably possesses a higher affinity for the carbohydrate than the catalytic domain and helps to localize the enzyme in a particular region of the complex carbohydrate substrate for a longer duration of time, thus allowing the catalytic domain to perform hydrolysis at multiple nearby sites through its own binding and catalytic functions.

1.1 Exo and Endoamylases:

Enzymes acting on starch are generally referred to as amylases. Thus, exoamylases are enzymes which act upon and hydrolyze, the terminal (from the non-reducing end) α -1,4 linked glucose in starch, leading to the formation of glucose and a lower molecular weight saccharide, reduced in size by one glucose

unit. Endoamylases, on the other hand, are enzymes which act on the internal α -1,4 linkages in starch, leading to the formation of saccharides of varying lengths. These lengths are unpredictable because the endoamylase is presumed to bind and hydrolyze glycosidic bonds at random. Both exo and endoamylases have been sub-categorized into different families based on their sequence and/or structural similarities, substrate specificity, etc. but they all follow either one or the other of the following two mechanism(s) of action.

1.1.1 Inverting mechanism:

This mechanism uses direct displacement, where the two active-site carboxylic acid residues (i.e., aspartate or glutamate residues) are oriented in such a way that one acts as a general base (which attacks water) while the other acts as a general acid (which facilitates cleavage of the glycosidic bond) to effectively cleave the glycosidic bond (Carter and Withers, 1994), through an oxocarbenium-ion-like transition state. The distance between the two aspartate/glutamate residues is approximately 0.95 nm or 6-11 Å to allow a water molecule and substrate molecule to bind simultaneously.

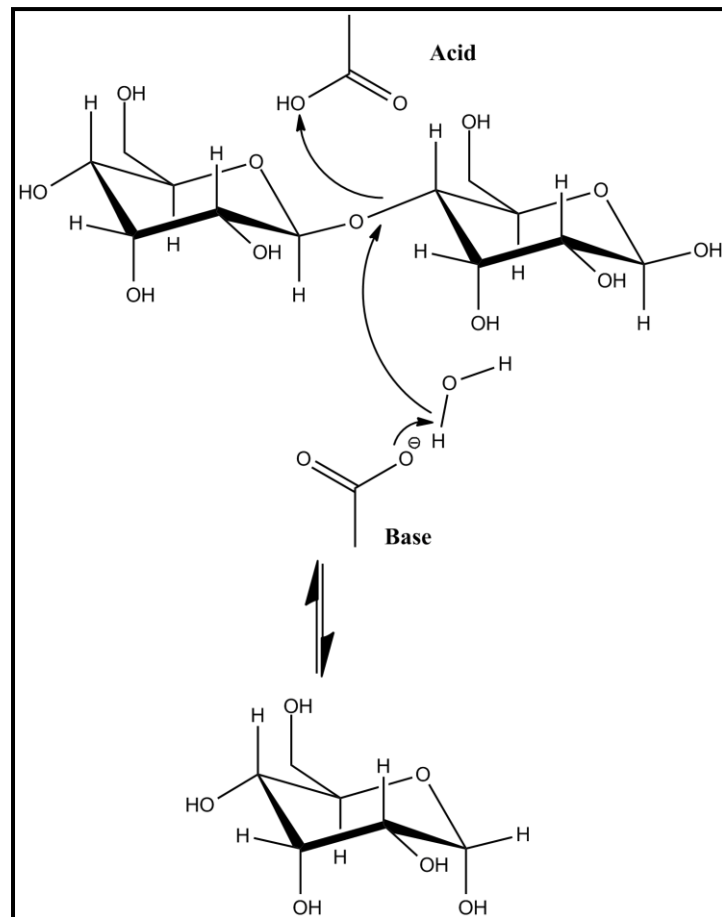


Figure 6: Schematic representing the inverting mechanism of amylases.

1.1.2 Retaining mechanism:

This mechanism exploits double displacement, which is a two-step mechanism (Koshland, 1953). A glycosyl-enzyme intermediate is first formed. One of the two active site carboxylic acid residues then acts as a nucleophile, attacking the sugar anomeric center while the other residue acts as an acid/base catalyst which first protonates the glycosidic oxygen and then helps deprotonate a water molecule. The distance between the two is approximately 0.55 nm or 5.5 Å required for the direct nucleophilic attack.

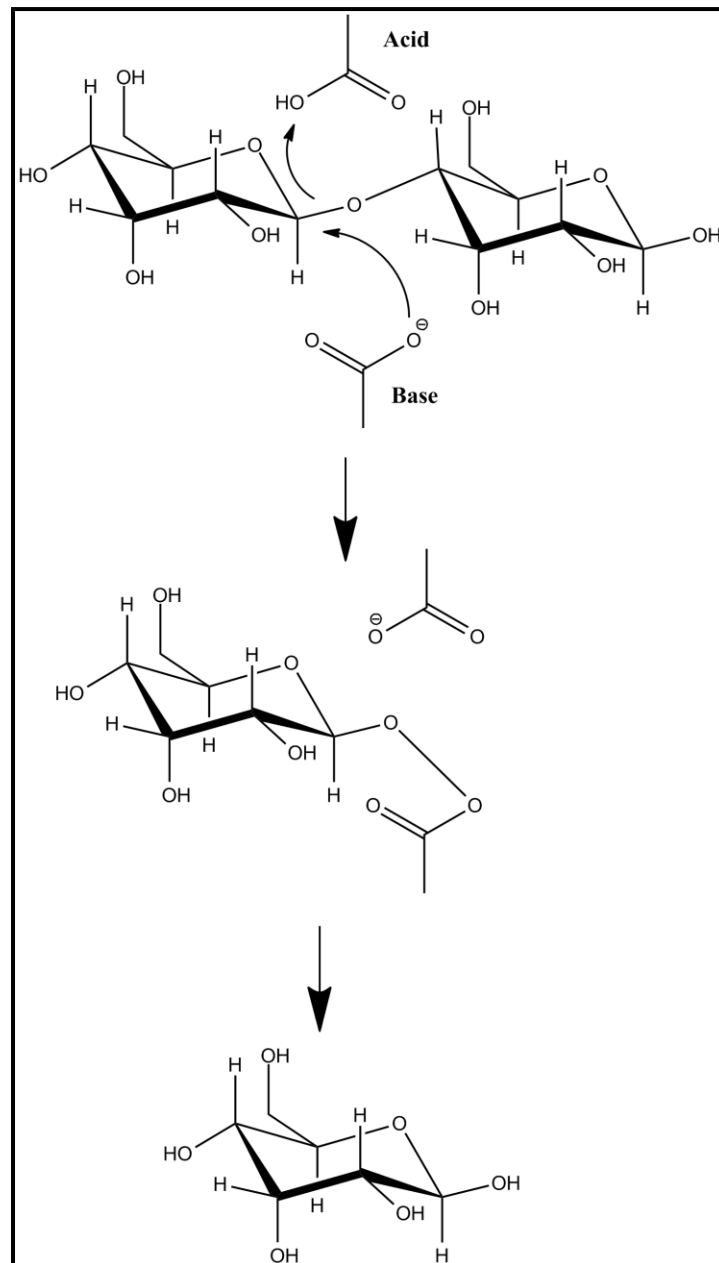


Figure 7: Schematic representing the retaining mechanism of amylases.

1.2 Different categories of Glucanotransferases

4- α glucanotransferases are enzymes catalyzing the transfer of a 1,4- α -D-glucan from a donor molecule to a new position in an acceptor carbohydrate molecule. Thus, these enzymes can both hydrolyze as well as form glycosidic linkages, but not much is known about the sites and mechanisms of these actions in terms of all the residues involved in promoting them in different enzymes, although there is a great deal is surmised about what happens at the level of the substrates which undergo modification through glucanotransferase action. The importance of

studying these enzymes lies in knowing their role *in vivo* as well as in exploiting them for different industrial processes. In 1999, Takaha and Smith classified them mainly on the basis of their reaction specificities into the following sub-categories or types:

Type I: Cyclodextrin Glucanotransferases (CGTases): CGTases hydrolyze α -1,4 glucosidic bonds in a donor molecule and join the hydrolytic product by creating new α -1,4 glucosidic bonds via an intramolecular glucan transfer reaction resulting in the formation of cyclic oligosaccharides known as cyclodextrins (CDs) (Figure 8). The number of glucose units in the ring may vary from six to seven or eight, which corresponds to α , β or γ -CDs, respectively. This cyclization reaction is reversible and the reverse reaction is called a coupling reaction. Some reports suggest that glucose is the smallest acceptor and the smallest transferred unit. The smallest donor molecule may vary from maltose to maltotriose as well as maltotetraose. Three-dimensional structures of CGTases show that they contain $(\beta/\alpha)_8$ -barrel type catalytic domain (Svensson, 1994). Multiple sequence alignment results show four highly conserved regions characteristic of the family GH13 (Svensson, 1994) and three acidic residues likely to be responsible for catalysis; Asp, Glu, and Asp, which have been validated by site-directed mutagenesis studies. These enzymes are mainly used by bacteria for utilization of exogenous glucans as a substrate for growth.

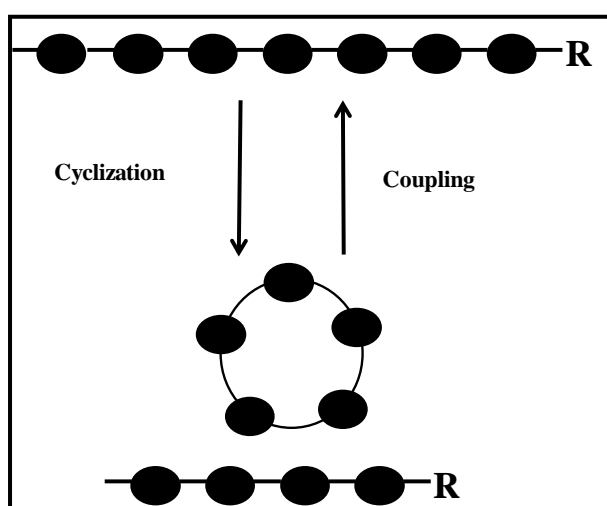


Figure 8: Schematic representation of the CGTase reaction.

Type II: Amylomaltases:

These were initially identified as maltose inducible enzyme in *Escherichia coli* but later also discovered as D-enzyme from potato tubers and other plants. The smallest donor molecule for plant D-enzyme is maltotriose and glucose is the smallest acceptor and the smallest transferred glucan unit is maltose. Plant D-enzymes do not catalyze the transfer of glucosyl units, whereas bacterial amyloamaltases do catalyze such transfer. Another notable feature of the D-enzyme reaction is that maltose is never produced by its action on any substrate because of the presence of two 'forbidden linkages' in maltooligosaccharides larger than maltotetraose. These are (i) the non-reducing end linkage and (ii) the bond penultimate to the reducing end. This enzyme plays a role in starch metabolism (Figure 9). In *E. coli*, amyloamaltase is part of a maltooligosaccharide transport and utilization system while in the case of other organisms, similar enzymes play important roles in glycogen metabolism.

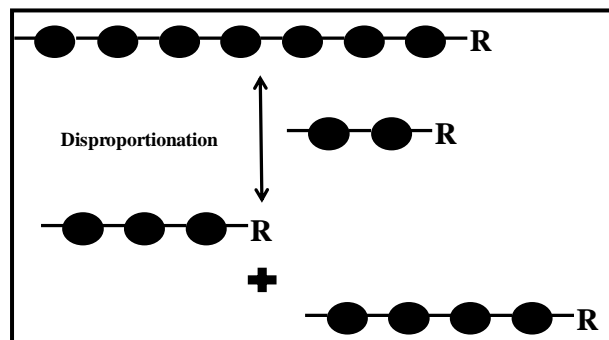


Figure 9: Schematic representation of amyloamaltase reaction.

Type III: Glycogen Debranching Enzyme (GDE):

These enzymes belong to family GH13. In eukaryotes, they perform both amylo-1,6-glucosidase, and 4 α GTase activities. Glycogen breakdown is mainly carried out by glycogen phosphorylase which sequentially removes glucose units in the form of glucose-1-phosphate from the non-reducing end but stops when branch lengths are reduced to four residues. Then the 4 α GTase activity of GDE catalyzes maltotriosyl transfer from the shortened branch to another non-reducing end on the glycogen molecule, which then serves as a new substrate for the phosphorylase. It then allows removal of the residual α -1,6-linked glucosyl residue by the glucosidase activity of GDE.

The smallest donor is maltotetraose, while the smallest acceptor is maltotriose and the smallest unit transferred is the maltose unit.

Type IV:

The enzyme which is known to belong to this class is from the hyperthermophilic bacterium *Thermotoga maritima* and belongs to family GH13. The smallest acceptor for the reaction catalyzed by this enzyme is maltose, the smallest donor is maltotetraose and the smallest transferred unit is maltose. Glucose is not produced as one of the products in the reaction.

Type V:

This group includes enzymes mainly from hyperthermophiles such as *Thermococcus litoralis*, *Pyrobaculum aerophilum*, *Aquifex aeolicus*, *Thermus thermophilus*, *Pyrococcus*, etc. belonging to family GH 57 or GH 77.

The smallest acceptor and transferred unit is glucose and the smallest donor is maltose.

In the work described in this thesis, we shall describe glucanotransferases that belong to new types, such as Type VI and even Type VII, in that we shall describe enzymes sourced from hyperthermophile archaea proteomes/genomes which display combinations of smallest donors, smallest acceptors and transferred glucan unit which is different from those given above.

1.3 Single active site v/s Split active sites:

The heart of an enzyme lies in its active site. In the case of glycoside hydrolases, the surface topology has a huge impact on the types of reaction they catalyze, type and amount of products produced and many other enzymatic properties. The active sites of glycosyl hydrolases have been categorized into the following types based on their topology (Davies and Henrissat, 1995; Davies and Williams, 2016).

1.3.1 Endo-acting enzymes: Since these enzymes have a preference for hydrolyzing internal glycosidic linkages, their active site is generally in the form of an open cleft facilitating access to bonds of elongated substrate molecules.

1.3.2 Exo-acting enzymes: They mainly hydrolyze terminal glycosidic linkages or may occasionally yield di- or trisaccharides, and so their active site is in the form of shallow caves or buried in pockets.

1.3.3 The tunnel: The active site is in form of a tunnel formed by enclosure of the catalytic center by an extended loop. It facilitates processive enzyme action by preventing the release of the polysaccharide substrate and allowing multiple reactions to take place.

1.3.4 The niche: The active site lies in a groove which is blocked at one end and permits exo/endo processive enzymatic activity.

1.4 Archaeal enzymes:

Carl Woese in 1970 established Archaea as the third domain of life (Alquéres *et al*, 2007). Initially, they were thought to inhabit only hostile environments but with the advent of 16S RNA sequencing, genome sequencing and genome-wide transcriptional analysis studies revealed them to be ubiquitous in nature.

Archaea depicts a hybrid of characteristics from both bacteria and eukaryotes. The single cell membrane of Archaea is composed of mainly glycerol-ether lipids. The cells range from 0.1 to over 15 μm in diameter and also possess flagella different in composition from bacteria. The transcriptional and translational machinery of Archaea is similar to Eukarya than Bacteria, like for translation archaea use eukaryotic initiation and elongation factors, and their transcription involves TATA-binding proteins and TFIIB (Alquéres *et al*, 2007 and Bräsen *et al*, 2014).

Owing to the features unique to Archaea, their metabolism is quite different from bacteria as well as Eukarya due to which it encodes very unusual enzymes in its genome. These enzymes can be and are being exploited in various industrial processes like: in trehalose production which used as a stabilizing agent and a preservative agent (Di Lernia *et al*, 2001), proteases used in detergent industries and polymer degradation.

Enzymes of carbohydrate metabolism and polymer-degrading enzymes have recently attracted much attention. Thermophilic and hyperthermophilic enzymes

from different sources are being used owing to their high heat stability, ability to function under harsh conditions of pH and their structural stability. The study that is undertaken in this thesis also deals primarily with one such enzyme from *Pyrococcus furiosus* and a second enzyme from *Thermococcus onnurineus*. Since *Pyrococcus furiosus* belongs to the Archaea, the study undertaken for the enzyme is in terms of understanding its mode of action which would help us to gain insights into the evolution of glucanotransferases as a whole, and also exploit its function in a better way.

1.5 Glucanotransferase from *Pyrococcus furiosus* originally annotated as an amylase (PfuAmyGT):

Members of starch/polysaccharide-modifying enzymes' families generally either hydrolyze or transglycosylate α -glucosidic linkages to produce α -anomeric mono- and oligosaccharides. There are four major types of reactions catalyzed by them:

- Hydrolysis of α -1,4 glycosidic linkages
- Hydrolysis of α -1,6 glycosidic linkages
- Transglycosylation to/for α -1,4 glycosidic linkages
- Transglycosylation to/from α -1,6 glycosidic linkages

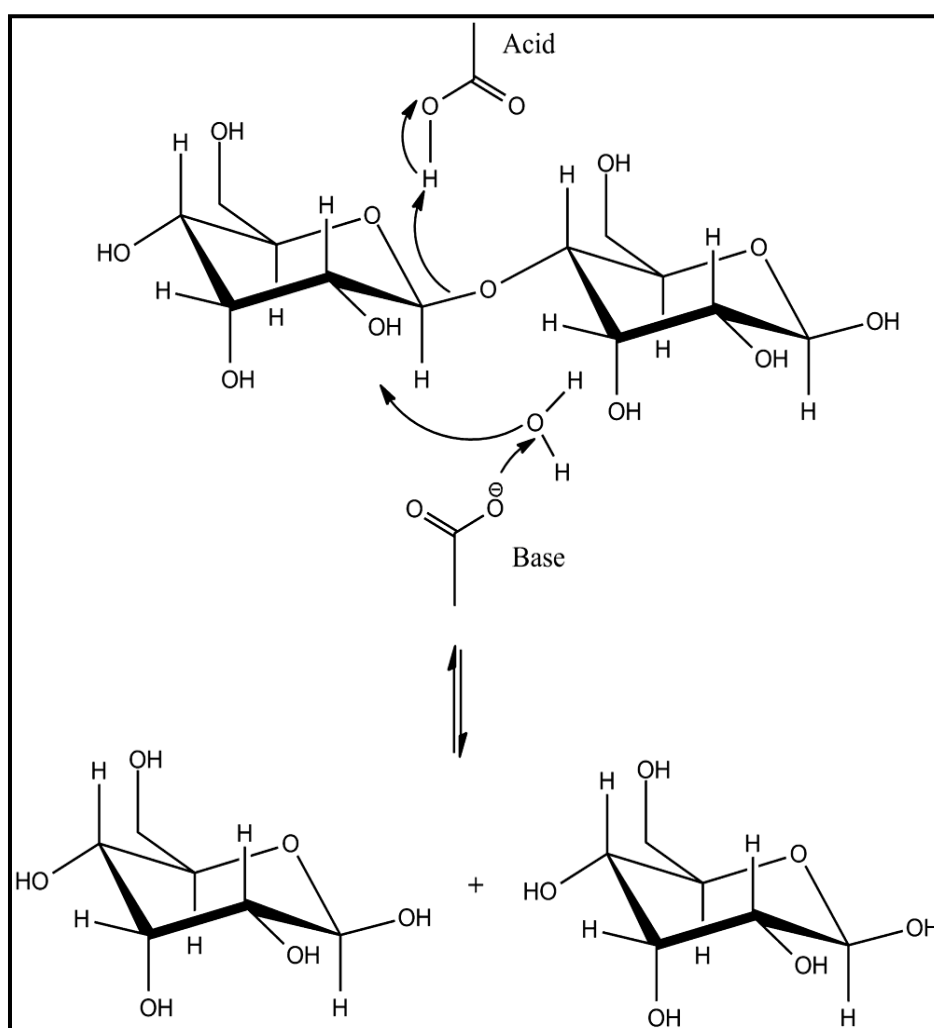
These reactions are typically carried out by α -amylases, pullulanases, cyclomaltoextrin glucanotransferases, and 1,4- α -D glucan debranching enzymes respectively.

But there have been many reports in the literature showing that many enzymes can have dual actions i.e. both hydrolysis and transglycosylation. For example, α -amylases generally show weak α -1,4 transglycosylation activity, CGTases show weak α -1,4 hydrolysis activity and some enzymes show hydrolysis of both α -1,4 and α -1,6 linkages. This is mainly possible because all of the above four reactions follow similar mechanisms. These reactions are carried out via a double-displacement mechanism proceeding with the initial formation of a covalent glycosyl intermediate formed by protonation of the oxygen in the glycosidic bond by the general acid catalyst, usually a Glu residue. Therefore the second step of the reaction, the nucleophilic attack determines the type of the enzyme i.e., in case of hydrolysis, the nucleophile is a water molecule whereas in case of

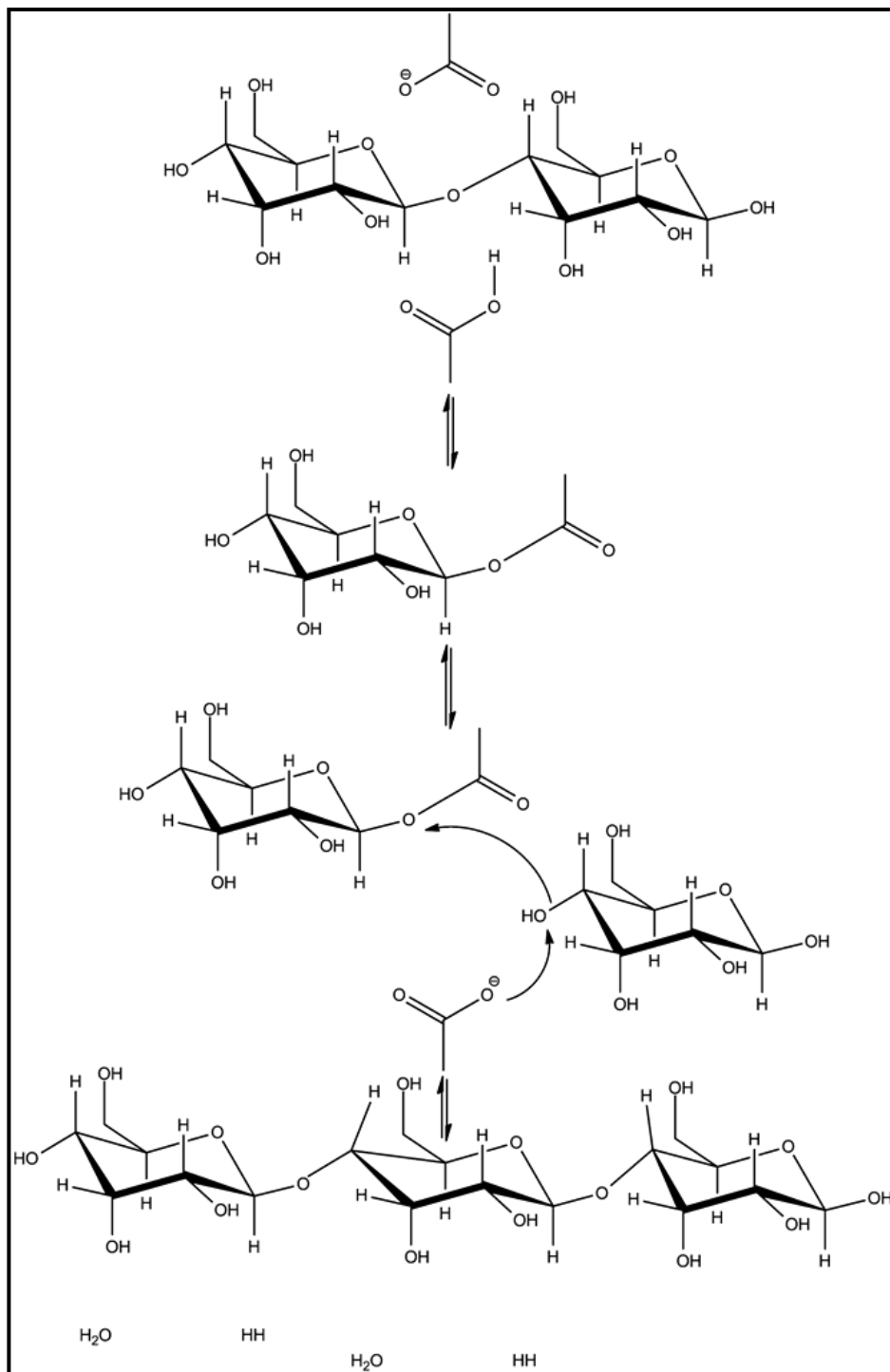
Introduction

transglycosylation reaction the nucleophile is generally the C-4 hydroxyl group of a glucose residue. This leads to the formation of linear oligosaccharides of varying lengths. When the glycosyl intermediate is transferred to the 4-hydroxyl of its own non-reducing end cyclic products are formed (Figure 10).

Structurally, these enzymes share the $(\beta/\alpha)_8$ -barrel protein fold in which the catalytic center lies in the loops joining some of the eight beta-alpha units. Most amylolytic enzymes function as monomers, utilizing a single active site. But there have been recent reports where enzymes belonging to this family show different specificities, like glucanotransferase and maltogenic amylase, etc. which are functional in the form of a dimer which, in turn, leads to the formation of allosteric or effector sites in addition to a main catalytic center. Binding of different substrate molecules at these sites leads to versatile transglycosylation activities in these enzymes.



A



B

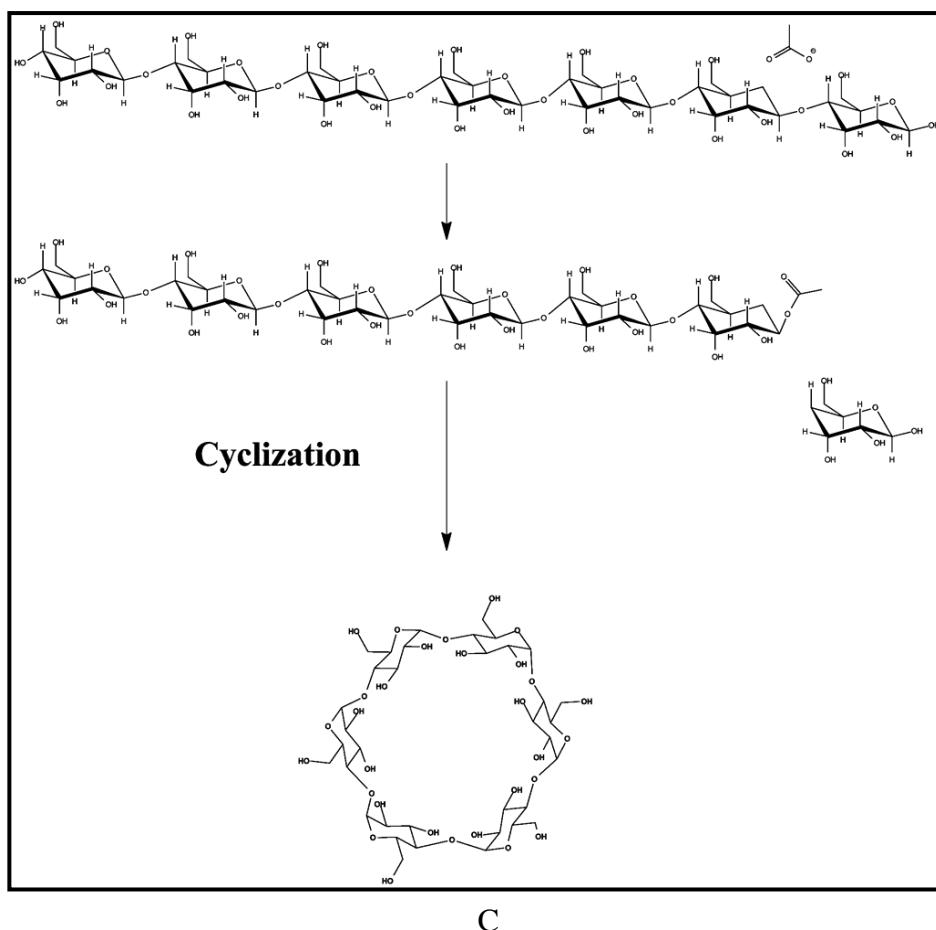


Figure 10: A: Schematic representation of hydrolysis.

B: Schematic representation of transglycosylation.

C: Schematic representation of cyclization reaction.

1.6 4- α -Glucanotransferase from *Pyrococcus furiosus*:

The protein dealt in this thesis is the NCBI's PF0272. This protein was initially cloned and recombinantly produced by Laderman *et al* in 1993. Its activity was then characterized as that of an α -amylase. Various characteristics like optimum temperature and pH of activity and thermal stability were also determined. This protein was then re-characterized as a 4- α -glucanotransferase by Han-Seung Lee *et al* in 2005. In all further literature, this protein is referred to as a glucanotransferase and its glucanotransferase activity has been demonstrated, but nothing explicit about its donor and/or acceptor sites and the mechanism of action has been said, and also nothing is said about its amylase activity other than to say that it has not been seen again. We have called this enzyme PfuAmyGT, in the

anticipation of discovering how it can be both an amylase and a glucanotransferase.

1.7 The multi-domain architecture of PfuAmyGT:

Since the crystal structure of this protein is not known, its structure was predicted by us using the program/server known as I-TASSER, essentially through its homology with a related enzyme from *Thermococcus litoralis* for which the structure has been determined. The domain architecture was determined using NCBI's Conserved Domain Database (CDD). This search showed two domains: An N-terminal catalytic domain of 4- α -glucanotransferase; glycoside hydrolase family 57 (GH57) from residue 13 to residue 289 and α -amylase/ α -mannosidase, GH57 family from 11 to 635, whereas, upon conducting a Pfam search on the UniProt ID P49067, we obtained a three-domain structure: from residue 8 to 292 is marked as belonging to family GH57, residues 312 to 385 as belonging to Domain of Unknown Function 1925 (DUF 1925) which adopts an immunoglobulin/albumin-binding domain-like fold and residues 393 to 644 belonging to DUF 1926 which adopt a beta-sandwich fold and have a proposed role in transglycosylation.

1.8 The multi-domain architecture of an enzyme from *Thermococcus litoralis* which is homologous to PfuAmyGT:

4- α -glucanotransferase from *Thermococcus litoralis* also belongs to family GH57 and catalyzes a disproportionation reaction on amylose to form large cyclic α -1,4-glucans. The crystal structure of this enzyme contains two domains: N-terminal domain contains a $(\beta/\alpha)_7$ barrel fold and C-terminal domain consist of a β -sandwich fold.

The active center is found in domain I and acarbose, an inhibitor (a non hydrolyzable maltotetraose analog) was found bound in the crystal structure, in one of the two chains of the dimeric enzyme, chain A. However, the authors could not explain why acarbose was not bound to the active site in chain B. Also, the structure contained a bound maltose, and the authors could not explain why maltose was found bound to a site far away from the active site, and that too in only chain B, but not in chain A, in the crystal structure (Imamura *et al* 2003).

Based on the work carried out in this thesis, we surmise that the enzyme picked up the maltose from the cytoplasm of the organism in which the enzyme was produced. We surmise that it contains a split active site. Imamura and coworkers would appear to have missed the possibility that the maltose and acarbose are actually bound to the two parts of this split active site because they found the acarbose bound to one chain and the maltose bound to the other chain. We transposed the maltose to the corresponding region of the first chain and discovered that the acarbose and the maltose actually face each other and are separated by a loop (whose role in transferring glucans we have established through mutagenesis).

1.9 The issues addressed in this thesis with PfuAmyGT and TonAmyGT:

1.9.1 The nature and extent of disproportionation:

In this thesis, we have explored the different types of substrates utilized by both the enzymes as well as the extent of utilization. The products formed by the reactions were visualized using thin layer chromatography.

1.9.2 The smallest donors and acceptors:

After carrying out a series of extensive and systematic reactions we have also examined and concluded the smallest donor and acceptor molecules for both PfuAmyGT (from *Pyrococcus furiosus*), and TonAmyGT (from *Thermococcus onnurineus*), again by using thin layer chromatography.

1.9.3 The likely mechanism of glucano-transfer:

For PfuAmyGT, we propose and validate our hypothesis of glucan transfer in the reaction. In many crystal structures of glycosyltransferases (mainly of eukaryotes), there has been found a loop region in the close vicinity of the sugar-nucleotide binding site which helps in the catalytic function. When the substrate molecule is bound, the loop takes up closed conformation covering the substrate to reduce its accessibility to the solvent while the acceptor site becomes more accessible to the solvent. After the catalysis, the loop comes back to its original open conformation.

Taking a cue from this mechanism we have identified a loop region in PfyAmyGT and carried out mutational studies to identify the residues helping in the transfer of the glucan moiety from the donor to the acceptor molecule.

1.9.4 The role of individual domains:

Since three domains were identified in both the enzymes. In the case of PfuAmyGT, we created truncated versions of the protein i.e. cloned and expressed each domain individually and in combinations to find out the function of each domain.

1.9.5 The binding sites for donor and acceptor molecules and their alignment:

PyMOL software was used to identify the putative donor and acceptor sites in PfuAmyGT as well as propose the mechanism of glucan transfer.

1.9.6 The structural-biochemical arrangement for glucan transfer:

Using bioinformatic tools, mainly PyMOL and the crystal structure available from *Thermococcus litoralis* we tried to determine the structural principles of glucose transfer from the donor to acceptor molecule for PfuAmyGT. After identifying the key residues involved in the transfer we mutated three of them (single point mutations) to alanine to examine the effects of the mutations upon activity and established the proposed mechanism.

1.9.7 The temperature and pH optima of activity:

The activity for both the enzymes was also characterized in terms of temperature, time and pH dependence.

1.9.8 The versatility of substrate acceptance:

The enzymes were also examined for both hydrolysis and transfer activities by using different substrate molecules like starch, pectin, cellulose, and xylan. Different combinations of donor molecules in varying concentrations were used to examine activity.

1.9.9 The hyperthermal stability of individual domains:

The stabilities of both the enzymes were examined in terms of thermal as well as chemical stabilities (urea and guanidinium chloride). Similarly, the stabilities of the individual domains of PfuAmyGT were also studied. All of these were found to be thermally stable and reasonably chemically stable.

1.9.10 Domain complementation attempts after independent expression:

The activity of an enzyme is considered to be a sum of the activities of its individual domains for multi-domain enzyme assembly. Since PfuAmyGT shows three individual domains we tried to clone these domains individually and check for activities using TLC.

References:

- Alquéres S.M.C., Almeida R.V., Clementino M.M., Vieira R.P., Almeida, W.I., Cardoso A.M., Martins, O.B. Exploring the Biotechnological Applications in the Archaeal Domain. 2007. *Brazilian Journal of Microbiology*, 38, 398-405.
- Athel Cornish-Bowden. Current IUBMB recommendations on enzyme nomenclature and kinetics. 2014. *Perspectives in Science. 1.* 74–87.
- Christopher Bräsen, Dominik Esser, Bernadette Rauch, Bettina Siebers. Carbohydrate Metabolism in Archaea: Current Insights into Unusual Enzymes and Pathways and Their Regulation. 2014. *Microbiology and Molecular Biology Reviews.* 89-175.
- Chung-wai Chiu and Daniel Solarek. Modification of Starches. 2009 *Starch: Chemistry and Technology, Third Edition Copyright.*
- D. E. Koshland, JR. Stereochemistry and the Mechanism of Enzymatic Reactions. 1953. *Brookhaven National Laboratory*, Upton, N. Y., U.S.A.
- Di Lernia I, Schiraldi C, Generoso M, De Rosa M. Trehalose production at high temperature exploiting an immobilized cell bioreactor. 2002. *Extremophiles.* 6(4), 341-347.
- Gideon Davies and Bernard Henrissat. Structures and mechanisms of glycosyl hydrolases. 1995. *Structure.* 3. 853-859.
- Gideon Davies and Spencer Williams. Carbohydrate-active enzymes: sequences, shapes, contortions, and cells. 2016. *Biochem. Soc. Trans.* 44. 79-87.
- H. N. Cheng and Qu-Ming Gu. Enzyme-Catalyzed Modifications of Polysaccharides and Poly(ethylene glycol). *Polymer.* 2012, 4, 1311-1330.
- Hiromi Imamura, Shinya Fushinobu, Masaki Yamamoto, Takashi Kumasaka, Beong-Sam Jeon, Takayoshi Wakagi, and Hiroshi Matsuzawa. Crystal Structures of 4 α Glucanotransferase from *Thermococcus litoralis* and Its Complex with an Inhibitor. 2003. *The Journal of Biological Chemistry.* 278 (21). 19378-19386.
- International Union of Biochemistry, 1961. Report of the Commission on Enzymes. Pergamon Press, Oxford.
- James N. Bemiller. One Hundred Years of Commercial Food Carbohydrates in the United States. 2009. *J. Agric. Food Chem.,* 57, 8125–8129.

- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The Carbohydrate-active enzymes database (CAZy). 2013. *Nucleic Acids Res.* 42. 490–495.
- Matti Leisola, Jouni Jokela, Ossi Pastinen, Ossi Turunen and Hans E. Schoemaker. Industrial use of enzymes. *Encyclopedia of Life Support System.*
- McCarter JD and Withers SG. Mechanisms of enzymatic glycoside hydrolysis. 1994 *Curr Opin Struct Biol.* 4(6), 885-92.
- Nasir Ahmad, Sumaira Mehboob, and Naeem Rashid. Starch-processing enzymes – emphasis on thermostable 4- α -glucanotransferases. 2015. *Biologia.* 709-725.
- Neelam Gurung, Sumanta Ray, Sutapa Bose, and Vivek Rai. A Broader View: Microbial Enzymes and Their Relevance in Industries, Medicine, and Beyond. 2013. *BioMed Research International.* 18 pages.
- Pedro Fernandes. Enzymes in Food Processing: A Condensed Overview on Strategies for Better Biocatalysts. 2010. *Enzyme Research.* 19 pages.
- Svensson B. 1994. *Plant Molecular Biology* 25, 141-157.
- Takeshi Takaha and Steven M. Smith The Functions of 4- α -glucanotransferases and their use for the Production of Cyclic Glucans.1999. *Biotechnology and Genetic Engineering Reviews*, 16:1, 257-280,
- Zheng Zhang, Ophir Ortiz, Ritu Goyal, and Joachim Kohn. Biodegradable Polymers. 2014. *Principles of Tissue Engineering.* 441–473.

2. Materials and Methods

Materials and Methods

This chapter describes the general material and methods used in the study undertaken in this thesis.

2.1 Cloning, Expression and Purification:

In this section we describe the material and methods involved in cloning our desired amplicons in suitable vectors, checking their expression and purifying proteins.

2.1.1 Materials

2.1.1.1 Bacterial strains:

The following table mentions the cell types used for the purpose of cloning and protein expression:

Cell type	Genotype
<i>E. coli</i> DH5 α	$F^- \phi 80lacZ\Delta M15 \Delta(lacZYA-argF)U169 recA1 endA1 hsdR17(rK^-, mK^+) phoA supE44 \lambda^- thi-1 gyrA96 relA1$ <ul style="list-style-type: none">• Used for general cloning and storage of plasmids.• Supports blue/white screening.• Provides increased insert stability.
<i>E. coli</i> XL1-Blue	$F' ::Tn10 proA^+ B^+ lacI^f \Delta(lacZ)M15/ recA1 endA1 gyrA96 (Nal^R) thi hsdR17 (rK^- mK^+) glnV44 relA1 lac$ <ul style="list-style-type: none">• Tetracycline resistant.• Provides blue/white selection.
<i>Top10</i>	$F^- mcrA \Delta(mrr-hsdRMS-mcrBC) \phi 80lacZ\Delta M15 \Delta lacX74 recA1 araD139 \Delta(ara-leu)7697 galU galK \lambda^- rpsL(Str^R) endA1 nupG$ <ul style="list-style-type: none">• Like DH5α it a general cloning plasmid.

Materials and Methods

	<ul style="list-style-type: none"> • Provides blue/white selection.
<i>E. coli</i> BL21(DE 3)	<p>$F^- omp^T hsdS_B (r_B^-, m_B^-) gal dcm$ (DE3)</p> <ul style="list-style-type: none"> • It is a T7 expression strain. • Has IPTG inducible lacUV5 promoter.
<i>E. coli</i> <i>BL21(DE3)pLysS</i>	<p>$F^- omp^T hsdS_B (r_B^-, m_B^-) gal dcm$ (DE3) pLysS(Cam^R)</p> <ul style="list-style-type: none"> • Chloramphenicol resistance. • Has IPTG inducible lacUV5 promoter. • Provides increased stability of the expressed protein. • Provides lower background expression of target genes when not induced.
<i>E. coli</i> <i>BL21Star(DE3)pLysS</i>	<p>$F^- omp^T hsdS_B (r_B^-, m_B^-) gal dcm rne131$ (DE3) pLysS(Cam^R)</p> <ul style="list-style-type: none"> • Provides high mRNA stability resulting in increased protein yield. • Has low background expression in uninduced cells.
<i>E. coli</i> M15	<p>$NaI^s Str^s Rif^s Thi^- Lac^- Ara^- Gal^+ Mtl^- F^- RecA^-$ $Uvr^+ Lon^+$</p> <ul style="list-style-type: none"> • Kanamycin selection. • Contains pREP4 plasmid which encodes lac repressor. • Inducible using IPTG.

2.1.1.2 Plasmids:

The following table highlights the vector systems/plasmids used for cloning as well as the expression of the recombinant proteins.

Materials and Methods

Plasmid	Genotype
pET 23a	<ul style="list-style-type: none"> • T7 promoter • N-terminal T7 tag • C-terminal His tag • pBR322 origin, f1 origin • Ampicillin resistance
pQE-30	<ul style="list-style-type: none"> • T5 promoter • N-terminal His tag • Ampicillin resistance

2.1.1.3 Chemicals and Kits:

Chemicals, kits and various substrates used in the study are mentioned in the. The reagents used were of analytical grade.

Reagents used	Source
Restriction enzymes	New England Biolabs(NEB), USA
Protein molecular weight markers	Thermo Fisher Scientific, USA
DNA ladders	BR Biochem, India
DNA polymerases	New England Biolabs(NEB), USA
T4 DNA ligase	New England Biolabs(NEB), USA
Plasmid isolation kit, Gel extraction kit, PCR purification kit	Qiagen, Hilden, Germany.
Ni-NTA Agarose/superflow	Qiagen, Hilden, Germany.
TLC Silica gel 60 F ₂₅₄	Merck, New Jersey, USA
Glucose	Himedia, India
Maltose	USB, Amersham Life Sciences
Starch	Himedia, India
Maltooligosaccharides	Santa Cruz Biotechnology, USA
Xylan	Sigma-Aldrich, USA
Pectin	Sigma-Aldrich, USA
Carboxy Methyl Cellulose	Sigma-Aldrich, USA

2.1.1.4 Antibiotics:

Ampicillin, Chloramphenicol, Kanamycin, and Tetracycline used in this study were obtained from Sigma Chemicals, USA. A 1000X stock solution for each of them was prepared as follows:

Antibiotic	Amount
Ampicillin	100 mg/ml in water
Chloramphenicol	35 mg/ml in methanol
Kanamycin	25 mg/ml in water
Tetracycline	12.5 mg/ml in 70 % ethanol

These solutions were then filter sterilized through 0.22 micron Millipore filters and the aliquots were stored in -20 °C.

2.1.1.5 Media:

Media used to grow bacterial cultures was prepared using Millipore Elix-3 water and autoclaved at 15 psi pressure for 15 minutes (Sambrook *et al* 1989 and 2001).

Luria-Bertani (LB) medium -

Component	Amount (g/L)
Tryptone	10.00
Yeast extract	5.00
NaCl	5.00

To prepare solid medium for culture plates 2% agar was added to the above-mentioned composition.

2.1.1.6 Oligonucleotide primers:

The primers used in the present work were procured from Integrated DNA Technologies (IDT) Inc., Coralville, IN, USA. To make a stock concentration of 100 μ M, the primers were dissolved in 10mM Tris, pH 8. The final concentration of primers used in all Polymerase Chain Reactions (PCR) was 1 μ M.

2.1.1.7 Buffers and Solutions used in recombinant DNA work:

TAE buffer (50X, pH 8.0) -

Component	Amount
Tris base	242.00 g
Glacial acetic acid	57.10 mL
EDTA(0.5M pH-8.0)	100.00 mL
Deionized water (DW)	Up to 1000 mL

6X DNA gel loading buffer (prepared in deionized water)-

Component	Amount
Bromophenol blue	0.25%
Glycerol	30%

1% w/v Ethidium Bromide stock solution-

Component	Amount
Ethidium bromide	0.1 g
Deionized water	10 mL

Buffer for preparation of chemically competent cells-

Component	Amount
Calcium chloride	60 mM
Glycerol	15% v/v
PIPES	10 mM

After adjusting the pH of the solution to 7.2 with NaOH it was filter sterilized using a 0.22-micron filter and was followed by autoclaving and stored at 4 °C.

Materials and Methods

Agarose gel (1%) -

Component	Amount
Agarose	0.50 g
DW	50.00 mL

2.1.1.8 Buffers and Solutions used for protein detection and analysis:

Tris-glycine buffer (Tank buffer) -

Component	Amount
Tris base	3.00 g
Glycine	14.40 g
SDS	1.00 g
DW	Upto 1000mL

Stacking gel (5%) -

Component	Amount (μ L)
Tris buffer (0.5M pH=6.8)	950.00
Bis acrylamide solution (30%)	490.00
SDS(10%)	37.50
APS(10%)	18.00
TEMED	10.00
DW (autoclaved)	1000.00

Materials and Methods

Resolving gel (12 %) -

Component	Amount (µL)
Tris buffer (1.5M pH=8.8)	1500.00
Bis acrylamide solution (30%)	2500.00
SDS (10%)	60.00
APS (10%)	30.00
TEMED	10.00
DW (autoclaved)	1900.00

APS (10%) -

Component	Amount
APS	100.00 mg
DW	Up to 1ml

Staining solution -

Component	Amount
Coomassie brilliant blue R250	0.25 g
Methanol	45.00 mL
Glacial acetic acid	10.00 mL
DW	45.00 mL

Destaining solution –

Component	Amount
Methanol	45.00 mL
Glacial acetic acid	10.00 mL
DW	45.00 mL

Materials and Methods

Bisacrylamide solution (30%) –

Component	Amount
Acrylamide	29.20 g
N,N-methylene-bis-acrylamide	0.80 g
DW	Up to 100 mL

Filtered using 2 μ m syringe filter and stored at 4 °C.

Lower Tris (4X), pH 8.8-

Component	Amount
Tris	18.17 g
10 % SDS	4 mL

Volume was made up to 100 mL with deionized water and pH was adjusted using 6N HCl.

Upper Tris (4X), pH 6.8-

Component	Amount
Tris	6.06 g
10% SDS	4 mL

Volume was made up to 100 mL with deionized water and pH was adjusted using 6N HCl.

Materials and Methods

5X Sample loading buffer (prepared in deionized water)-

Component	Amount
Tris.Cl pH 6.8	0.15 M
SDS	5 %
Glycerol	25 %
B-mercaptoethanol	12.5 %
Bromophenol blue	0.06 %

2.1.1.9 Buffers and Solutions used for protein purification:

Native/ Non-denaturing 6X His-tagged proteins:

Lysis buffer-

Component	Amount
NaH ₂ PO ₄ pH 8.0	0.05 M
NaCl	0.3 M
Imidazole	0.01 M

Wash buffer-

Component	Amount
NaH ₂ PO ₄ pH 8.0	0.05 M
NaCl	0.3 M
Imidazole	0.02 M

Elution buffer-

Component	Amount
NaH ₂ PO ₄ pH 8.0	0.05 M
NaCl	0.3 M
Imidazole	0.25 M

Denaturing 6X His-tagged proteins:

Lysis Buffer B-

Component	Amount
Urea/GdmHCl	8 M/6 M
NaH ₂ PO ₄	0.1 M
Tris-Cl	0.01 M

pH was adjusted to 8.0 with NaOH.

Wash Buffer C-

Component	Amount
Urea/GdmHCl	8 M/6 M
NaH ₂ PO ₄	0.1 M
Tris-Cl	0.01 M

pH was adjusted to 6.3 with HCl.

Elution Buffer D-

Component	Amount
Urea/GdmHCl	8M/6M
NaH ₂ PO ₄	0.1 M
Tris-Cl	0.01 M

pH was adjusted to 5.9 with HCl.

Elution Buffer E-

Component	Amount
Urea/GdmHCl	8 M/6 M
NaH ₂ PO ₄	0.1 M
Tris-Cl	0.01 M

pH was adjusted to 4.5 with HCl.

2.2 Methods:

2.2.1 Cloning:

Cloning is a method of producing multiple identical copies of the desired gene by inserting it into a vector and propagating it in a host organism. Conventional cloning includes the following steps:

2.2.1.1 Gene amplification:

The desired gene is amplified either from genomic DNA of an organism or any other template source using a set of specific primers (short DNA stretches containing an overhang region, desired restriction site and complementary region with a free 3' end for extension) using a technique called Polymerase Chain Reaction (PCR). The amplified product is also referred to as an insert.

2.2.1.1.1 Polymerase Chain Reaction:

It is a procedure of amplifying a DNA sequence in vivo using a set of heat stable DNA polymerase, deoxyribonucleotides; dATP, dCTP, dGTP, dTTP and primers. In this study, 1-2 ng of *Thermococcus onnurineus* genomic DNA or pET-23a containing PfuAmyGT gene were used as a template. The reaction mixture contained 500 µM of each dNTP, 1.25 µM of each primer, 2-4 mM MgSO₄ (if mentioned), 1X reaction buffer and 0.5 units of Vent (NEB)/ Deep Vent (NEB)/ Phusion (Thermo Scientific) DNA polymerase. The final volume was made up to 25 µL with PCR grade water. The reactions were carried out using a Mastercycler gradient and vapoprotect from Eppendorf, USA. The programme used were:

S.No.	Step	Deep Vent/ Vent	Phusion	Cycles
1	Initial denaturation	95 °C, 5 minutes	98 °C, 5 minutes	1
2	Denaturation	95 °C, 30 seconds	98 °C, 30 seconds	30
3	Annealing	Depends on the T _m of the primers	Depends on the T _m of the primers	
4	Extension	72 °C, 30sec-	72 °C,	

		1min/kb	30secs/kb	
5	Final extension	72 °C, 5 minutes	72 °C, 5 minutes	1

The PCR product was run on 1 % agarose gel along with DNA ladders for size analysis.

2.2.1.1.2 Splicing by Overlap Extension Polymerase Chain Reaction (SOE-PCR):

SOE-PCR helps to generate single and/or multiple mutations, substitutions, insertions and deletions in a stretch of DNA with the help of oligonucleotide primers (Heckman & Pease, 2007). This method was first described in the 1980s (Ho *et al*, 1989) and since then have been widely used to create desired DNA products. The schematic below (Figure 11) briefly describes the steps followed to carry out point mutations in the desired DNA stretch. The desired mutation is incorporated in the primers. In the first PCR reaction, the desired gene is amplified using mutated forward primer and the wild-type reverse primer. Similarly, in the second PCR reaction, wild-type forward primer and mutated reverse primers are used. In the final (third) PCR reaction the products of the first two reactions are used as templates which contain an overlap complementary region (sticky region) which is further amplified using the wild-type forward as well as wild-type reverse primers. This results in the final point mutated amplicon which can then be digested and ligated into desired vector systems.

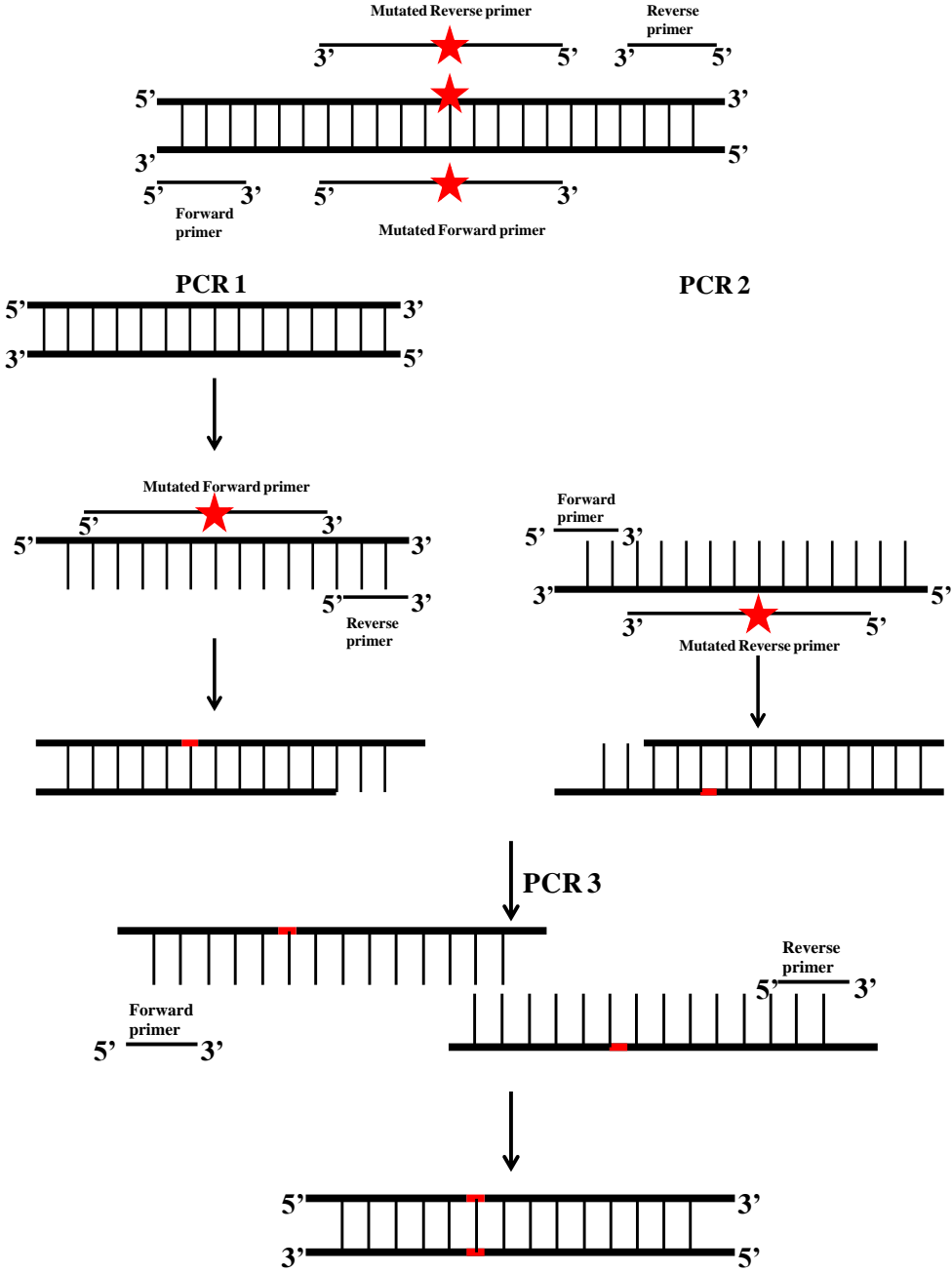


Figure 11: Schematic representation of SOE-PCR.

2.2.1.1.3 Primers used for various constructs:

S.No.	Construct	Primer
1	PfuAmyGT full-length F	AATAATCATATGATCAATGGTTGGACC
2	PfuAmyGT full-length R	AATAATCTCGAGACCAGACGCTTC
3	PfuAmyGT domain 1 F	ATATATCATATGATTAATGGGTGGACGGAAGTG G
4	PfuAmyGT domain 1 R	ATATATCTCGAGCTGCTTTGCTGGTAATGACC
5	PfuAmyGT domain 1 R with a stop codon	ATATATCTCGAGTCACTGCTTTGCTGGTAATGA CC
6	PfuAmyGT domain 2 F	AATATATCATATGTACCGTGTTTTCGTTCGTGG
7	PfuAmyGT domain 2 R	AATATATACTCGAGCAGGTTGTTCCAGATCGC
8	PfuAmyGT domain 3 F	ATACGAAATGGATCCAGTTATGTAAGCCTCGGA AAG
9	PfuAmyGT domain 3 R	AATAAGTAAGCTTCTATCCCGAGGTTTCCTC
10	Pfu D to A F	GCAACGCCGCGTACTG
11	Pfu D to A R	CGTTGCGGCGCATGAC
12	Pfu H to A F	GCGTACTGGGCCGGTCTGTTC
13	Pfu H to A R	CGCATGACCCGGCCAGACAAG
14	Pfu W to A F	GACGCGTACGCGCACGGTCTG
15	Pfu W to A R	CTGCGCATGCGCGTGCCAGAC
16	TonGT F	ATATATGCTAGCATGGTCAATTCATCTTTGG
17	TonGT R	ATATATGCGGCCGCCAACTCCTTAAAGCTGAGC

2.2.1.1.4 Agarose gel electrophoresis:

DNA fragments were separated on 1% agarose gel. The samples were loaded by addition of 6X gel loading buffer to a final concentration of 1X. Electrophoresis was carried in 1X TAE buffer at ~90 volts. For staining, 0.5 µg/mL ethidium bromide was added to agarose. The gels were visualized using a UV transilluminator (Hoefer, USA) and documented using GelDoc EZ Imager (BioRad).

2.2.1.1.5 Purification of DNA fragments from agarose gels:

After agarose gel electrophoresis analysis, the desired DNA fragments were excised out of the gel and purified using gel extraction kit from Qiagen as described below-

1. The gel slice was weighed and QG buffer was added to it, according to the size of DNA fragment: 6 volumes for <100 bp; 3 volumes for 100 bp – 4 kb; 3 volumes with 2 volumes of water for >4 kb.
2. It was then incubated at 50 °C for 10 minutes to solubilize the agarose.
3. To this, one gel volume of isopropanol was added and mixed by inverting the tube.
4. Placed a column in a 2 mL collection tube.
5. To bind DNA, the sample was applied to the column and centrifuged at 6000 ×g for 1 minute.
6. Flow through was discarded and the column was placed back in the same collection tube.
7. To wash, 0.75 mL of buffer PE was added to the column and centrifuged at 6000 ×g for 1 minute.
8. After discarding the flow through centrifugation was carried out at 6000 ×g for 1 minute.
9. Placed column in a clean 1.5mL micro-centrifuge tube.
10. To elute DNA, 50 µL of EB buffer or water was added and centrifugation for 1 minute at 13000 ×g was carried out.

2.2.1.1.6 PCR purification of DNA fragments:

The PCR product obtained from the above steps was also sometimes purified using QIAquick PCR purification from Qiagen. The procedure followed is briefly described below:

1. Five volumes of PB buffer was added to one volume of PCR reaction and mixed.
2. Placed a column in a 2 mL collection tube.
3. To bind DNA, the sample was applied to the column and centrifuged at 6000 \times g for 1 minute. The flow-through was discarded and the column was placed back in the same tube.
4. To wash, 750 μ L PE buffer was added to the column. Centrifugation at 6000 \times g for 1 minute was carried out. The flow-through was discarded.
5. Centrifuged once more to remove residual wash buffer.
6. Placed the column in a clean 1.5 mL microcentrifuge tube.
7. To elute DNA, 50 μ L EB buffer or water was added to the center of the column and let it was allowed to stand for 1 minute. Then, centrifugation at 13000 \times g for 1 minute was carried out.

2.2.1.1.7 Quantification of DNA:

The purified DNA fragments were quantified for concentration estimation either by visually comparing them with the bands of DNA ladder (which depict a definite amount of DNA when a definite amount of ladder is loaded) or by placing 1 μ L of DNA on Nanodrop spectrophotometer (Thermo Scientific) after taking a blank using water or Tris. In the latter case, DNA absorbance is measured at 260 nm and an absorbance of 1, the concentration is equal to 50 ng/mL double-stranded DNA. The purity was also confirmed by checking the ratio of OD₂₆₀/OD₂₈₀ which should be around 1.8-2.0.

2.2.1.1.8 Digestion of the vector and insert:

The insert and vector (the vehicle used to insert the gene into a host organism) were digested with restriction enzymes to generate complementary ends. Restriction digestion was carried out using enzymes from either NEB or Thermo Scientific (Fast digest) for 3 hours and 30 minutes respectively at 37 °C. The

reaction mixture contained 1 µL enzyme per µg of plasmid DNA or 200 ng of PCR product. After digestion, the DNA fragments were separated by agarose gel electrophoresis and desired fragments were purified and quantitated.

2.2.1.1.9 Ligation:

In this process, the digested insert and vector are joined together with the help of T4 DNA ligase enzyme (NEB) which catalyzes the formation of the phosphodiester bond between the two by utilizing ATP. A ratio of 3:1 for insert:vector was used for the reaction which also contained 1X buffer (supplied with enzyme) and units of the enzyme. The reaction was carried out at 25 °C for 3 hours or 16 °C for 16-18 hours.

For a particular amount of digested vector (50 ng), the amount of insert (in the ratio of 3:1 respectively) to be used is calculated by the following formula-

$$\frac{50 * \text{size of insert} * 3}{\text{size of vector} * 1}$$

2.2.1.1.10 Transformation:

After ligation, the reaction mixture was transformed into cloning host cells (for which they are made competent) by the procedure described below-

1. Competent cells were thawed by incubating on ice for 10 – 15 minutes.
2. Ligation mixture (usually 10µL) was added to 100 µL competent cells in microcentrifuge tubes.
3. The tubes were incubated on ice for 20-45 minutes.
4. Then, the tubes were transferred to a preheated 42 °C circulating water bath for 90 seconds.
5. Rapidly transferred the tubes on ice and allowed the cells to cool for 1 – 2 minutes.
6. 1000µL of LB medium was then added to each tube and they were transferred to a shaking incubator set at 37 °C for 1 hour.
7. 100µL of transformed competent cells was spread on LB agar plates containing the appropriate antibiotic.
8. Plates were incubated at 37 °C for 12 – 16 hours.

2.2.1.1.11 Preparation of competent cells:

E. coli DH5 α and *E. coli* BL21pLysSstar competent cells were prepared using the CaCl₂ method. The procedure followed is briefly described below:

1. 200 mL LB medium was inoculated (1%) with an overnight culture of cells and incubated at 37 °C for 3 hours or till the O.D₆₀₀ reached 0.3-0.4.
2. The culture was then incubated on ice for 30 minutes.
3. Cells were pelleted by centrifugation at 1500 \times g for 10 minutes.
4. 10 mL of 0.1 M CaCl₂ was added to each tube and pellet was re-suspended with the help of a pipette.
5. Centrifuged at 1500 \times g for 10 minutes at 4 °C.
6. 10 mL of 0.1 M CaCl₂ was added to each tube and the pellet was re-suspended.
7. Tubes were kept on ice for 45 minutes and intermittently shaken every 5 minutes.
8. Centrifuged at 1500 \times g for 5 minutes at 4 °C.
9. The supernatant was discarded and the tubes were kept in ice.
10. Added 1.5 mL of ice-cold 0.1 M CaCl₂. The pellet was suspended in this solution and 100 μ l aliquots were made and stored at -80 °C till further use.

2.2.1.1.12 Confirmation of transformants -

Colonies observed on LB agar plates were confirmed by-

1. Plasmid isolation and comparing it with the respective undigested plasmid on an agarose gel.
2. Amplifying the isolated plasmid using outer primers specific for the vector and/or the gene.
3. Re-digesting the plasmid using restriction enzymes and running it on an agarose gel to observe the desired size of the insert.
4. DNA sequencing.

2.2.2 Expression of recombinant proteins in *E coli*:

The genes cloned in pQE-30 were transformed in XL1-B or M15 competent cells while those cloned in pET-23a were transformed in BL21 Star(DE3)pLysS for expression. Overnight grown cultures were freshly inoculated in LB media in the presence of required antibiotics at 37 °C to OD₆₀₀ of 0.6-0.7. At this point,

induction was given using 1 mM IPTG and cells were further grown for 5-6 hours. For expression check, a small amount of induced culture was harvested and sonicated. The pellet and supernatant were run on SDS PAGE and analyzed for the presence of desired protein.

2.2.3 Glycerol stock preparation:

1500 μ L of overnight grown cultures of desired bacteria (37 °C, 220 rpm) was mixed with 500 μ L of 60 % glycerol to make the final concentration of glycerol to be 15 %. This stock was stored at -80 °C till further use.

2.2.4 Nondenaturing/ Native protein purification:

Induced bacterial cultures were pelleted at 8000 rpm using a centrifuge. This pellet was then resuspended in 1X native lysis buffer and subjected to sonication using Misonix Ultrasonic Liquid processor followed by heating at 80 °C for 15-20 minutes. This solution was then centrifuged at 16000 rpm for 30-60 minutes and the clear supernatant solution was loaded on to a pre-equilibrated (with 1X Native lysis buffer) Ni-NTA column. It was followed by washing and elution with increasing imidazole concentrations. All the fractions collected during purification were analyzed on SDS PAGE.

2.2.5 Refolding of proteins:

2.2.5.1 On-column refolding:

1. After overnight incubation in 8 M urea and centrifugation at 12,000 rpm for 20 minutes, the supernatant was loaded on pre-equilibrated Ni-NTA column.
2. Washing was done with approximately 20 ml of 4 M urea, 5 % glycerol and 20 mM imidazole in native lysis buffer, pH 8.0.
3. Another washing step was carried out using 2 M urea, 5 % glycerol and 20 mM imidazole in native lysis buffer, pH 8.0.
4. Elution was done in 250 mM imidazole and 5 % glycerol in 1X native lysis buffer, pH 8.0.
5. This protein can then be dialyzed in the desired buffer to remove imidazole and its secondary structure was assessed.

2.2.5.2 Refolding by dialysis:

If the protein is eluted in 8 M urea, pH 4.5 refolding is carried out by dialysis where the urea concentration is gradually decreased from 8 M to 4 M, 2 M, 1 M, 0.75 M, 0.25 M and finally the protein is brought in the desired buffer.

2.2.6 Mass Spectrometry:

To identify the eluted protein it was subjected to SYNAPT G2-Si Mass Spectrometer, Waters in MALDI-TOF mode.

2.2.6.1 Matrix used:

For protein samples: CHCA: α -Cyano-4-hydroxycinnamic acid

For carbohydrate samples: DHB: 2, 5-Dihydroxybenzoic acid

2.2.6.2 The sample preparation is described below:

2.2.6.2.1 Peptide Mass Fingerprinting (PMF identification):

For in-solution digestion lyophilized trypsin (Proteomic grade, Sigma) was reconstituted with 20 μ L of 1 mM HCl to make a final concentration of 1 mg/mL. For in-gel digestion trypsin was reconstituted with 100 μ L of 1mM HCl to a final concentration of 0.2 mg/ml.

2.2.6.2.1.1 In- solution digestion:

A ratio of 1:100 to 1:20 (w/w) of protein or peptide to the enzyme is recommended. The substrate was dissolved in 100 mM Ammonium bicarbonate (ABC) pH 8.5. Trypsin was added according to substrate amount, mixed well and incubated at 37 °C for 16-18 hours.

2.2.6.2.1.2 In-gel digestion:

The desired protein band was excised out of a gel and cut into small pieces. The pieces were transferred to un-autoclaved micro-centrifuge tubes.

1. These pieces were then dehydrated with 40-100 μ L of 1 M Acetonitrile (ACN) solution. This step was repeated thrice.

Materials and Methods

2. To destain the gel pieces 40-100 μL of 30 % ACN solution in 25 mM ammonium bicarbonate (ABC) was added. The samples were incubated at 30 $^{\circ}\text{C}$ with mild shaking for 30 minutes. This step was again repeated three times.
3. After removing the above-mentioned solution, the gel pieces were further dehydrated once with 40 μL of 1 M ACN.
4. Then, 50 μL of 25 mM Dithiothreitol (DTT) was added and incubated at 56 $^{\circ}\text{C}$ for 20 minutes.
5. After cooling the samples to room temperature and pipetting off the residual solution, 50 μL of freshly prepared 55 mM iodoacetamide (IDA) in 25 mM ABC was added for alkylation. The samples were alkylated for 20 minutes in dark.
6. After this, the gel pieces were washed with de-ionized water to remove IDA and further dehydrated using 200 μL of 50 % ACN in 25 mM ABC for 5 minutes and then with 100 % ACN for 30 seconds.
7. To these de-stained and dehydrated gel pieces, 50 μL of trypsin was added and incubated at 37 $^{\circ}\text{C}$ for 16-18 hours.
8. After this, the supernatant obtained by centrifugation at 9500 rpm for 1 minute was subjected to MALDI-MS analysis.

2.2.6.2.2 Carbohydrate samples:

To identify the products obtained after enzymatic degradation, the reaction mixture was diluted 100 times using mass spectrometry grade water and mixed with matrix 2, 5-Dihydroxybenzoic acid (DHB) in the ratio of 1:1. The matrix was prepared by dissolving DHB to a concentration of 10g/L in 10% (v/v) ethanol-water solution. 2 μl of the mixture of DHB and sample was spotted on the MALDI plate and it was subjected to SYNAPT G2-Si Mass Spectrometer, Waters in MALDI-TOF mode with the laser source at 355 nm.

2.2.7 Protein concentration estimation:

Absorption spectra in the range of 200-600 nm were collected using a Cary 50 UV-Vis spectrophotometer. Protein concentration was then estimated by comparing the obtained absorbance value at 280 nm with that obtained from ProtParam (ExPASy bioinformatic resource portal).

2.2.8 Gel filtration chromatography:

After Ni-NTA purification, the eluted proteins obtained were subjected to size exclusion chromatography using AKTApurifier system from GE. The column Superdex-200 10/300 GL was equilibrated with 50 mM Tris pH 8.0 (unless otherwise mentioned). The column has a bed volume of 24 mL. 500 μ L sample was loaded onto the column for analysis. The elution profile of the desired protein sample (which depends upon the hydrodynamic radius and molecular weight of the protein) was compared to the elution profiles of standard protein mixtures run on the same column for analysis.

2.2.9 Fluorescence spectroscopy:

Emission spectra of proteins were collected using a Cary Eclipse fluorimeter in the range of 300-400 nm by excited them at 295 nm. The slit width of 5 nm was used for both excitation and emission slits. Scan speed used was 100 nm/min for all spectra collected and an average of 5-10 spectra was done.

2.2.10 Circular Dichroism (CD) spectroscopy:

Far-UV spectra were collected using Chirascan (Applied Photophysics) CD spectrophotometer using 0.1 cm path length CD cuvette. Spectral range was 250-190/200 nm. For single wavelength data, an average of 3-5 scans was done with an integration time of 1 second. Raw ellipticity obtained was then converted to mean residual ellipticity using the following formula-

$$[\Theta] = \frac{\Theta_{\text{observed in mdeg}} \times 100 \times \text{MRW}}{1000 \times \text{protein concentration in } \frac{\text{mg}}{\text{mL}} \times \text{path length of the cuvette in cm}}$$

Mean Residue Weight (MRW)

$$= \frac{\text{Total molecular weight of the protein}}{\text{Total number of amino acids in the protein}}$$

2.2.10.1 Heat denaturation:

For the determination of temperature-induced unfolding transitions, the protein samples (0.15-0.2 mg/ml) were heated from 20 °C to 90 °C with a constant heating rate of 1 ° C/minute in a sealed quartz 0.1 cm cuvette to avoid evaporation.

2.2.10.2 Chemical denaturation:

For determination of chemical-induced unfolding transition using urea and guanidium hydrochloride (GdmHCl), 0.15-0.2 mg/mL of final protein concentration was used along with urea concentrations varying from 0 to 8 M and GdmHCl concentrations varying from 0 to 6 M. Samples containing denaturants were mixed and incubated at room temperature for 16-18 hours.

The samples were then monitored by CD at 222 nm. The decrease in signal strength at 222 nm was directly proportional to the unfolding of the proteins and the corresponding T_m and C_m were determined by plotting and fitting the data with a sigmoidal Boltzmann function using OriginPro 8.5.

2.2.11 Dynamic Light Scattering (DLS):

To determine the hydrodynamic radius of the proteins, they were first filtered through 0.22-micron polyvinylidene difluoride syringe filters (Millipore) and then subjected to Zetasizer Nano ZS90 model instrument (Malvern Instruments). DLS measures the changes in light scattering resulting from molecules in solution with time to determine the translational diffusion coefficient and consequently find the hydrodynamic radius using the Stokes-Einstein equation. The R_H depends on the shape of the molecule and can be determined by comparison with a calibration curve of R_H of known proteins.

2.2.12 Differential Scanning Calorimetry (DSC):

This technique was used to determine the thermal stability of protein molecules. It determines changes in the partial molar heat capacity of the protein at constant pressure. It is calculated by subtracting a buffer baseline done during creating a thermal history for the instrument (MicroCal VP-DSC, GE Healthcare). As the protein is heated over a temperature range, it unfolds to rise to an endothermic peak and when the unfolding is complete the heat absorption decreases and a new

baseline is established. This data was then plotted and fitted using Origin and the T_m was calculated.

2.2.13 Activity Assays:

2.2.13.1 Starch-iodide method:

Originally discovered by Colin and Claubry in 1814, this method is used to test the presence of starch in a given solution both qualitatively and quantitatively. Starch contains both amylose (only α 1-4 linked glucose units) and amylopectin (in addition to linear glucose chains it also consists of branching at α 1-6). Only amylose can produce a helical structure of 8.0 Å pitch with an outer diameter of 13.0 Å. This gives rise to a 5 Å wide cavity with 6 glucose units per turn that can accommodate iodine atoms. The iodine atoms can line up to produce a linear polyiodide chain with an average I-I separation of 3.1 Å. Since molecular iodine is not soluble in water potassium iodide is added which leads to the formation of charge transfer complexes. The negatively charged iodide in these compounds acts as charge donor and the neutral iodine as a charge acceptor. Electrons in this charge-transfer complex are excited to a higher energy level by light and its complementary color is observed.

For the study undertaken here a stock solution of 4% potassium iodide with 1.25% of iodine was prepared in deionized water and stored in a dark bottle.

2.2.13.2 Zymography:

In this assay, hydrolytic activity of an enzyme can be visualized in the form of clearing of the substrate on a PAGE gel. It was introduced in 1962 by Gross and Lapière to observe the degradation of collagen (Vandooren *J et al* 2013).

In our case, we co-polymerized 0.2 % starch with 8 % native (non-SDS containing) PAGE gel as well as 0.2 % starch with 200 mM maltose. We then loaded appropriate amount (increasing concentrations) of PfuAmyGT in it and finally stain the gel with an iodine solution to observe zones of clearance.

2.2.13.3 Thin Layer Chromatography (TLC):

Thin layer chromatography is a form of liquid chromatography which helps to separate a mixture of substances by loading them on to a thin sorbent layer (stationary phase), in this case, silica (which is highly polar) and allowing them to move by capillary action along with the movement of a solvent (mobile phase). The substances constantly move to and fro between the mobile phase and stationary phase and are separated on basis of their polarity with respect to each other, to the stationary phase as well as the mobile phase. The mobile phase used in this study is composed of Butanol:Ethanol: Water in the ratio of 50:30:20 respectively.

2.2.14 Structure analysis tools:

2.2.14.1 Clustal Omega:

It is a multiple sequence alignment tool for analyzing three or more DNA and/or protein sequences provided by EMBL-EBI. The results obtained were further subjected to ESPript 3.0.

2.2.14.2 I-TASSER (Iterative Threading ASSEmbly Refinement):

It is a tool for predicting protein secondary structure. It chooses templates from entries submitted to Protein Data Bank based on fold recognition.

2.2.14.3 TM-align:

It was used for aligning secondary structures of proteins by superimposition. This program returns a TM score between 0 and 1. A score of 0.2 or below is considered to be insignificant because of random matching whereas a score of 0.5 and above is considered significant.

2.2.14.4 PyMOL:

For visualization and analysis of protein secondary structure and locate specific substructures and residues, PyMOL was used.

References:

- Karin L Heckman, Larry R Pease. Gene splicing and mutagenesis by PCR-driven overlap extension. 2007. *Nature Protocols*. 2(4).
- Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. 2010. *Nature Protocols*. 5. 725-738.
- Sambrook J, Fritsch E.F., Maniatis T. Molecular cloning: a laboratory manual. 1989. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
- Sambrook J, Russel D. Molecular cloning: a laboratory manual. 2001. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
- Sievers F, Wilm A, Dineen DG, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. 2011. *Molecular Systems Biology*.
- Steffan N. Hoa, Henry D. Hunt, Robert M. Horton, Jeffrey K. Pullen, Larry R. Pease. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. 1989. *Gene*. 77. 51-59.
- Vandooren J, Geurts N, Martens E, Steen P, Opdenakker G. Zymography methods for visualizing hydrolytic enzymes. 2013. *Nature Methods*. 10 (3). 211-220.
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: Protein structure and function prediction. 2015. *Nature Methods*. 12. 7-8.
- Zhang Y. I-TASSER server for protein 3D structure prediction. 2008. *BMC Bioinformatics*.

3. Results and Discussion

3.1 Physico-chemical characteristics of PfuAmyGT and its domains:

3.1.1 Cloning, expression and purification:

The full-length gene was sub-cloned by amplification of a synthesized gene (optimized for expression in *E. coli* in terms of codon usage) to obtain the desired amplicon of 1972 bp. This amplicon was then inserted into the pET-23a vector system after suitable restriction-digestion, using enzymes *NdeI* and *XhoI*. After transformation in XLI-Blue competent cells, several colonies were screened through the performance of PCR reactions, using T7 forward and reverse primers. To confirm the presence of our desired insert, the plasmid (from selected colonies, showing amplification and production of an amplicon of around 2 kb) was isolated and subjected to digestion with *NdeI* and *XhoI*, as shown in Figure 12 A, B.

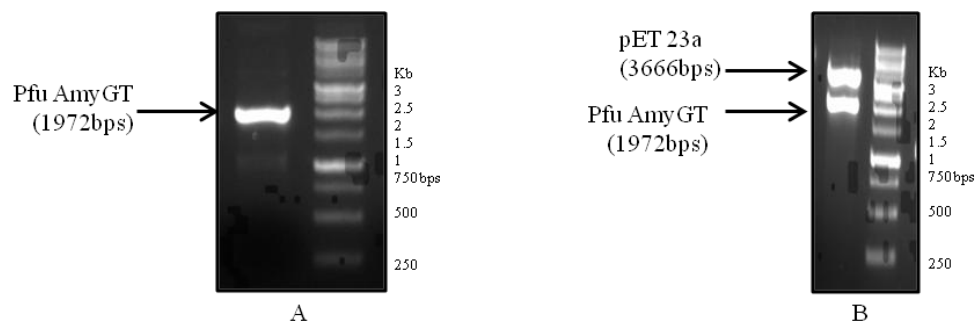


Figure 12: Panel A shows amplified PCR product of 1972 bps. Panel B shows the *NdeI* and *XhoI* digested pET-23a containing PfuAmyGT gene.

Once the synthetic (codon-optimized) gene of PfuAmyGT encoding the full-length wild-type protein sequence was cloned, it was used as a template for cloning and amplification of the protein’s domains as shown in Table 1 (column 2 and 3). These domains were identified with the help of the ‘Pfam’ database (Figure 13) with the initial view that the functions of full-length PfuAmyGT could potentially be attributable to its different domains (which turned out not to be the case, since the split active site of the enzyme was later eventually inferred to lie at domain interfaces). We thought that the individual domains would provide us with meaningful insights regarding the mechanism of the function, and activity, of full-length PfuAmyGT. Therefore, the domains and their boundaries were predicted, using the UniProt ID: P49067 of full-length PfuAmyGT and the

Results and Discussion

gene sequences encoding the domains were cloned for expression and production as separate polypeptides. The enzyme was predicted to have 3 domains: Domain 1- Belonging to family GH 57, starting at amino acid residue number 8 and ending at residue 292; Domain 2 & Domain 3- belong to DUF (Domain of Unknown Function) 1925 and 1926 respectively. We have successfully cloned the different domains (full length, individual domains and also truncated variants) of PfuAmyGT as described in table 1.

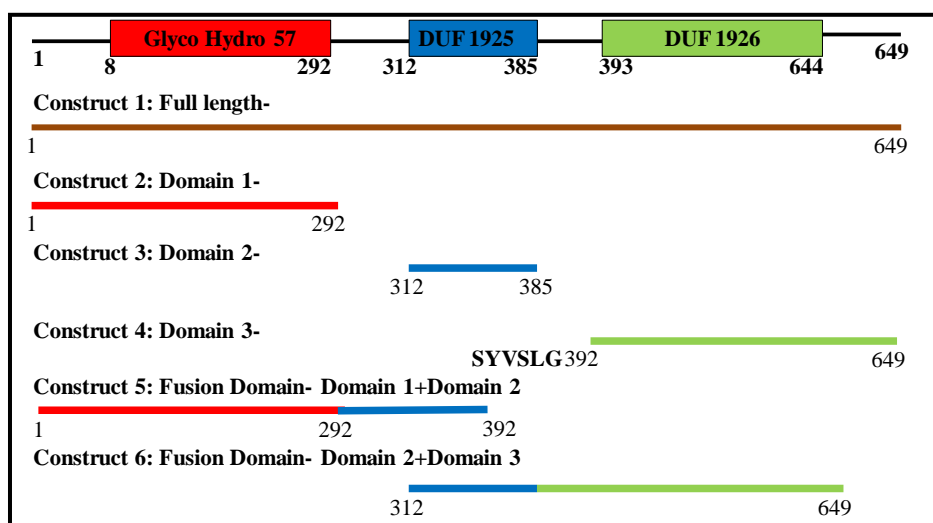


Figure 13: It is a schematic depicting the boundaries of different domains predicted by Pfam database and the truncated variants created by cloning.

Results and Discussion

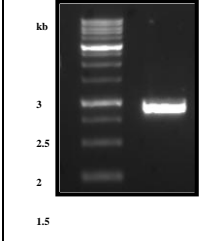
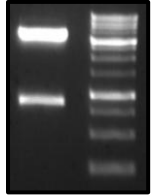
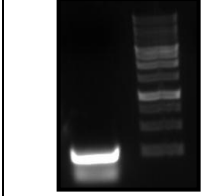
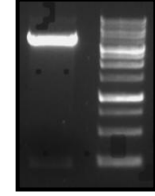
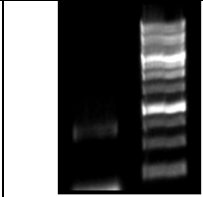
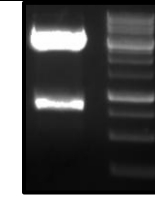
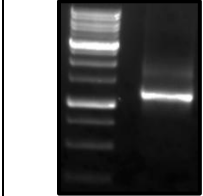
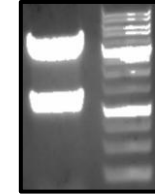
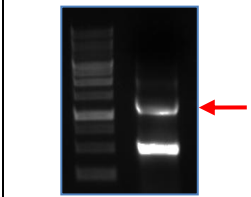
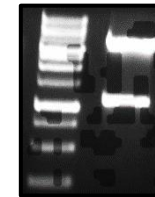
Domain/ Boundary	Amplification	Cloning	Expected size (bps)	Cloning host and restriction sites	Expression host
1 (8-292)			876	pET-23a <i>NdeI</i> and <i>XhoI</i>	BL21 Star(DE3)pLys S
2 (312- 385)			222	pET-23a <i>NdeI</i> and <i>XhoI</i>	BL21 Star(DE3)pLys S
3 (393-644)			795	pQE-30 <i>BamHI</i> and <i>HindIII</i>	pQE-30
1+2			1194	pET-23a <i>NdeI</i> and <i>XhoI</i>	BL21 Star(DE3)pLys S
2+3			1039	pET 23a <i>NdeI</i> and <i>XhoI</i>	BL21 Star(DE3)pLys S

Table 1: Amplification of DNA encoding all of the predicted PfuAmyGT domains, their restriction digestion checks, and the vectors used as cloning and expression hosts.

After confirmation by sequencing, these expression-competent plasmids were transformed into the respective competent cells to check for protein expression. These cells were then grown on a large scale (in secondary cultures) and induced at a culture O.D₆₀₀ of 0.6 with 1 mM IPTG, to express the cloned genes for 6-8 hours of further growth at 37 °C, with shaking at 220 rpm. The culture was then

Results and Discussion

sedimented through centrifugation at 8000 rpm, and cells thus obtained were re-suspended in native lysis buffer. For purifying full-length PfuAmyGT, cells were subjected to heating at 90 °C for 20-30 minutes, followed by sonication. Heating during purification helped to degrade *E. coli*-derived proteases, precipitate and sediment most other proteins of *E. coli* origin, and help the hyperthermophilic protein (PfuAmyGT) to fold under near-native temperature conditions, to achieve native conformation (i.e., the protease-susceptible regions of the protein became buried through folding at high temperatures, since the enzyme evolved in a hyperthermophile, and the heating simultaneously prevented degradation by host proteases). The lysate obtained after heating was then centrifuged at 16,000 rpm and the supernatant was subjected to Ni-NTA affinity chromatography. The protein is expressed along with the 6X-His affinity tag, derived from the vector after insertion of the gene into the vector's multiple cloning sites. The fractions collected during purification were analyzed on SDS PAGE (Figure 14).

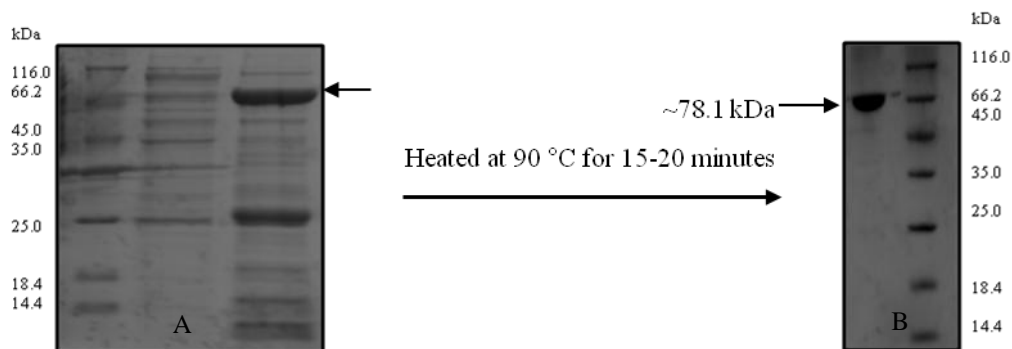


Figure 14: Panel A shows SDS PAGE run after Ni-NTA purification of PfuAmyGT. Panel B shows SDS PAGE run after Ni-NTA purification carried out after heating the lysate at 90 °C for 20 minutes which shows a pure, single band of approximately 78 kDa.

The domains of PfuAmyGT 1 and 3 were also purified in the same manner. For domain 1+2, 0.2 % sarkosyl was added to native lysis buffer during re-suspension of the cell pellet. Sarkosyl which also known as sarcosyl or sodium lauroyl sarcosinate, is an anionic surfactant that helps in solubilizing proteins from inclusion bodies by encapsulating and disrupting protein aggregates (Tao *et al*, 2010). These proteins were then purified by Ni-NTA purification using a standard protocol. For domain 2+3, 1 % glucose was added at the time of induction along with IPTG, which helps to repress induction of the *lac* promoter by lactose. For domain 2, purification was carried out under denaturing (8 M urea) conditions.

Results and Discussion

Since all the proteins possess a 6X-His affinity tag (from their respective vectors) they were finally purified by Ni-NTA affinity chromatography and analyzed by SDS-PAGE (Figure 15). The general properties and purification conditions are summarized in table 2.

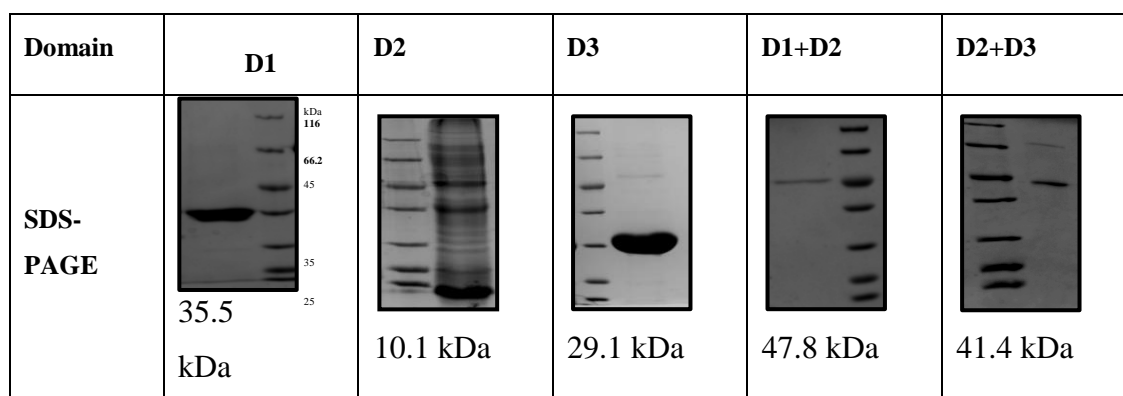


Figure 15: SDS-PAGE analysis of purified domains of PfuAmyGT each of which shows ~95 % pure protein band. The molecular weight marker is same for all gels.

Constructs	No. of aa	Protein size (Da)	pI	No. of W	No. of C	M ⁻¹ cm ⁻¹	Purification conditions
PfuAmyGT full length (Wild type)	664	78167.58	5.58	18	2	155140	Non denaturing + heating at 90 °C for 20 minutes
D 1	300	35561.39	5.0	12	1	92720	Non denaturing
D 2	82	10132.73	10.14	3	1	26150	Denaturing
D 3	252	29143.63	4.95	3	0	34990	Non denaturing
D 1 + 2	403	47892.66	6.05	15	2	120150	Non denaturing + 0.2 % sarcosyl
D 2 + 3	353	41440.70	6.76	6	1	62420	Non denaturing +1 % glucose

Table 2: Some properties and purification conditions of PfuAmyGT full-length and its domains.

After obtaining the bands of the correct size(s) on SDS-PAGE, refolding was attempted for PfuAmyGT domain 2 by both the dialysis method as well as through on-column refolding (as mentioned in Material & Methods) but we were unable to refold it, perhaps because it is too small and rich in hydrophobic residues to refold independently into native structure while exposing large

hydrophobic surfaces (which are otherwise engaged in interacting with the domains 1 and 3). Therefore, we gave up on attempts to make domain 2 independently. However, as the figures above show, we made domains 1 and 3 independently and also fusions of domain 2 with both domain 1 and domain 3.

3.2 Confirmation of the chemical identities of all domain constructs by mass spectrometry:

Samples were prepared for Peptide Mass Fingerprinting (PMF) of the purified protein (after excision from SDS-PAGE, in the form of gel bands) as discussed in the Materials and Methods chapter. The PMF data were obtained and analyzed and found to match with expectation. *De novo* sequencing was also done for one of the several peptide fragments of size matched with expectation, as shown in Figure 16. Panel A shows the *in silico* digested tryptic peptides (for PfuAmyGT full-length, wild-type sequence) using the ExPASy tool of which the ones marked in red were observed in the mass spectrum (MS) as shown with a red arrow in Figure 16. Panel B, in the same figure, shows the spectrum obtained and sequence of the peptide obtained after MS/MS of a peptide of molecular mass 1103 Da.

Similarly, all the domains were confirmed by this process as shown in Figure 17.

Results and Discussion

A

Mass	Position	Sequence			
1774.8475	226-239	FGIWPGTYEWWYEK	678.3304	249-254	ISSDEK
1752.8802	546-560	LVNDGFVEYIVNNK	658.3882	107-111	IEQIR
1717.8643	201-215	VLEYLHSLIDGDESK	639.2620	651-656	FEEASG
1529.7920	255-266	INLMLYTEYLEK	593.3293	307-311	GIFEK
1481.7634	518-530	VPYSYELLDGGIR	564.3504	354-357	YLLR
1249.5881	1-11	MINGWTEVGDK	563.3300	267-270	YKPR
1247.7146	191-200	YLIPFRPVDK	560.3191	318-322	GGIWK
1237.6827	295-304	LFVEFVNELK	537.3031	596-600	YAVGK
1212.5241	605-613	FEDEMEVWK	533.2718	115-118	EWAK
1168.5092	328-336	YPESNYMHK	531.3038	240-243	GWLR
1161.6626	437-446	LVNYVDVLAR	520.3242	314-317	VFVR
1116.5320	216-225	VAVFHDDGEK	506.2973	614-617	YPVK
1103.5480	509-517	QQEIGEFPR	504.2664	641-644	LEDK
1053.5575	534-542	EHLGIEVEK			
998.5669	133-140	VWQPELVK			
971.5672	380-387	AIWNNLIK			
938.4941	388-396	ANSYVSLGK			
860.4624	126-132	GVWLTER			
817.4301	589-595	EIVVDDK			
765.3889	119-125	SIGFDAR			
742.4093	472-477	IPDEIR			
724.3512	479-484	EVAYDK			
718.3559	323-327	NFFYK			
713.3253	244-248	EFFDR			
708.3783	338-343	MLMVSK			
700.3373	347-352	NNPEAR			

B

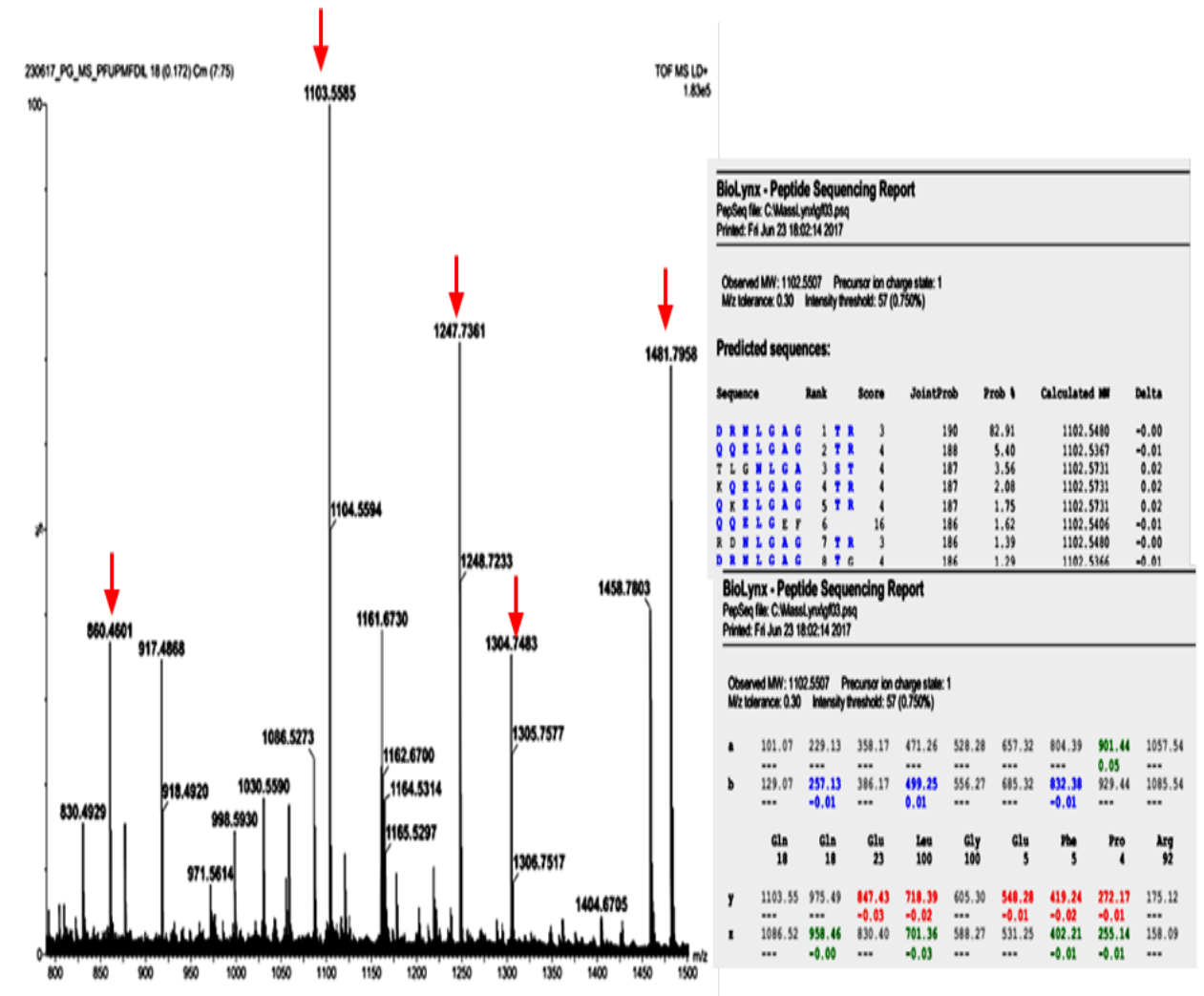
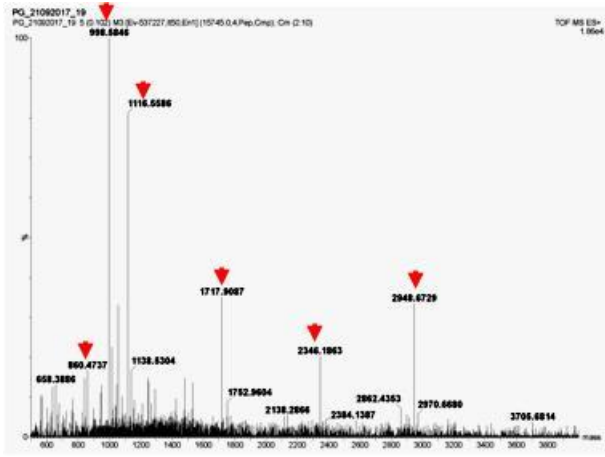
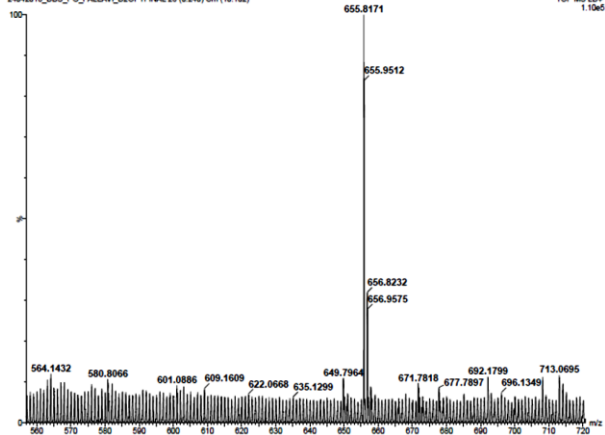
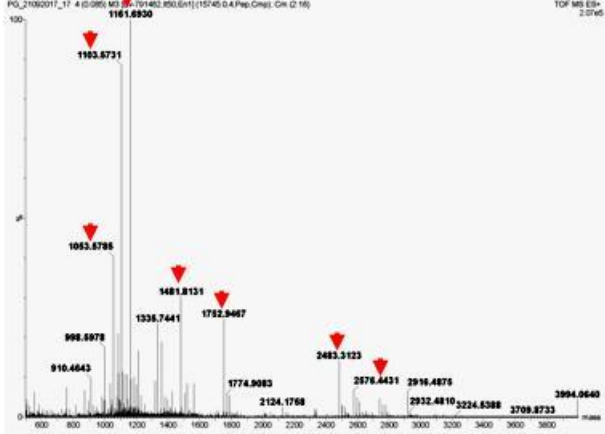


Figure 16: Panel A shows the *in silico* (tryptic) digested peptides. Panel B shows the mass spectrum obtained (PMF) and *de novo* sequence of mass peak 1103 by MS/MS.

Domains	MS spectrum (and matched peptide)																																				
D1	 <p>MS spectrum for domain D1. The x-axis represents m/z from 600 to 3800, and the y-axis represents relative intensity from 0 to 100. The base peak is at m/z 998.8846. Other significant peaks are marked with red triangles and labeled with their m/z values.</p> <table border="1"> <thead> <tr> <th>m/z</th> <th>Relative Intensity (approx)</th> </tr> </thead> <tbody> <tr><td>658.3886</td><td>5</td></tr> <tr><td>860.4737</td><td>10</td></tr> <tr><td>998.8846</td><td>100</td></tr> <tr><td>1116.6586</td><td>85</td></tr> <tr><td>1138.8304</td><td>15</td></tr> <tr><td>1717.9087</td><td>40</td></tr> <tr><td>1752.9604</td><td>10</td></tr> <tr><td>2138.2666</td><td>5</td></tr> <tr><td>2346.1963</td><td>25</td></tr> <tr><td>2384.1387</td><td>5</td></tr> <tr><td>2862.4353</td><td>10</td></tr> <tr><td>2948.6729</td><td>45</td></tr> <tr><td>2970.6690</td><td>5</td></tr> <tr><td>3706.6814</td><td>5</td></tr> </tbody> </table>	m/z	Relative Intensity (approx)	658.3886	5	860.4737	10	998.8846	100	1116.6586	85	1138.8304	15	1717.9087	40	1752.9604	10	2138.2666	5	2346.1963	25	2384.1387	5	2862.4353	10	2948.6729	45	2970.6690	5	3706.6814	5						
m/z	Relative Intensity (approx)																																				
658.3886	5																																				
860.4737	10																																				
998.8846	100																																				
1116.6586	85																																				
1138.8304	15																																				
1717.9087	40																																				
1752.9604	10																																				
2138.2666	5																																				
2346.1963	25																																				
2384.1387	5																																				
2862.4353	10																																				
2948.6729	45																																				
2970.6690	5																																				
3706.6814	5																																				
D2	 <p>MS spectrum for domain D2. The x-axis represents m/z from 560 to 720, and the y-axis represents relative intensity from 0 to 100. The base peak is at m/z 655.8171. Other significant peaks are marked with red triangles and labeled with their m/z values.</p> <table border="1"> <thead> <tr> <th>m/z</th> <th>Relative Intensity (approx)</th> </tr> </thead> <tbody> <tr><td>564.1432</td><td>5</td></tr> <tr><td>580.8066</td><td>5</td></tr> <tr><td>601.0886</td><td>5</td></tr> <tr><td>609.1609</td><td>5</td></tr> <tr><td>622.0668</td><td>5</td></tr> <tr><td>635.1299</td><td>5</td></tr> <tr><td>649.7964</td><td>5</td></tr> <tr><td>655.8171</td><td>100</td></tr> <tr><td>655.9512</td><td>85</td></tr> <tr><td>656.8232</td><td>10</td></tr> <tr><td>656.9575</td><td>10</td></tr> <tr><td>671.7818</td><td>5</td></tr> <tr><td>677.7897</td><td>5</td></tr> <tr><td>692.1799</td><td>5</td></tr> <tr><td>696.1349</td><td>5</td></tr> <tr><td>713.0695</td><td>5</td></tr> </tbody> </table>	m/z	Relative Intensity (approx)	564.1432	5	580.8066	5	601.0886	5	609.1609	5	622.0668	5	635.1299	5	649.7964	5	655.8171	100	655.9512	85	656.8232	10	656.9575	10	671.7818	5	677.7897	5	692.1799	5	696.1349	5	713.0695	5		
m/z	Relative Intensity (approx)																																				
564.1432	5																																				
580.8066	5																																				
601.0886	5																																				
609.1609	5																																				
622.0668	5																																				
635.1299	5																																				
649.7964	5																																				
655.8171	100																																				
655.9512	85																																				
656.8232	10																																				
656.9575	10																																				
671.7818	5																																				
677.7897	5																																				
692.1799	5																																				
696.1349	5																																				
713.0695	5																																				
D3	 <p>MS spectrum for domain D3. The x-axis represents m/z from 600 to 3800, and the y-axis represents relative intensity from 0 to 100. The base peak is at m/z 1181.6930. Other significant peaks are marked with red triangles and labeled with their m/z values.</p> <table border="1"> <thead> <tr> <th>m/z</th> <th>Relative Intensity (approx)</th> </tr> </thead> <tbody> <tr><td>910.4643</td><td>5</td></tr> <tr><td>998.6878</td><td>10</td></tr> <tr><td>1063.6795</td><td>35</td></tr> <tr><td>1181.6930</td><td>100</td></tr> <tr><td>1193.8731</td><td>85</td></tr> <tr><td>1336.7441</td><td>15</td></tr> <tr><td>1481.8131</td><td>25</td></tr> <tr><td>1752.9487</td><td>20</td></tr> <tr><td>1774.9083</td><td>5</td></tr> <tr><td>2134.1768</td><td>5</td></tr> <tr><td>2483.3123</td><td>15</td></tr> <tr><td>2576.4431</td><td>10</td></tr> <tr><td>2916.4875</td><td>5</td></tr> <tr><td>2932.4810</td><td>5</td></tr> <tr><td>2924.6388</td><td>5</td></tr> <tr><td>3708.8733</td><td>5</td></tr> <tr><td>3994.0640</td><td>5</td></tr> </tbody> </table>	m/z	Relative Intensity (approx)	910.4643	5	998.6878	10	1063.6795	35	1181.6930	100	1193.8731	85	1336.7441	15	1481.8131	25	1752.9487	20	1774.9083	5	2134.1768	5	2483.3123	15	2576.4431	10	2916.4875	5	2932.4810	5	2924.6388	5	3708.8733	5	3994.0640	5
m/z	Relative Intensity (approx)																																				
910.4643	5																																				
998.6878	10																																				
1063.6795	35																																				
1181.6930	100																																				
1193.8731	85																																				
1336.7441	15																																				
1481.8131	25																																				
1752.9487	20																																				
1774.9083	5																																				
2134.1768	5																																				
2483.3123	15																																				
2576.4431	10																																				
2916.4875	5																																				
2932.4810	5																																				
2924.6388	5																																				
3708.8733	5																																				
3994.0640	5																																				

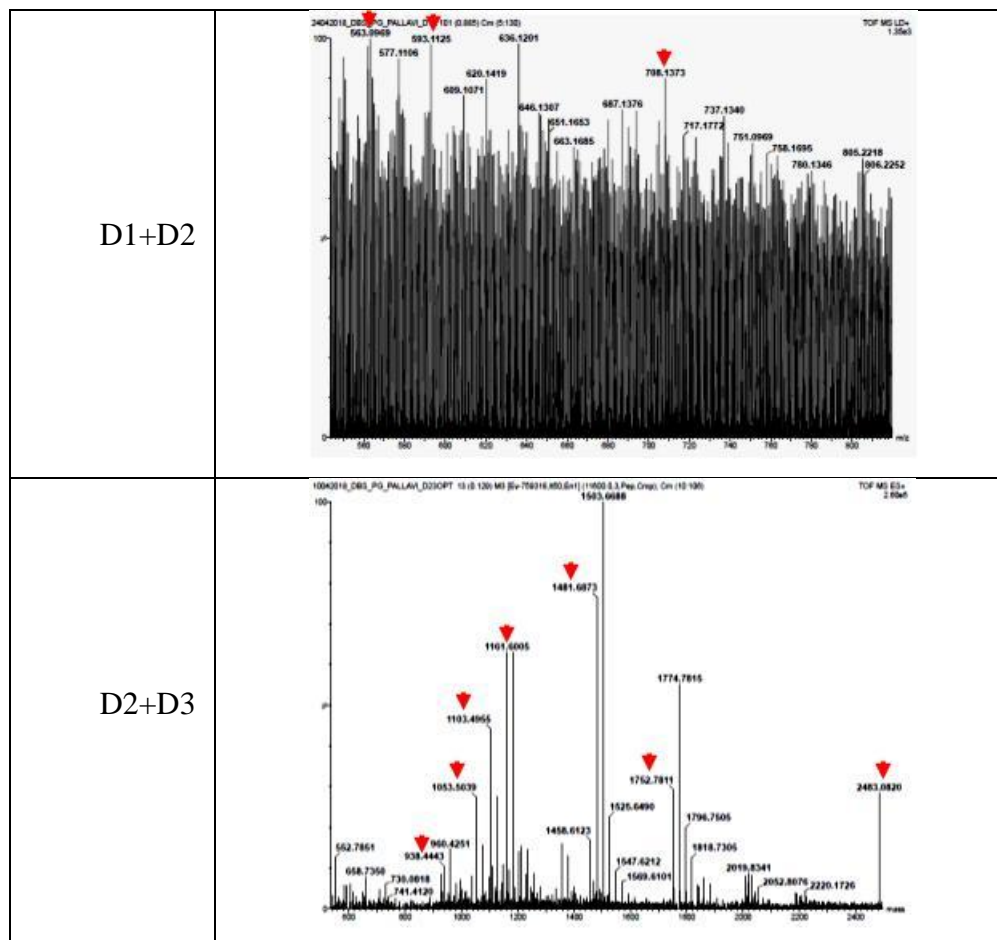


Figure 17: MS spectra and corresponding peaks obtained during the identification of PfuAmyGT domains.

3.3 Secondary, tertiary and quaternary structures of PfuAmyGT and its domains:

3.3.1 Determining the oligomeric status of the proteins:

3.3.1.1 Gel filtration chromatography:

The proteins purified by Imidazole Metal Affinity Chromatography (IMAC) using Ni-NTA resin were subjected to gel filtration chromatography in order to further purify these, in preparative chromatographic mode, as well as to determine their oligomeric status in analytical chromatographic mode. The gel filtration column used, i.e., Superdex-200 Increase 10/300 GL, was first equilibrated with 50 mM Tris, pH-8.0, at a flow rate of 0.5 ml/minute. Then, 500 μ l of each protein was injected. We obtained an elution volume of \sim 11.356 ml for PfuAmyGT (Figure 18). The reported homolog of this protein (4- α -glucanotransferase from *Thermococcus litoralis* with PDB ID's 1K1X and 1K1Y) is said to be dimeric in

Results and Discussion

nature and on comparison with the calibration curve (obtained by running proteins of known molecular weights) we found native PfuAmyGT to be a ~115 kDa protein which corresponds roughly to a dimeric molecular weight, or at any rate to a molecular weight which is well above that of a monomer, although below that of a dimer.

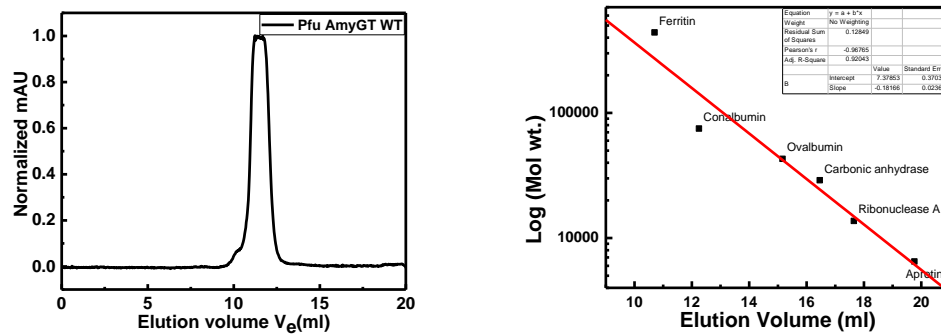


Figure 18: Gel filtration profile of PfuAmyGT on Superdex-200 Increase column with an elution volume of ~11.35 ml. The right-hand side of the figure shows the calibration curve of the column.

As summarized in Figure 19 (column one represents domain name, column two shows the gel filtration profile and column three shows the observed elution volume for each domain), the gel filtration profile was also determined for all of the other domains (individual and truncated versions) except for domain 2.

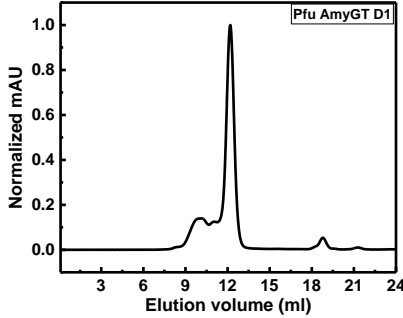
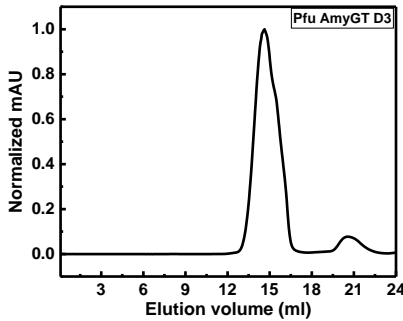
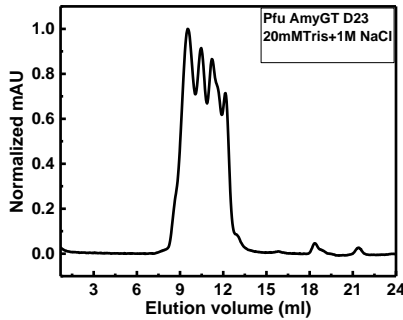
Domain	Gel filtration profile	Observed Elution volume (ml)
D1		12.2
D3		14.6
D2 + D3		No distinct peak observed

Figure 19: Gel filtration peaks of different (individual and truncated version) PfuAmyGT and their respective elution volumes.

3.3.1.2 Dynamic Light Scattering (DLS):

The major population obtained from gel filtration chromatography of full-length PfuAmyGT was subjected to the prediction of hydrodynamic radius by DLS. We observed a hydrodynamic radius of ~5.25 nm (Figure 20), or a diameter of ~10.5 nm. In comparison with the crystal structure of *T. litoralis* homolog (i.e., through PYMOL software-based analysis of the longest dimension of the dimeric form, which is ~11.3 nm), it appears that the determined hydrodynamic radius of PfuAmyGT falls in the expected range of sizes for a dimer of this protein.

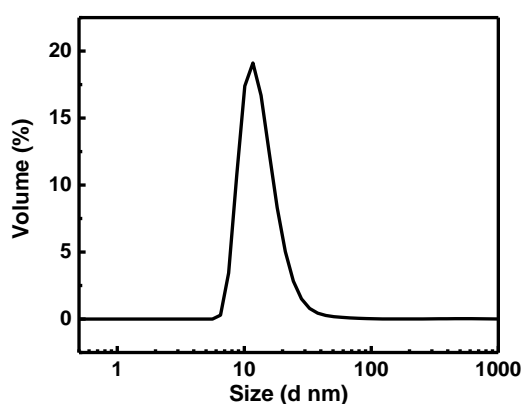


Figure 20: DLS analysis of full-length PfuAmyGT.

3.3.2 Determination of the secondary structure of PfuAmyGT:

To determine the secondary structure of PfuAmyGT, the purified protein was examined using circular dichroism (CD) spectroscopy, with spectrum collected at a protein concentration of 0.15 mg/ml in a 0.1 cm cuvette. The CD spectrum shows two major negative bands at 222 nm and 208 nm, which mainly correspond to the alpha-helical structure. Based on comparison with the *T. litoralis* homolog, the protein is expected to be a mixture of alpha helical and beta sheet structures. When this is the case, the alpha-helical signature still dominates the CD spectrum but the mean residue ellipticity (MRE) is grossly reduced from that expected (-35,000 to -40,000 deg cm² dmol⁻¹) for a pure alpha-helical protein. In keeping with this expectation, the MRE of the negative bands of PfuAmyGT was determined to be approximate -10,000 deg cm² dmol⁻¹ (Figure 21).

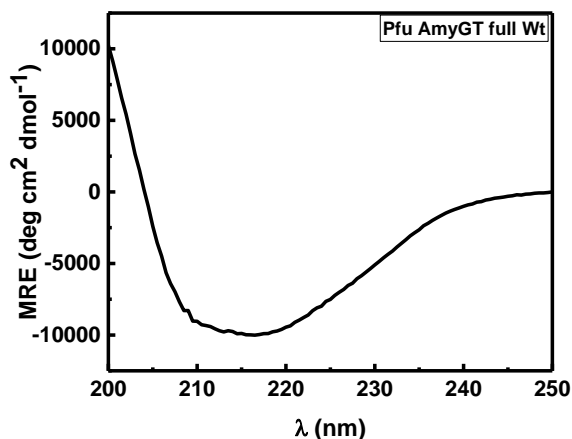


Figure 21: CD spectrum of PfuAmyGT.

3.3.3 Determination of the presence of tertiary structure:

To determine the presence of tertiary structure (or ‘foldedness’) of the protein, we exploited the presence of the aromatic amino acids tryptophan, tyrosine and phenylalanine in its sequence. The overall folding of the native protein was determined using intrinsic tryptophan fluorescence.

PfuAmyGT contains 18 tryptophan residues which make it good for analysis of intrinsic fluorescence, since tryptophan fluorescence is highly intense and also influenced by its microenvironment, providing clues to the presence of the tertiary structure. In hydrophobic environments, when tryptophans are buried inside the protein structure (i.e., with the indole side chain being the primary source of absorbance and emission) the emission maximum of the overall fluorescence emission envelope is at around 308-330 nm, whereas when these tryptophans are exposed to the aqueous solvent, the emission maximum is red-shifted to longer wavelengths and when they are completely exposed to the solvent to around ~350-353 nm. The protein sample of PfuAmyGT (0.15 mg/ml) was excited at 295 nm (specifically exciting tryptophan over tyrosine) to obtain the emission spectrum shown in Figure 22. The wavelength of maximum emission is 339 nm, suggesting that the bulk of the tryptophan residues of PfuAmyGT are buried within a reasonably well-folded protein.

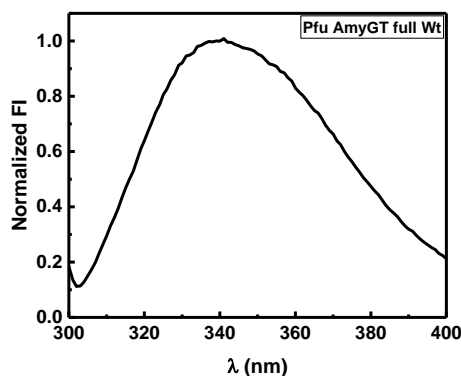


Figure 22: Fluorescence emission spectrum of PfuAmyGT.

The secondary and tertiary structural states, or status, of the individual and truncated versions of PfuAmyGT, were also similarly examined, using CD and fluorescence spectroscopy, respectively (Figure 23). All of them were found to be reasonably well-folded. The predicted structure of the PfuAmyGT domain 1 (D1), based on homology-based modeling, and the structure of the *T. litoralis* homolog, is a TIM-barrel like domain consisting of seven beta-alpha units (β/α)₇. The emission fluorescence of this construct peaks at 348 nm showing it to be folded domain. In the case of domain 3 (D3) which is predicted to be a β -sandwich, the CD spectrum shows a negative band at ~218 nm, in keeping with its domination by beta-sheet content. The fluorescence emission of this construct peaks at ~338 nm. Further, D1 + D2 is a combination of the TIM-barrel like fold and loops which is seen in its CD spectrum. The fluorescence peak of this construct is at ~339 nm. With D2 + D3, which is a combination of loops, a small α -helix, and a β -sandwich, this content to is reflected in its CD spectrum and a fluorescence peak at ~336 nm suggests that it can also fold independently.

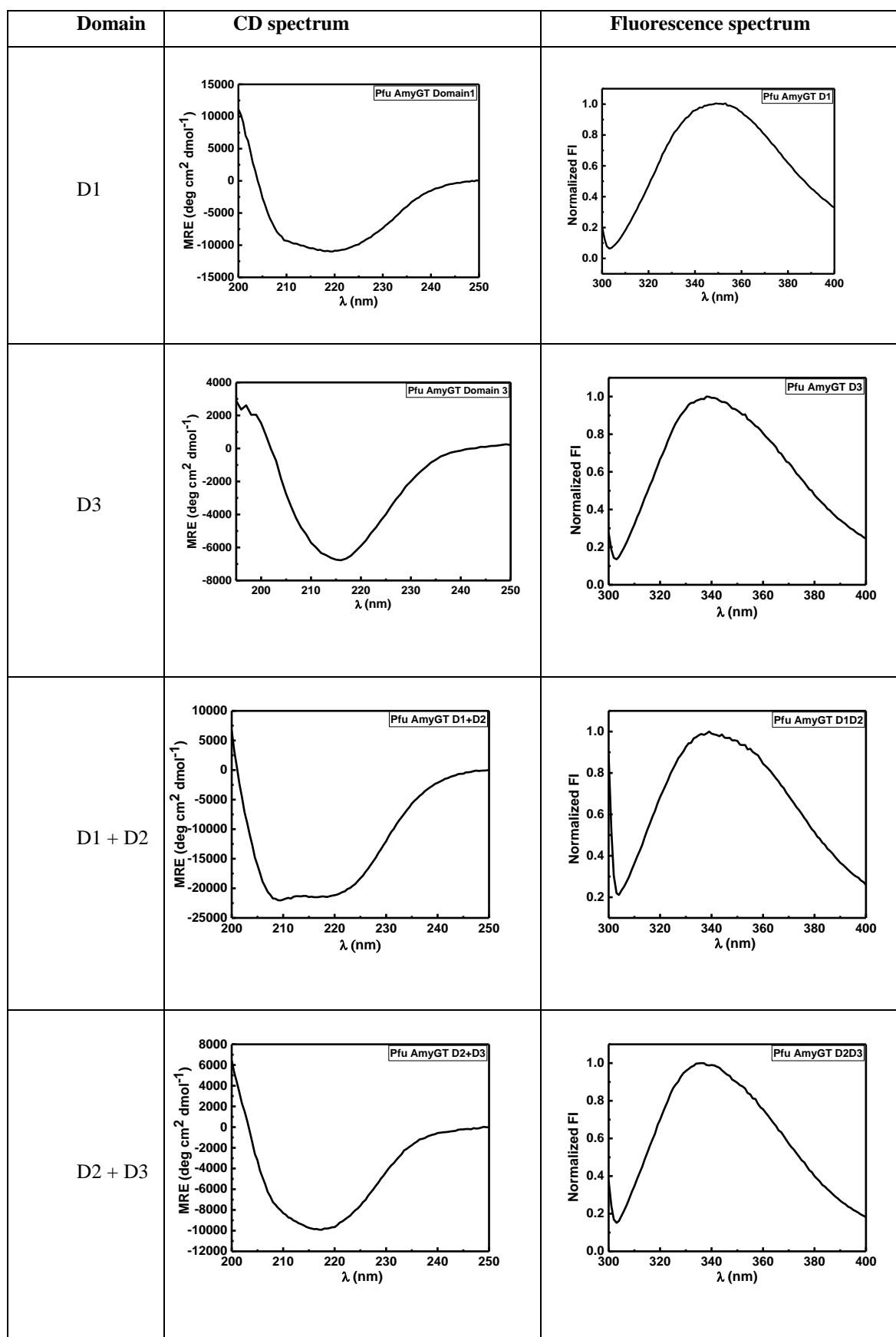


Figure 23: CD spectra and fluorescence emission spectra of different domains of PfuAmyGT.

3.3.4 Stability of PfuAmyGT and its domains:

3.3.4.1 Stability to Thermal Denaturation:

The thermal stability of the natively purified proteins (full-length PfuAmyGT and its variants) were determined by CD spectroscopy. The proteins were heated from 20 °C to 90 °C, with a constant heating rate of 1 °C/minute. No significant change in the secondary structure of the proteins was observed, in keeping with the expectation that this should be so, since the domains are sourced from a hyperthermophile protein and it should, therefore, be thermostable. Further, the mid-transition temperature (or T_m) for full-length PfuAmyGT could not be calculated, since there is no substantive unfolding and 87 % of the population remains folded even at 90 °C (as discerned from the changes in MRE) (Figure 24). This is not surprising for a protein of hyperthermophile origin.

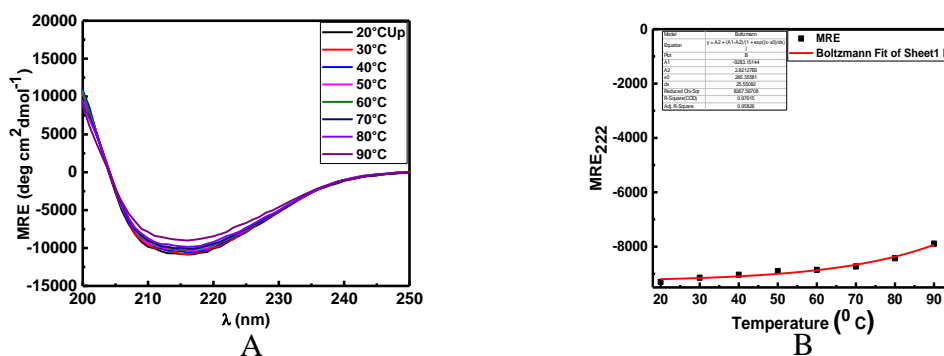


Figure 24: Thermal stability profile of PfuAmyGT using CD spectroscopy.

Thermal stability was also assessed by subjecting full-length PfuAmyGT to Differential Scanning Calorimetry (DSC). Samples with different concentrations of PfuAmyGT (0.25 mg/ml and 0.5 mg/ml) were heated from 20 °C to 120 °C at a scan rate of 90 °C/hour, as shown in Figure 25.

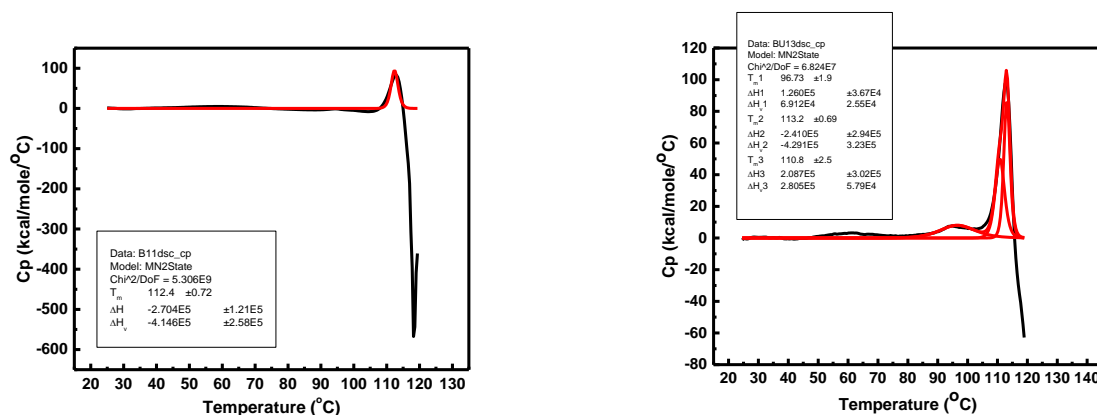


Figure 25: DSC transition curves of full-length PfuAmyGT at different concentrations. The left panel shows a melting curve and fit of PfuAmyGT at a concentration of 0.5 mg/ml and right panel shows a melt and fitting at concentration of 0.25 mg/ml.

When 0.25 mg/ml concentration was used the transition curve could be fitted to show three transition temperatures at 96 °, 110 °, and 113 °C, whereas when 0.5 mg/ml of the protein was used only one transition temperature of 112 °C was seen, along with evidence of heat-induced aggregation. Owing to its hyperthermophilic origin, such high melting temperatures are to be expected. The three transition peaks may also represent the melting of individual domains which show co-operative unfolding individually, within the full-length protein.

The thermal stabilities of individual domains 1 (D1) and 3 (D3) were also checked independently by secondary structure analysis using CD spectroscopy. A suitable amount of protein was heated from 20 °C to 90 °C with a constant heating rate of 1 °C/minute as shown in Figure 26 which suggests that these domains are as stable as the full-length protein as no significant change in secondary structure is observed.

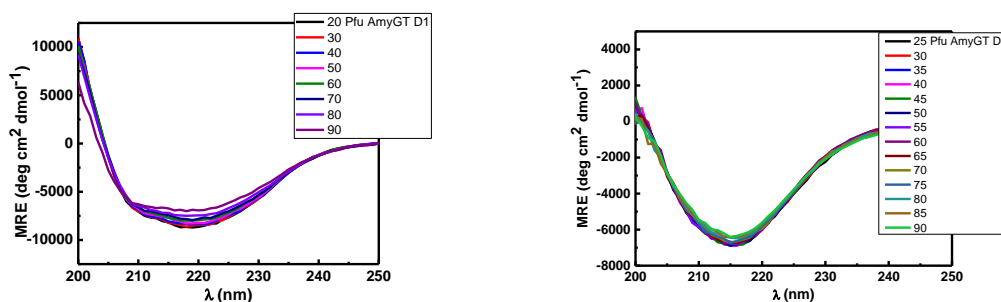


Figure 26: Thermal stability profile for PfuAmyGT domain 1 (D1) and domain 3 (D3).

3.3.4.2 Stability to Chemical Denaturation:

To study the chemical stability of the protein, 0.15 mg/ml of full-length PfuAmyGT protein was incubated with varying concentrations of Urea (0 M to 8 M) and Gdm.HCl (0 M to 6 M) and incubated overnight at room temperature. CD spectra were then recorded.

Urea is known to perturb the structures of proteins in two ways: through an “indirect” mechanism in which urea substitutes into water hydrogen bond network and through a “direct” mechanism in which urea interacts with protein backbone to weaken its hydrogen-bonding and hydrophobic interactions (Kuharski *et al*, 1984 and Soper *et al*, 2003).

Gdm.HCl at high concentrations mainly destabilizes electrostatic interactions in proteins (Greene and Pace, 1974) and also weakens hydrogen bonds. At low concentrations it presumably masks positively and negatively charged amino acid side chains, thus masking electrostatic repulsions and increasing stability (Oscar *et al* 1994). The concentration of Gdm.HCl required for protein denaturation as compared to urea is roughly half (Greene and Pace, 1974). Results obtained and shown in Figure 27 reveal that urea does not have a significant impact on the structure of full-length PfuAmyGT (Panel A). Further, a $C_m \sim 4.5$ M was obtained with Gdm.HCl, which in accordance with Greene and Pace (1974), the concentration of Gdm.HCl required for protein denaturation is roughly half as compared to urea.

Results and Discussion

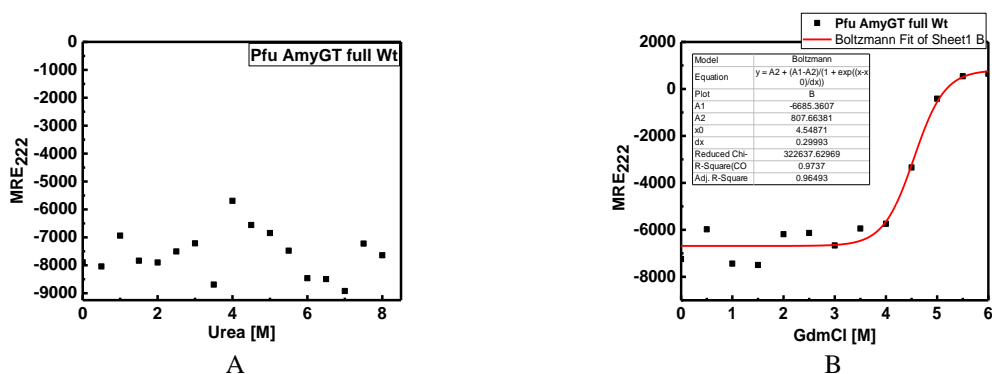


Figure 27: Panel A: Chemical stability of PfuAmyGT with urea Panel B: Chemical stability of PfuAmyGT with Gdm.HCl.

For domains 1 (D1) and 3 (D3), the protein constructs were incubated in varying concentrations of urea and analyzed by CD spectroscopy for stability against denaturation by chemical means. Figure 28, shows an apparent C_m (since the complete loss of structure is not seen) of urea for domain 1 (D1) to be 7.3 M and for domain 3 (D3) to be 7.2 M.

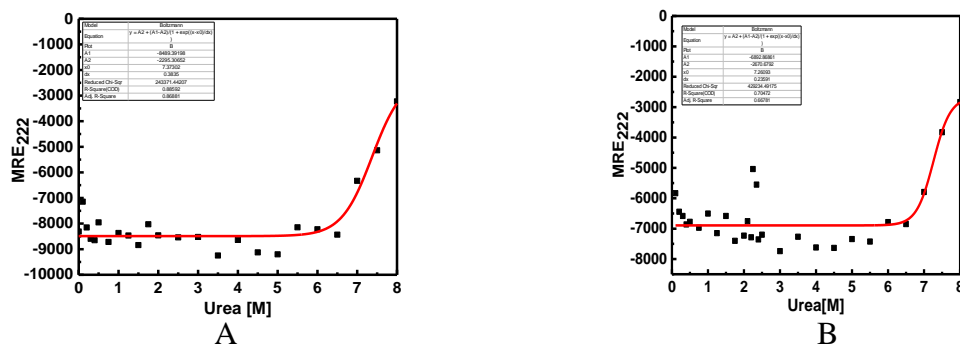


Figure 28: Panel A: Apparent C_m of urea for PfuAmyGT domain 1 (D1) while Panel B shows it for PfuAmyGT domain 3 (D3).

In the case of Gdm.HCl, for D1 C_m , was observed to be ~3.3 M. Gdm.HCl was unable to unfold D3 completely and the apparent C_m obtained for it was ~2.3M (Figure 29).

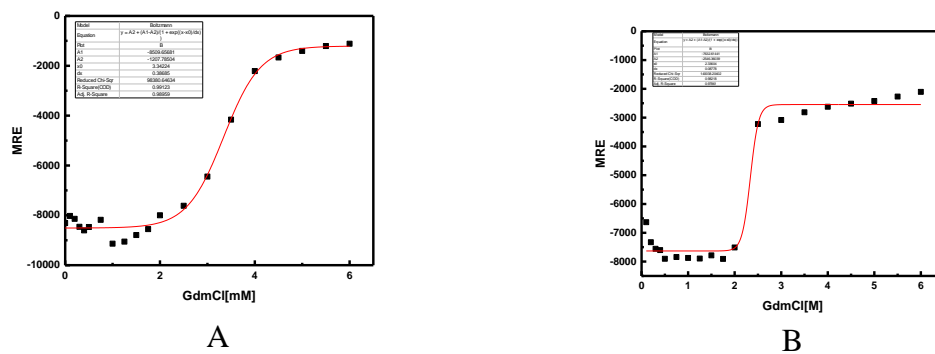


Figure 29: Panel A: C_m of Gdm.HCl for PfuAmyGT domain 1 (D1) while Panel B shows it for PfuAmyGT domain 3 (D3).

3.4 Amylase activity in PfuAmyGT and its domains:

3.4.1 Starch-iodine and TLC based assays of amylase activity:

The activity of the full-length protein of wild-type sequence and its truncated variants was first assessed as an α -amylase (Ladermann *et al*, 1993). The standard starch iodine method was used for analysis of hydrolyzed substrate (Fuwa, 1954).

As shown in Figure 30, this enzyme showed amylase activity on 0.2 % starch only after incubation at higher temperatures i.e. 70 °C, 80 °C and 90 °C (after 120 minutes of incubation). Since this enzyme is derived from a hyperthermophilic organism this result is in accordance with its nature of origin. Activity here is seen as a degradation of starch which upon reaction with iodine gives a pale yellow color whereas un-degraded starch gives blue-black color. This activity was quantified by measurement of absorbance at 620 nm (shown later).

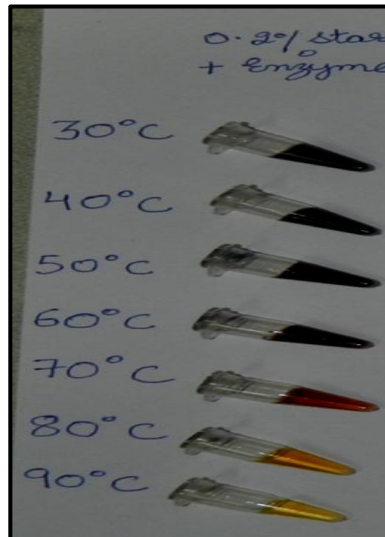


Figure 30: A visual representation of amylase activity obtained by the action of full-length PfuAmyGT on starch at different temperatures.

3.4.2 Plate-based assay:

In a plate-based starch degrading assay, 0.2 % starch was co-solidified with LB agar. Different concentrations of the enzyme were placed in wells cut on the plate which enabled it to diffuse and hydrolyze the starch substrate which was observed in the form of a zone of clearance (Figure 31). The radii of these clearance zones increased with increasing enzyme concentrations.

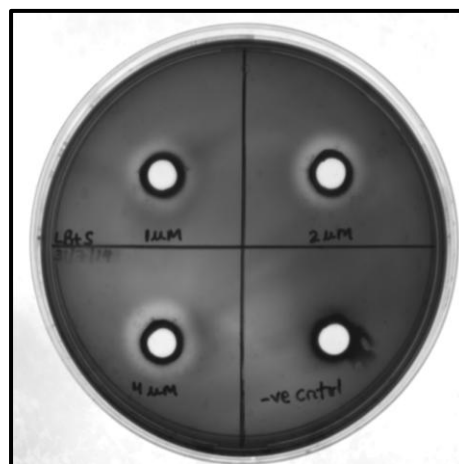


Figure 31: Amylase activity observed as clear zones around cut wells which contain different concentrations of full-length PfuAmyGT.

3.4.3 Zymogram:

In a similar manner as above, 0.2 % starch was copolymerized with native or non-denaturing (i.e., not containing SDS) polyacrylamide gel and protein was loaded in it at varying concentrations. As the protein moved across the gel it simultaneously degraded starch which was visualized by staining the gel with iodine solution. This solution stained the background blue as it would contain undegraded starch while the areas where starch had degraded because of the action of the enzyme were seen as clear zones (Figure 32).



Figure 32: Panel A: It shows bromophenol blue stained native PAGE of full-length Pfu AmyGT which was loaded in different (increasing) concentrations. Panel B: It shows iodine stained native PAGE gel which shows activity in form of clear zones.

3.4.4 Thin Layer Chromatography (TLC):

We further investigated the amylase activity of this enzyme using thin layer chromatography. Using a mobile phase composed of Butanol: Ethanol: Water in the ratio of 50:30:20 respectively. We were able to separate the products of the enzymatic reaction on the basis of their polarity and they were compared to a set of standards run with it shown in Figure 33.

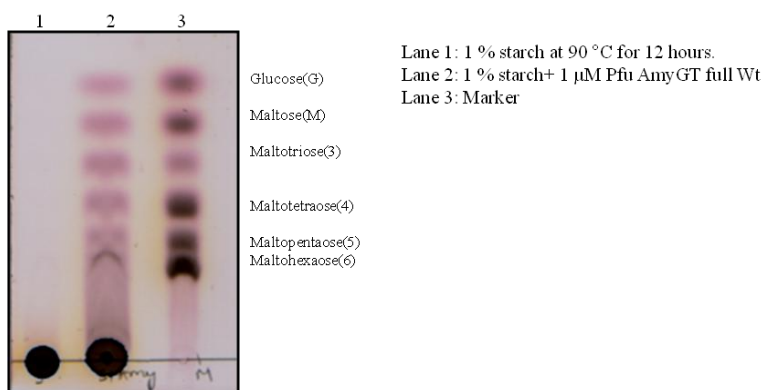


Figure 33: TLC depicting the reaction products of full-length PfuAmyGT on 1 % starch.

Results and Discussion

Starch being almost equally polar to silica did not move from the spot in this solvent system. The products of the reaction contained different species of different polarity and thus they have been separated in this solvent system. Since glucose is the least polar molecule in the mixture it is carried the farthest on the TLC plate followed by maltose, maltotriose and so on.

The presence of glucose as one of the products confirms this enzyme to possess an exo-amylase activity, but the presence of other oligosaccharides cannot be explained if this enzyme is categorized only as an exo-amylase and there is a need to conceive of it as a glucanotransferase capable of first hydrolyzing starch through exo-amylase action, and then disproportionating the starch into a pool of maltooligosaccharides through glucanotransferase action.

The truncated variants, however, did not show any amylase activity on starch (under similar experimental conditions, Figure 34).

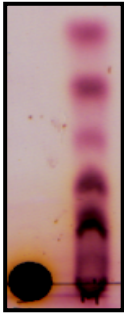


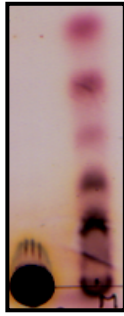
Domain	D1	D3	D1 + D2	D2 + D3	
TLC					In each case: Lane 1: 1 % starch+ 1μM enzyme Lane 2: Marker

Figure 34: Activity of different domains of PfuAmyGT on 1 % starch. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.4.5 Amylase activity is significant at higher temperatures:

As is also seen in Figure 30, the amylase activity of full-length PfuAmyGT is observed at higher temperatures (70 °C to 90 °C). TLC also shows similar results as shown in Figure 35.

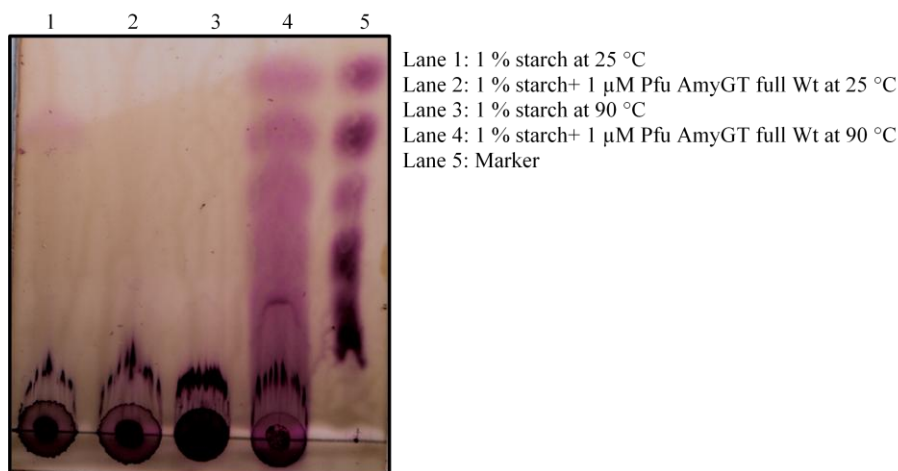


Figure 35: TLC showing the amylase activity of full-length PfuAmyGT (25 °C and at 90 °C). The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.4.6 Amylase activity is aided by the addition of maltose:

Lee *et al* in 2005 reported global transcriptional patterns (using whole-genome DNA microarrays) of various *Pyrococcus furiosus* derived genes when maltose is used as a primary carbon source for it. They have reported that the expression of gene PF0272 (which expresses PfuAmyGT) increases most dramatically when the organism is grown on maltose and that this enzyme is not able to degrade starch.

Taking a cue from this report we decided to check the activity of PfuAmyGT full-length protein in presence of maltose.

Therefore, the starch-iodine assay, plate-based assay, and zymogram were carried out in a manner similar to that described above, but now in the presence of excessive maltose concentrations. Since we are providing maltose as an activator, it is acting as a primer to the enzyme, triggering the amylase reaction to occur at a faster rate. As we observed in the starch-iodine assay (Figure 36), starch is now

Results and Discussion

being degraded even at ambient temperatures in the presence of maltose which was earlier occurring only at high temperatures.

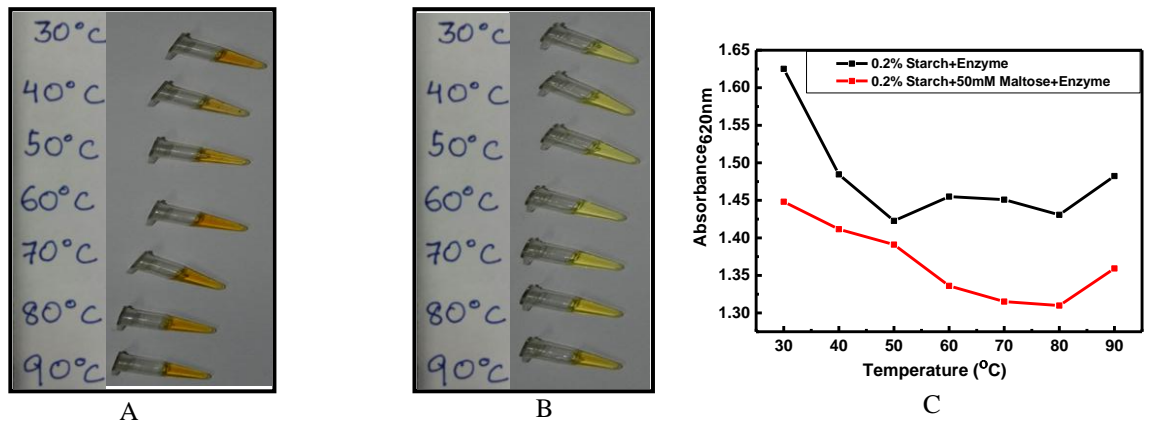


Figure 36: Panel A: Shows the activity of PfuAmyGT on 0.2 % starch in presence of 50 mM maltose. Panel B: It is a control reaction showing the color of only maltose with iodine reagent. Panel C: It shows the absorbance of the reactions at 620 nm.

Absorbance taken at 620 nm for the two reactions (decrease in absorbance is equivalent to increase in activity) i.e. 0.2 % starch + PfuAmyGT and 0.2 % starch+ 50 mM maltose + Pfu Amy GT shows a fold difference of approximately 10 fold at 30 °C and around 8.5 fold at 90 °C. Hence we observed that the amylase activity of PfuAmyGT is increased in the presence of maltose.

Similar results were observed qualitatively in a plate-based starch degrading assay and zymogram as shown in Figure 37.

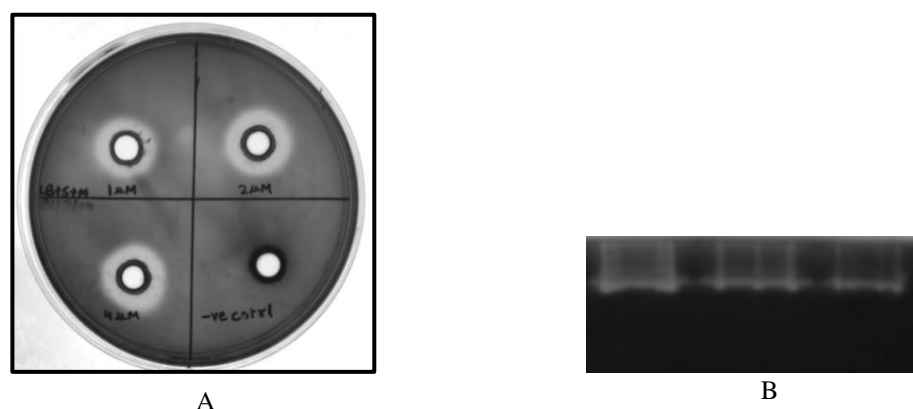


Figure 37: Panel A: Shows clear zones on an agar plate containing 0.2 % starch and 50 mM maltose. Panel B: Shows clear zones on native PAGE gel containing 0.2 % starch and 50 mM maltose.

TLC also showed the intensification of oligosaccharide bands obtained as products (Figure 38) in the presence of different maltose concentrations along with 1 % starch.

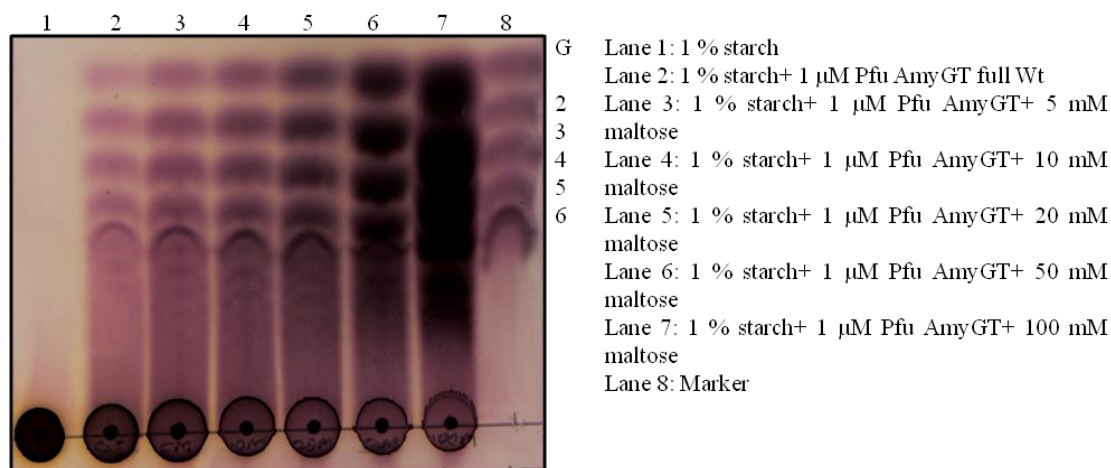


Figure 38: TLC showing an increasing intensity of oligosaccharide products with increasing maltose concentration.

We can conclude from the above data that this enzyme has an exo-amylase activity to a small extent since it is able to degrade starch and produce glucose.

Based on these results we also dismissed our first hypothesis that maltose may be acting as an allosteric activator of the amylase activity of PfuAmyGT, because the concentrations of maltose that increase the activity of PfuAmyGT both as an apparent amylase, and as a glucanotransferase (creating a pool of maltooligosaccharides out of starch) are far higher (continuously increasing up to several hundred millimolar maltose) than the concentrations at which small molecules act as allosteric activators of enzymes for micromolar enzyme concentrations. We were thus forced to consider alternative explanations for the effect of the addition of maltose, e.g., the possibility that maltose, by acting as a sink at the acceptor site of a glucanotransferase, for glucose excised at the donor site, somehow increases the rate of excision of glucose at the donor site.

3.4.7 Amylase activity co-operates with a glucanotransferase activity to create multiple oligosaccharides from a single substrate:

One of the most well-characterized enzymes with the specificity to carry out glucosyl transferase reactions is amylomaltase from *E. coli* also known as MalQ. This enzyme is able to transfer a non-reducing dextrinyl moiety to glucose or

another maltooligosaccharide. However, it can neither transfer maltose, nor glucose, to another glucose, or maltose, when present in a reaction mixture, nor can it generate maltose as the main product in a reaction. This information is revealed by radioactivity based kinetic studies carried on this enzyme by Palmer *et al* in 1976.

In light of this information, the above results can be interpreted as glucosyl transferase reactions. In the above reactions, starch (acting as a donor molecule) is first degraded by exo-amylase action to generate glucose which then acts as an acceptor molecule in the first reaction to generate maltose, and thereafter in cycles of growing length through disproportionation reactions, to generate oligosaccharides of varying lengths by addition of glucose molecules, as shown in the schematic representation in Figure 39. The different products thus formed can also now act as both donors and acceptors (shown later) to yield different oligosaccharides.

In the case of PfuAmyGT, we see a good concentration of maltose being formed as a product which is unlike the action of MalQ.

Palmer *et al* argued that a trace amount of maltotriose and/or higher oligosaccharides are always present as contaminants when using maltose as a substrate or acceptor molecule to prime the reaction and initiate formation of oligosaccharides. This argument can be extrapolated in our case to support the existence of trace amounts of glucose or maltose being present in starch as contaminants to prime and initiate the transferase reaction and that there is no actual exo-amylase activity in PfuAmyGT of its own. On the other hand, it could also be that there is an exo-amylase activity which is increased by the presence of suitable acceptor glucans.

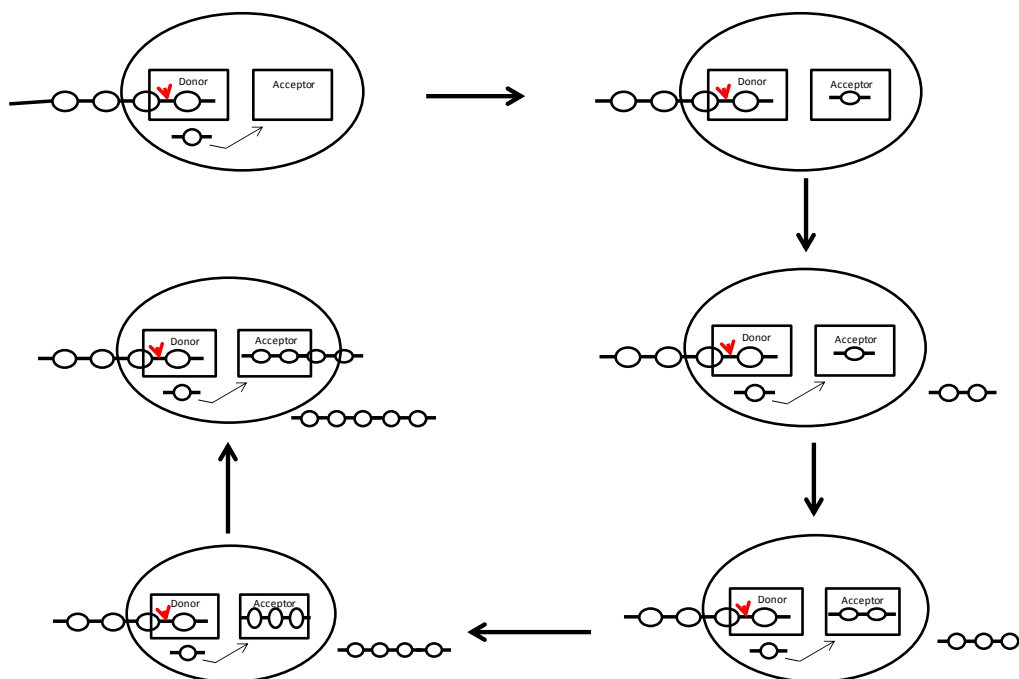


Figure 39: Schematic representation of the formation of different oligosaccharides by the action of PfuAmyGT.

3.5 Characterization of Glucanotransferase activity of PfuAmyGT and its domains:

3.5.1 Conversion of starch to multiple oligosaccharides is temperature dependent:

The ability of PfuAmyGT to generate different oligosaccharides from starch depends on the temperature at which the reaction is carried out. As shown in Figure 35, the extent of reaction (here seen in terms of intensity of oligosaccharide produced) is higher when carried out at 90 °C than at 25 °C for the same time period. This observation supports the fact that PfuAmyGT is derived from a hyperthermophilic organism but it has a wide temperature range of operation and that maltose is equally a good primer at 25 °C as at higher temperatures. The difference between the activities at the lower and higher temperatures in the presence and absence of maltose suggest that the presence of maltose as an acceptor increases the catalytic rate (Figure 40).

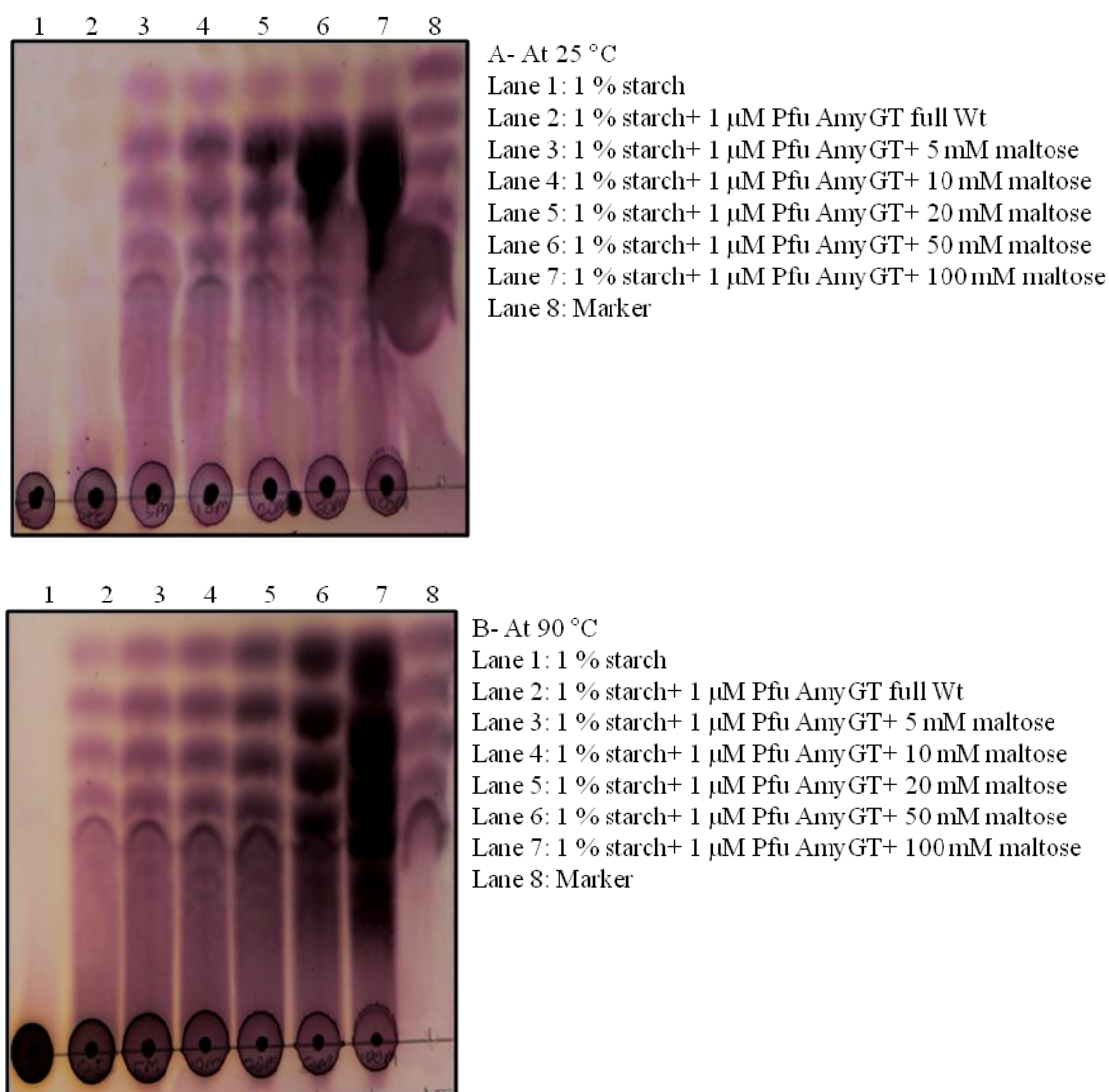


Figure 40: Panel A shows the transferase reaction carried out by PfuAmyGT at 25 °C while Panel B shows the same reaction carried out at 90 °C. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.5.2 Maltose accelerates the reaction by acting as an acceptor:

When excessive maltose is already provided in the reaction mixture along with starch it appears to act as a primer to increase the rate at which multiple oligosaccharides are formed. In Figure 41, we observe that the amount of different oligosaccharides produced (i.e., the band intensities of individual oligosaccharides) is higher when maltose is added at the start of the reaction, within a given duration of the reaction. Both reactions (1 % starch + 1 μ M PfuAmyGT and 1 % starch + 20 mM maltose + 1 μ M PfuAmyGT) were

incubated at 90 °C and samples were taken at intervals of 2 hours for 12 hours from the start of the reaction. Hence we confirm that maltose accelerates the transferase reaction of PfuAmyGT.

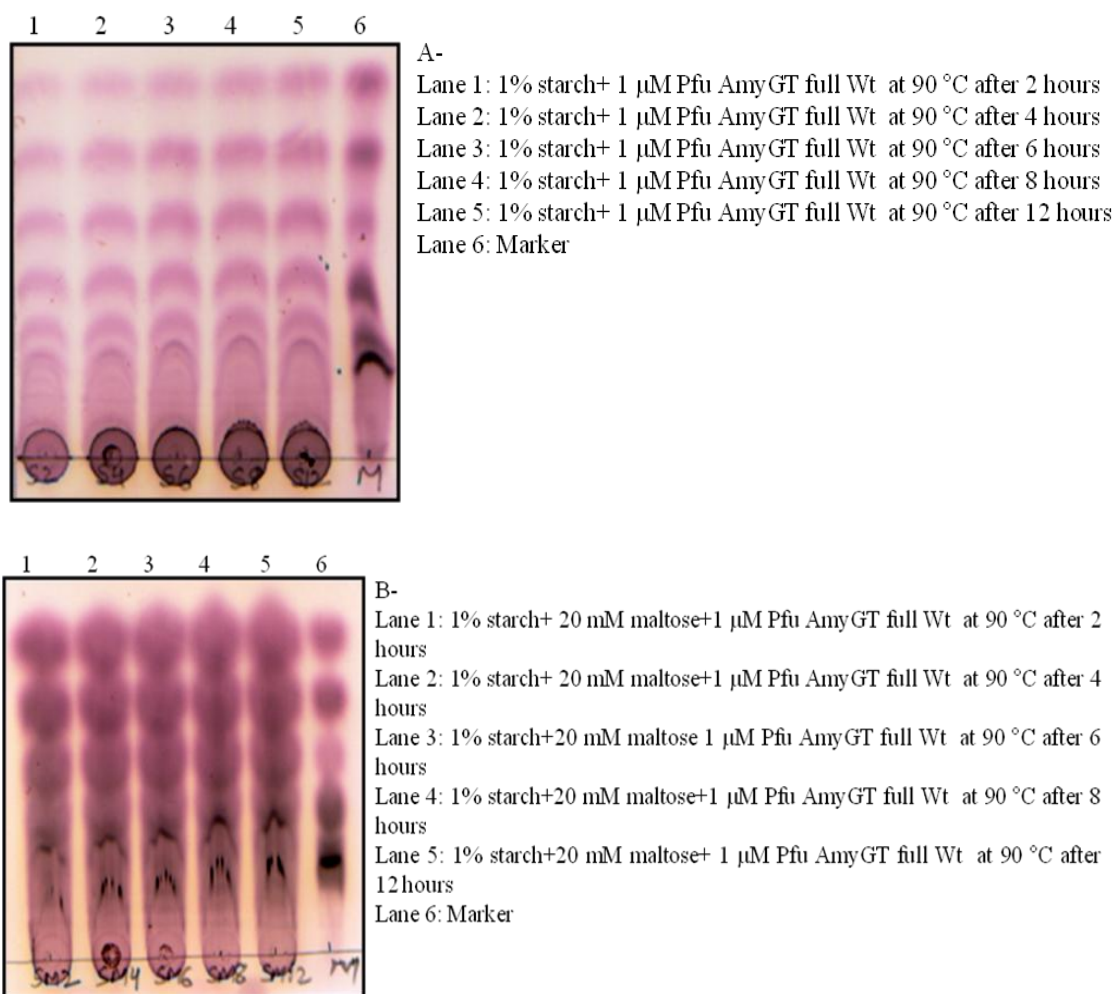


Figure 41: Panel A shows the action of PfuAmyGT on starch at different time intervals and Panel B shows its action when maltose is added at the start of the reaction. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.5.3 The transferred unit is glucose:

In all the above TLCs we can see the oligosaccharides formed differ in size by one glucose molecule thus, it can be concluded that the smallest transferred unit between a donor and acceptor molecule is glucose.

3.5.4 The smallest donor capable of transferring glucose is maltotriose:

Starch is a polysaccharide of glucose molecules. So to study the minimum length of glucose units which can act as a donor molecule in this reaction we carried out a series of experiments trying different combinations of donor and acceptor molecules (Figure 42).

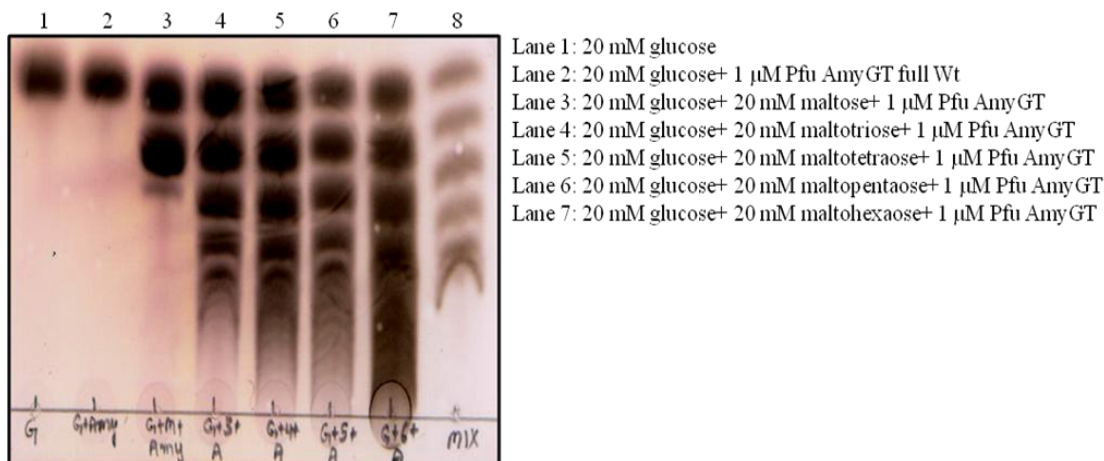


Figure 42: Determining the smallest donor and acceptor molecules for PfuAmyGT. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

In lane 2 of the above figure, which contains 20 mM glucose + PfuAmyGT, no formation of oligosaccharides is seen and so is the case when glucose and maltose are present with PfuAmyGT. Thus maltose cannot act as a donor molecule for PfuAmyGT. But when maltotriose is present with glucose we see formation of all the other (and higher) oligosaccharide species of different lengths.

Here a band of maltose is seen to contain some maltotriose as a contaminant (lane 3), and we are also providing glucose but still we do not see production of all other oligosaccharides, presumably because the concentration of maltotriose is not enough to trigger and drive the transferase reaction and to drive out maltose, by competition, from occupying the donor site like the case argued by Palmer *et al* for MalQ.

3.5.5 The smallest acceptor capable of an accepting the transferred glucose is glucose itself:

In order to find out the smallest unit capable of accepting the glucose molecule by PfuAmyGT, we used different concentrations of glucose with starch as seen in Figure 43. Here we observed that glucose is an equally good acceptor molecule along with starch and formation of different oligosaccharides is seen, exactly as seen earlier with maltose addition to starch and PfuAmyGT. Formation of maltose is also seen again in all these reactions.

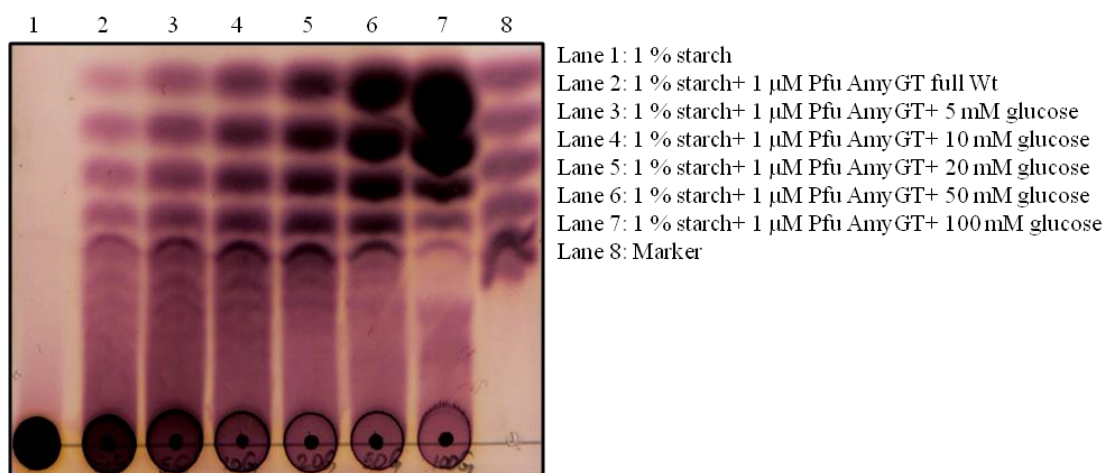


Figure 43: Glucose is the smallest acceptor molecule required for PfuAmyGT action. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

After establishing the smallest donor molecule to be maltotriose and smallest acceptor molecule to be glucose, a series of experiments were done which led to the conclusion that maltotriose is able to act both as donor and acceptor (Figure 44) i.e. when 20 mM maltotriose is present with PfuAmyGT, we see the formation of all other oligosaccharide species. And even when maltotriose is present with maltotetraose, maltopentaose and maltohexaose formation of high molecular weight species are seen, in fact, to such an extent that the material is seen to streak all the way down till the spot of sample loading.

Results and Discussion

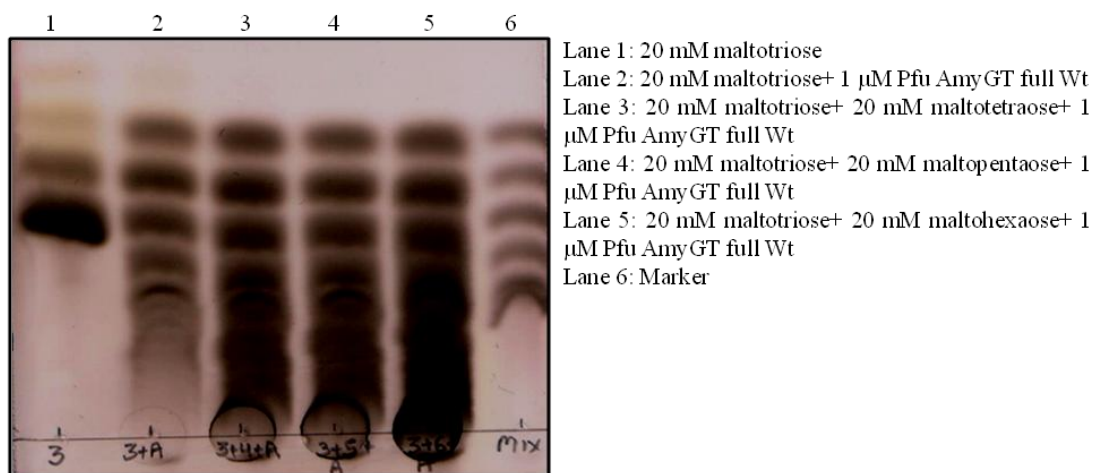
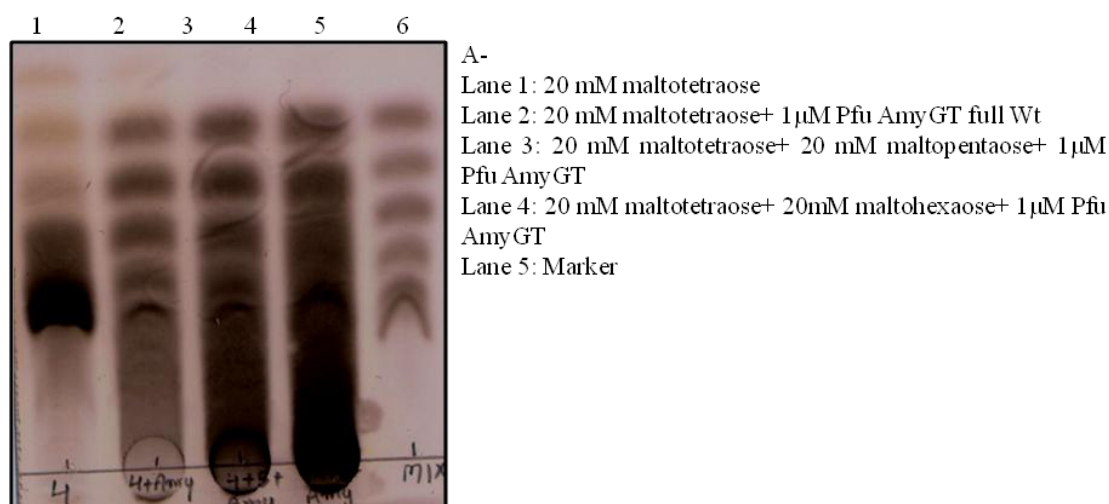


Figure 44: The above TLC shows that maltotriose can act both as a donor molecule as well as an acceptor molecule to give rise to all other oligosaccharide species. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

The maltooligosaccharides such as maltotetraose, maltopentaose, and maltohexaose can also be used both as donor and acceptor molecules to generate higher maltooligosaccharides (Figure 45).



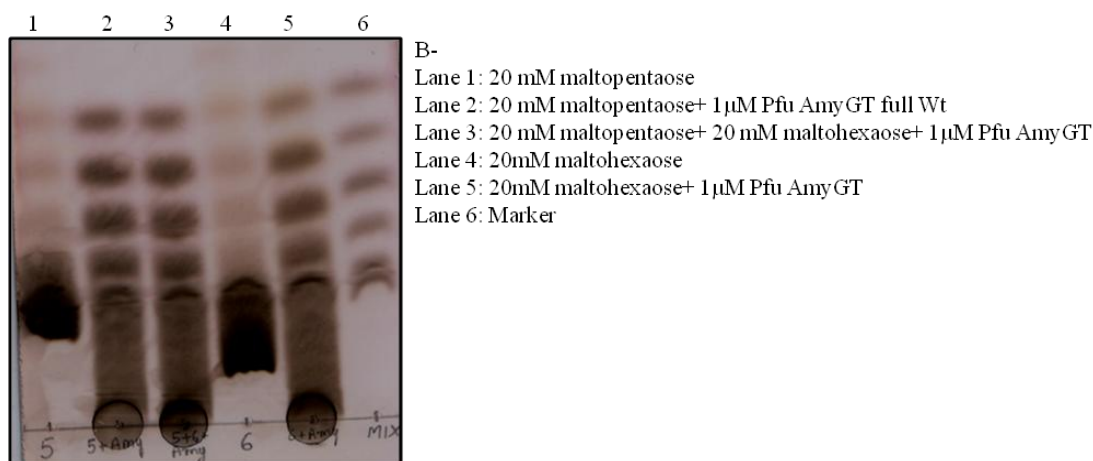


Figure 45: The TLCs in this figure show the formation of different oligosaccharide species when donor molecules other than maltotriose are used. In Panel A maltotetraose is used in combination with higher saccharides while in Panel B maltopentaose is used. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

The schematic in Figure 46 below shows the molecules which can act as donor and acceptor molecules for PfuAmyGT.

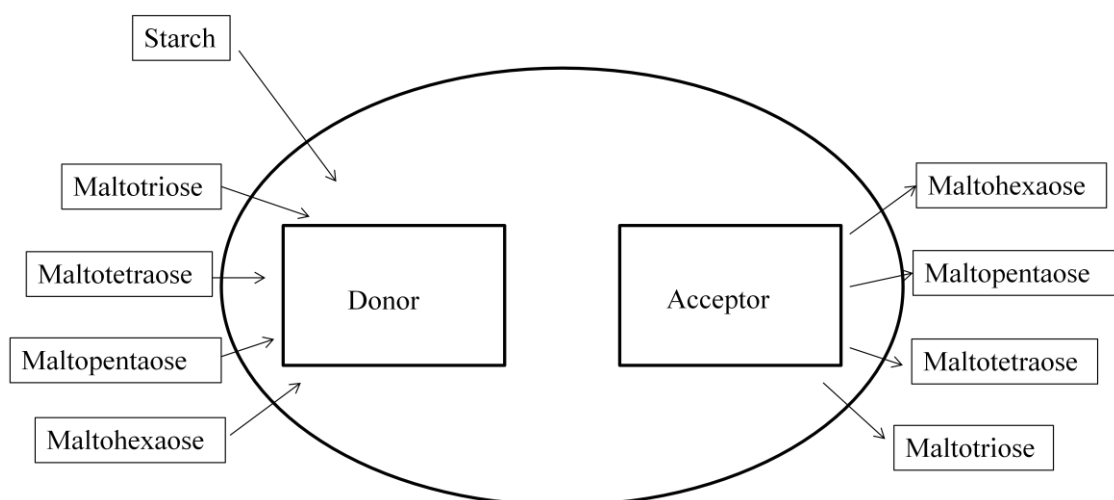


Figure 46: A schematic to show that PfuAmyGT can accommodate different oligosaccharides both as donor and acceptor molecules.

Keeping in mind that maltotriose is able to act as both donor and acceptor for the activity of PfuAmyGT, the transferase activity was then characterized (for ease of experimentation) in terms of time, temperature, pH, and effect of organic solvent using maltotriose as the sole substrate.

3.5.6 Time-dependence of PfuAmyGT activity:

An enzyme sample of 1 μM concentration was incubated with 20 mM maltotriose at 90 °C and aliquots were taken at different time intervals to examine products, using TLC. We observed that the enzyme is active after 2 hours of incubation and remains active even after incubation for 12 hours at 90 °C (Figure 47).

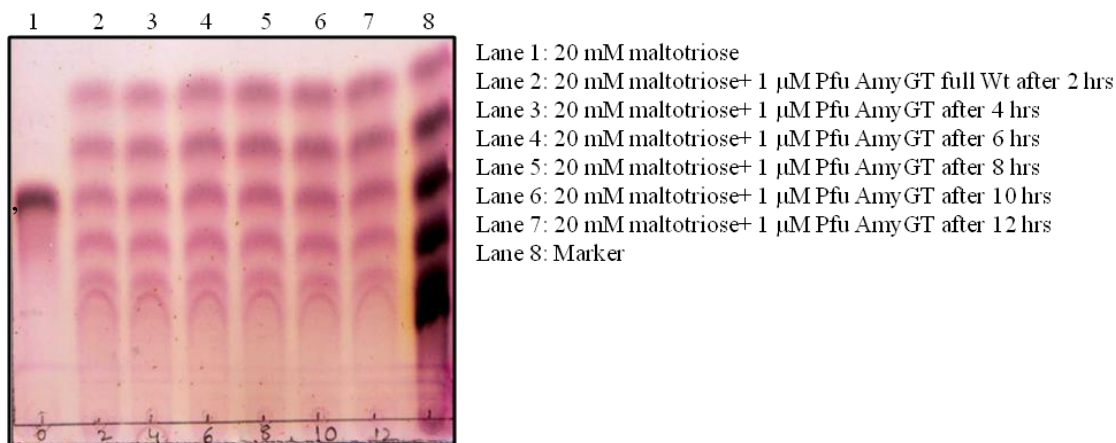
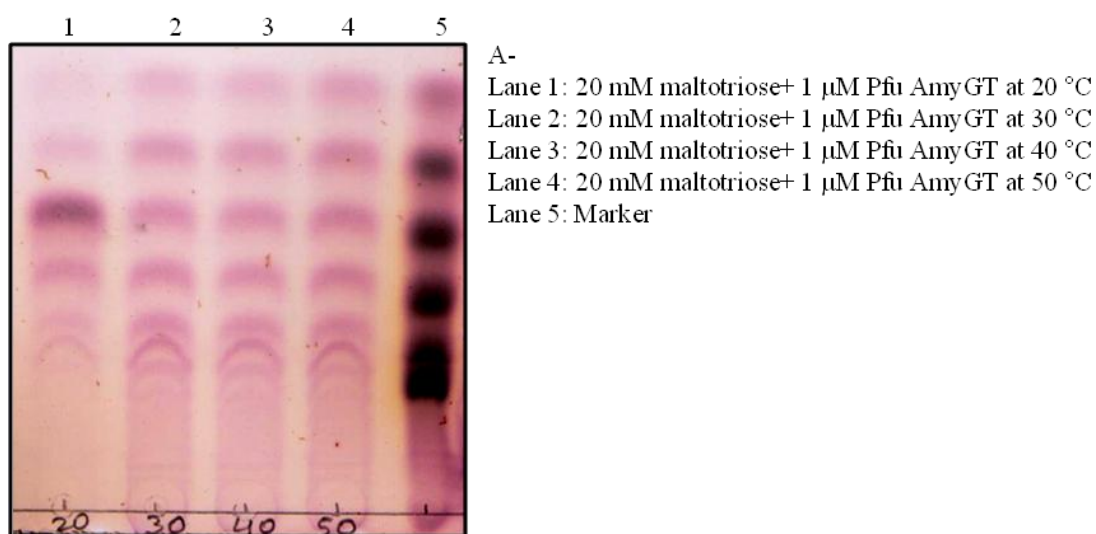


Figure 47: Time-dependent profile of PfuAmyGT activity. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.5.7 Temperature-dependence of PfuAmyGT activity:

Similarly, the activity of this enzyme was examined from 20 °C to 100 °C. We see that the enzyme is comparatively less active at 20 °C as compared to 100 °C but becomes significantly active after 50 °C (Figure 48).



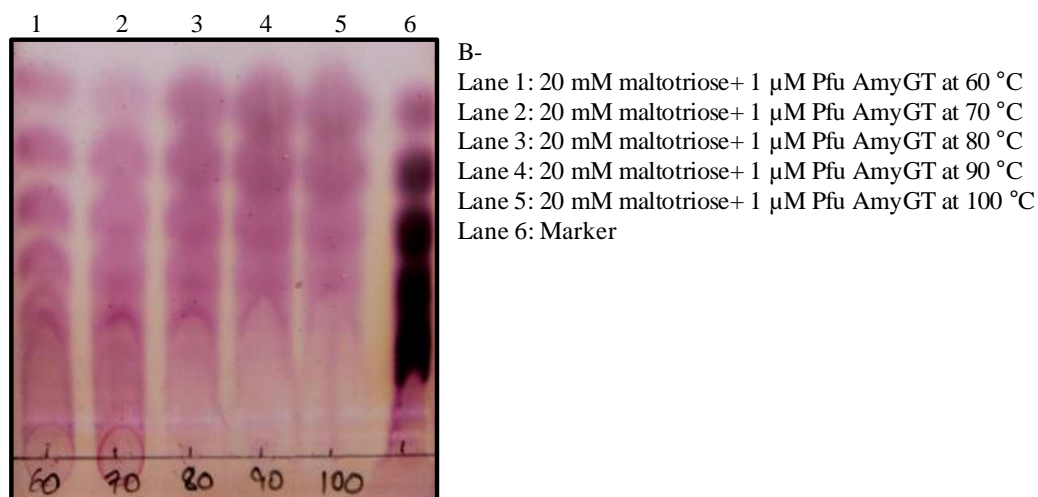
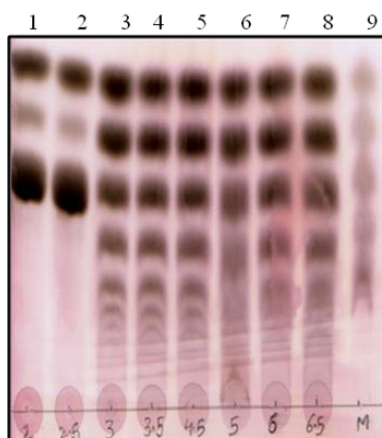


Figure 48: The above TLCs show the temperature profile of PfuAmyGT. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

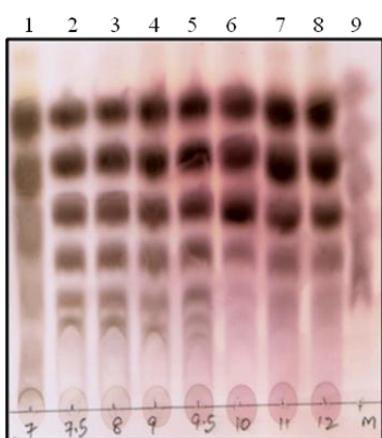
3.5.8 pH-dependence of PfuAmyGT activity:

To assess the working pH range of this enzyme, 1 μ M enzyme was incubated at 90 °C for 12 hours with 20mM maltotriose + 20mM glucose and 1 % starch+ 20 mM glucose in separate reactions and in buffers ranging from pH 2.0 to pH 12.0 (Figure 49).

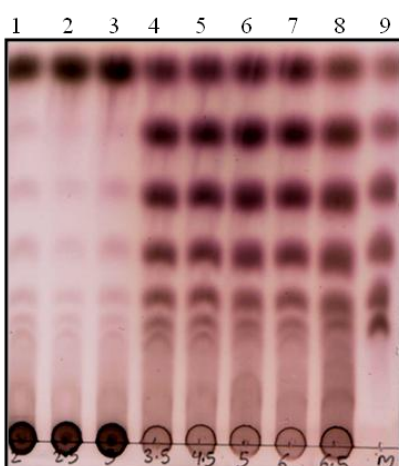
Results and Discussion



- A-
- Lane 1: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT full Wt at pH 2
 - Lane 2: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 2.5
 - Lane 3: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 3
 - Lane 4: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 3.5
 - Lane 5: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 4.5
 - Lane 6: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 5
 - Lane 7: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 6
 - Lane 8: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 6.5
 - Lane 9: Marker



- B-
- Lane 1: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT full Wt at pH 7
 - Lane 2: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 7.5
 - Lane 3: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 8
 - Lane 4: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 9
 - Lane 5: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 9.5
 - Lane 6: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 10
 - Lane 7: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 11
 - Lane 8: 20 mM glucose+ 20 mM maltotriose+ 1 μ M Pfu AmyGT at pH 12
 - Lane 9: Marker



- C-
- Lane 1: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT full Wt at pH 2
 - Lane 2: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 2.5
 - Lane 3: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 3
 - Lane 4: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 3.5
 - Lane 5: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 4.5
 - Lane 6: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 5
 - Lane 7: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 6
 - Lane 8: 20 mM glucose+ 1 % starch+ 1 μ M Pfu AmyGT at pH 6.5
 - Lane 9: Marker

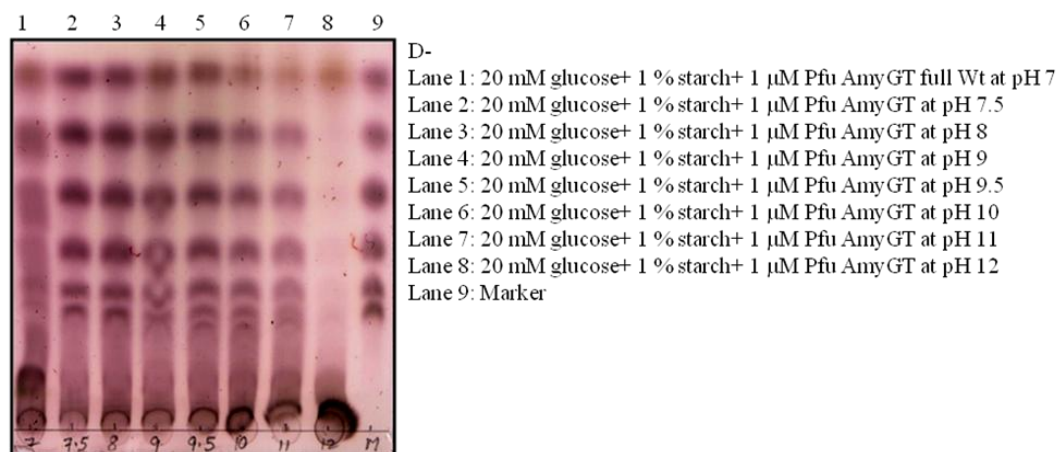


Figure 49: Panel A shows PfuAmyGT action on maltotriose+glucose from pH 2.0 to 6.5
 Panel B shows PfuAmyGT action on maltotriose+ glucose from pH 7.0 to 12.0
 Panel C shows PfuAmyGT action on starch+ glucose from pH 2.0 to 6.5 and
 Panel D shows PfuAmyGT action on starch+ glucose from pH 7.0 to 12.0

In the case of the reaction involving 20 mM maltotriose + 20 mM glucose, significant activity is seen from pH 3 to pH 9.5. At pH values of 2.0 and 2.5, some pH-dependent hydrolysis of the substrate is seen but no activity of the enzyme is seen, whereas at pH values of 10, 11, and 12, the number of oligosaccharide species formed is less and the activity is considerably reduced. Similar is the case when starch+glucose are used as substrates. Initially, some pH dependent hydrolytic products are seen and the activity is seen in the pH range from 3.5 to 10. Thus, it is evident that this enzyme has a wide range of pH over which it shows significant activity.

3.5.9 Effect of organic solvents:

In general, for a transferase reaction to take place, first a hydrolysis step is required in order for a glucose molecule to be released which is then transferred to a donor by the formation of a glycosidic linkage. Theoretically, such reactions can also be reversed (Cote and Tao 1990). In the case of our enzyme, if we replace bulk water in the solvent system with an organic solvent, one anticipates that it should reduce the amylase activity of the enzyme and consequently also hinder the transferase activity, because in the absence of amylase activity no acceptor molecule will be generated (unless provided separately in the reaction

Results and Discussion

mixture) to generate different oligosaccharide species. We tested this by examining activity in the presence of organic solvents.

For this analysis, 1 μ m enzyme was incubated with 1 % starch and 20 mM maltotriose in separate reaction mixtures and in the presence of 0.1 %, 1 %, 2 %, 5 %, 10 %, 20 %, and 50 % acetone, butanol and ethanol at 90 °C for 12 hours.

Starch and/or maltotriose did show any significant auto degradation in the presence of varying concentrations of acetone. Activity was observed up to 20 % acetone, with a drastic reduction at 50 % acetone, suggesting that the enzyme can tolerate quite a high percentage of acetone; higher than 20 % and less than 50 % (Figure 50).

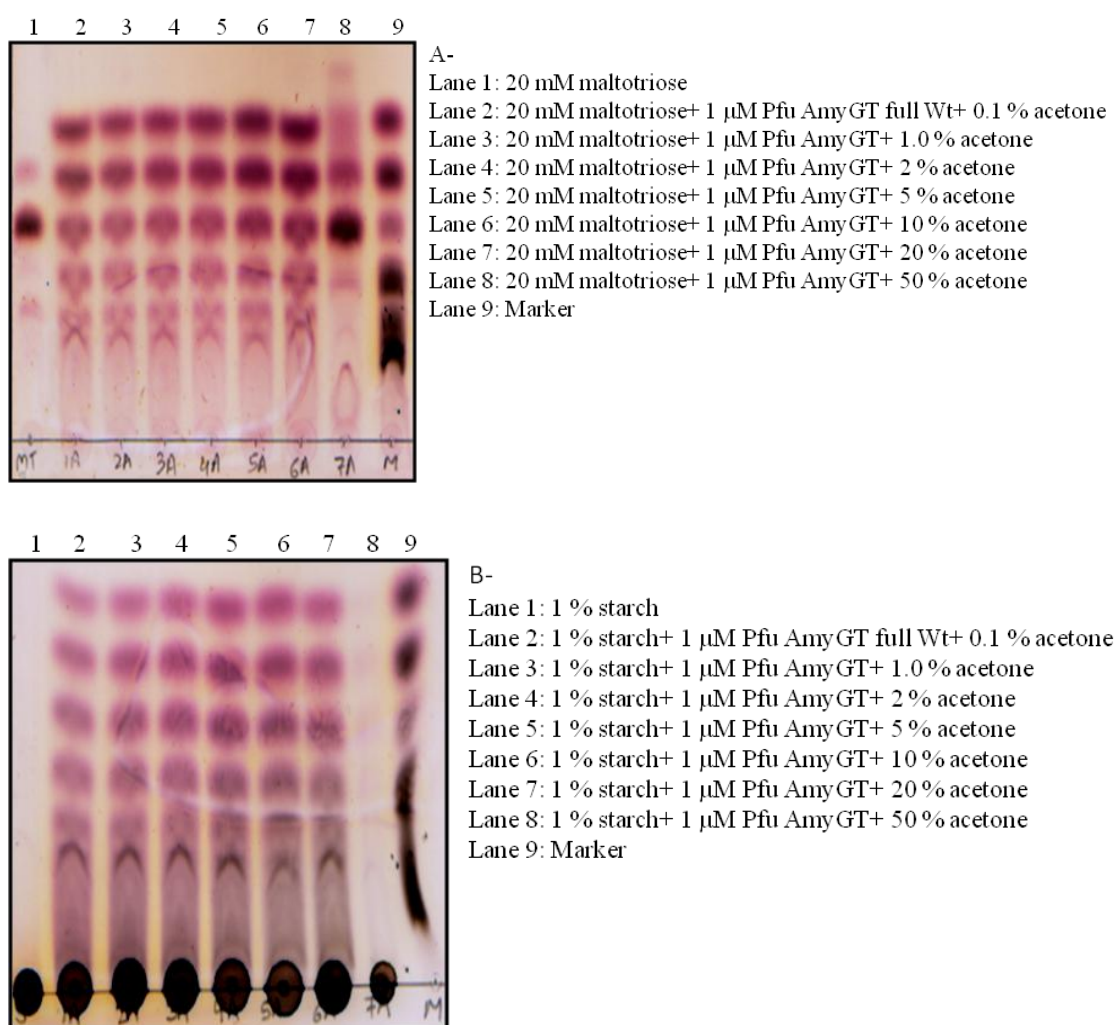


Figure 50: Panel A shows the activity of PfuAmyGT on 20 mM maltotriose in presence of different acetone concentrations and Panel B shows the activity on 1 % starch.

To ensure that the enzyme is in its native conformation in such high percentages of acetone, CD spectra were collected as shown in Figure 51.

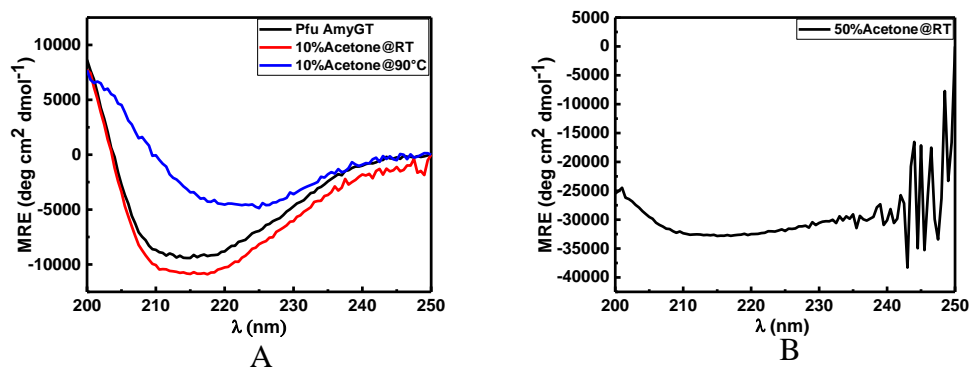
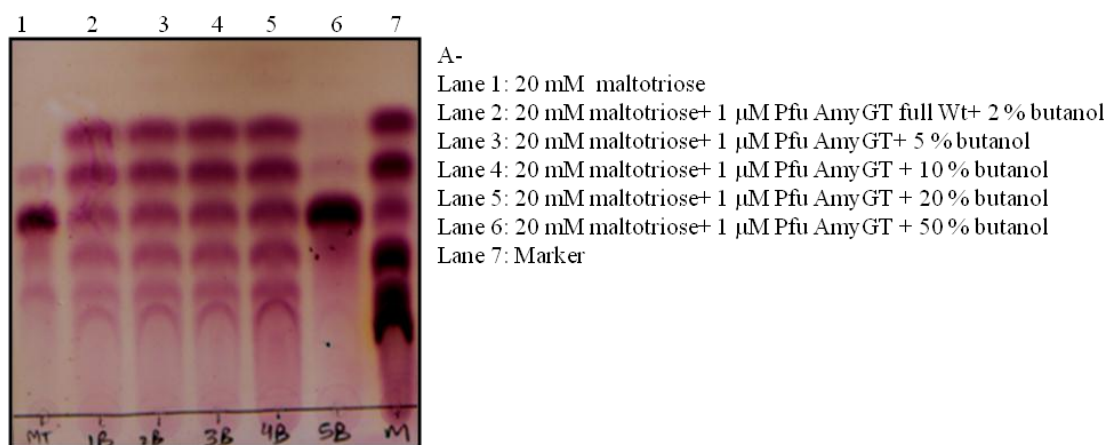


Figure 51: Panel A shows the CD spectra of native PfuAmyGT in presence of 50 mM Tris (pH 8.0) and in the presence of 10 % acetone after an overnight incubation both at room temperature and at 90 °C. Panel B shows the spectrum of PfuAmyGT incubated overnight at room temperature in 50 % acetone.

It is observed that secondary structure is reasonably retained by PfuAmyGT in 10 % acetone, both at room temperature and at 90 °C, but mostly lost in 50 % acetone. These observations explain the activity observed up to 20 % acetone and loss of activity at 50 % acetone.

Similar activity assay was carried out in presence of butanol (Figures 52 and 53) and ethanol (Figures 54 and 55).



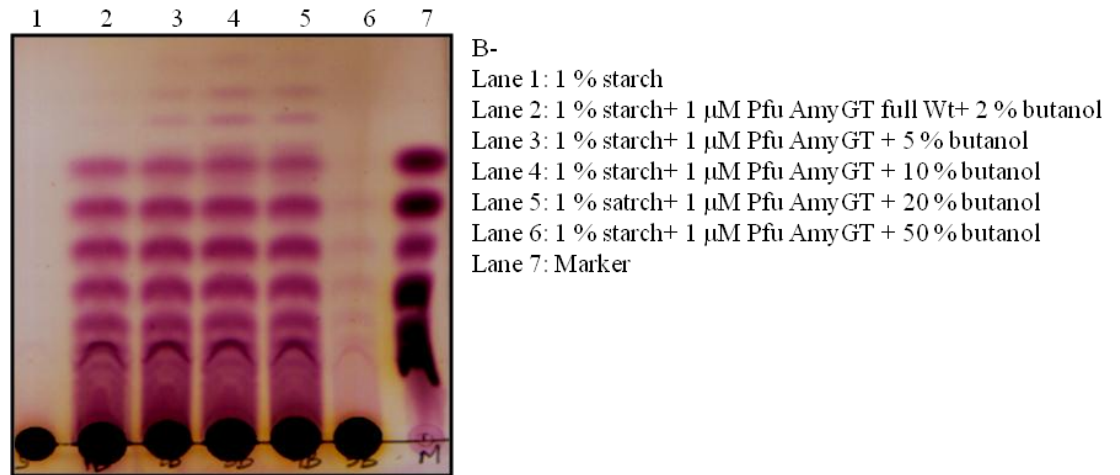


Figure 52: Panel A shows the activity of PfuAmyGT on 20 mM maltotriose in presence of different butanol concentrations and Panel B shows the activity on 1 % starch. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

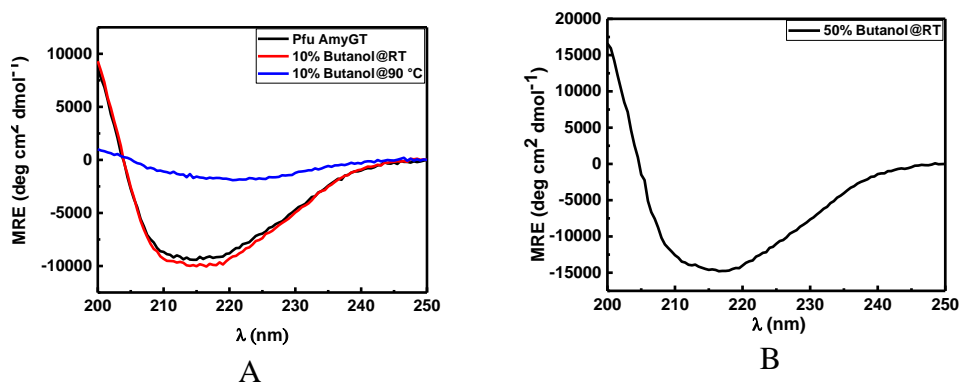


Figure 53: Panel A shows the CD spectra of native PfuAmyGT in presence of 50 mM Tris (pH 8.0) and in the presence of 10 % butanol after an overnight incubation both at room temperature and at 90 °C. Panel B shows the spectrum of PfuAmyGT incubated overnight at room temperature in 50 % butanol.

In case of butanol (Figure 52), activity is abolished in the presence of 50 % butanol and formation of sugar alcohols (more polar than glucose, with more mobility on the TLC) is seen when starch is used as a substrate in the presence of 5, 10 and 20 % butanol. The enzyme's structure does not change significantly after incubation in 10 % butanol at room temperature but decreases when incubated at 90 °C (Figure 53). However, it seems to increase in the presence of 50% butanol, which can be because the protein may reside in the water phase

(butanol-aqueous mixtures may phase-separate in such a long period of time of incubation, even though the incubation was carried out under stirring conditions) but activity can be seen to have almost been abolished.

Similar observations were seen with ethanol; decreased activity at 20 % and 50 % ethanol in case of maltotriose and at 50 % ethanol in case of starch. For starch, sugar alcohol formation is seen at 5, 10 and 20 % ethanol whereas for maltotriose it is seen at 10 % ethanol (Figure 54).

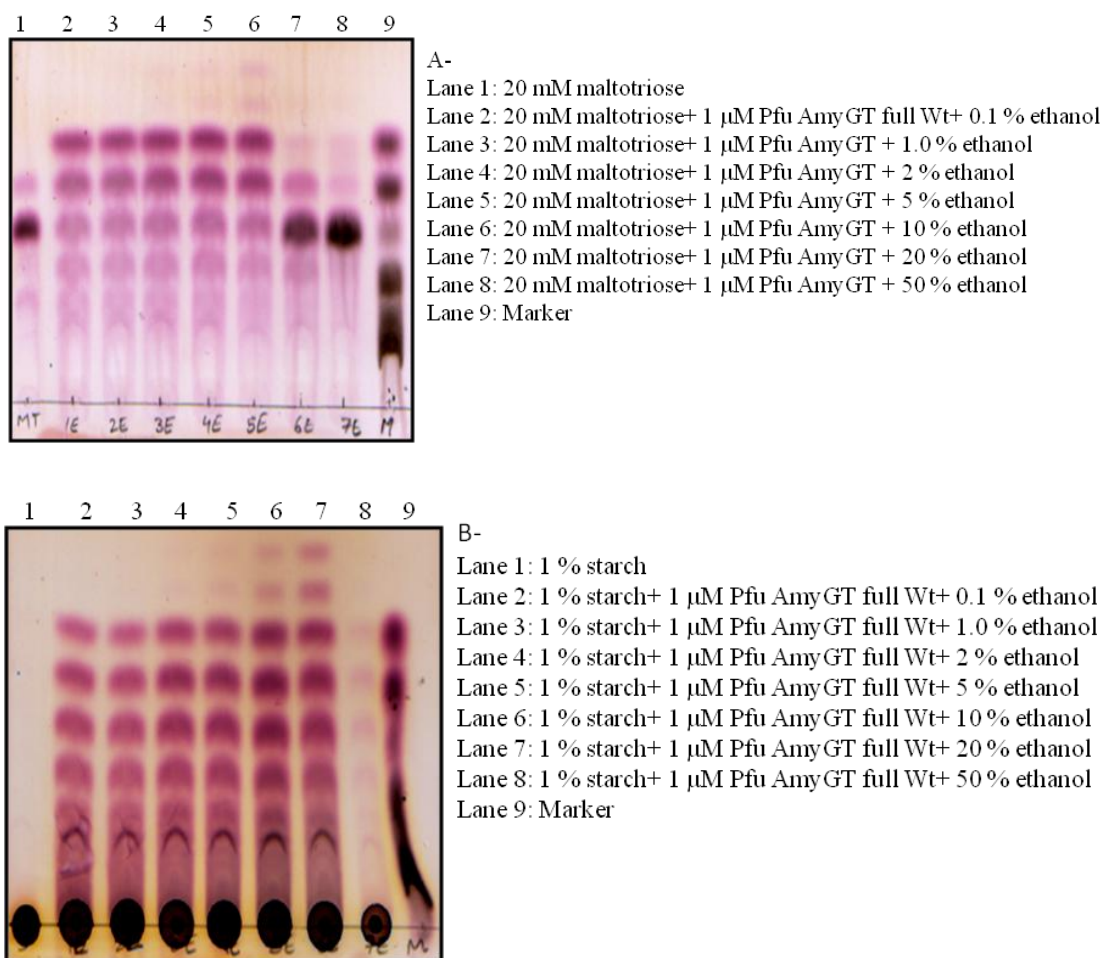


Figure 54: Panel A shows the activity of PfuAmyGT on 20 mM maltotriose in presence of different ethanol concentrations and Panel B shows the activity on 1 % starch. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

This ability of PfuAmyGT can be exploited industrially in biofuel production. Especially in case of ethanol, it is tolerant to 20 % ethanol when starch is used as a substrate and also produces sugar alcohols from it so it can be used

independently or in a cocktail of enzymes for biomass degradation. As seen in Figure 55, the structure of PfuAmyGT is also maintained in presence of ethanol.

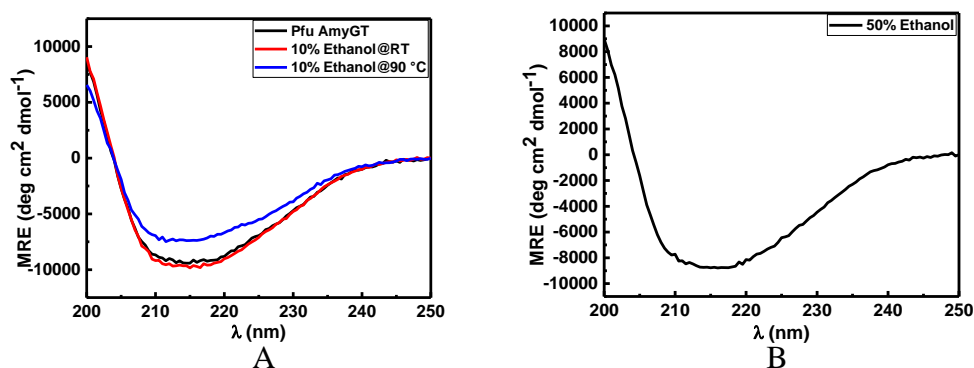
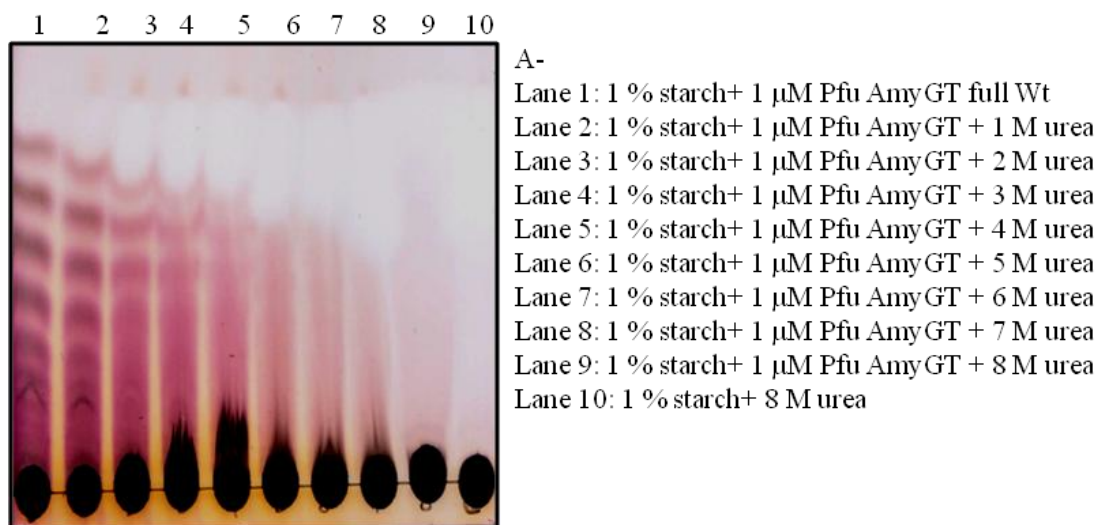


Figure 55: Panel A shows the CD spectra of native PfuAmyGT in presence of 50 mM Tris (pH 8.0) and in the presence of 10 % ethanol after an overnight incubation both at room temperature and at 90 °C. Panel B shows the spectrum of PfuAmyGT incubated overnight at room temperature in 50 % ethanol.

3.5.10 Activity of PfuAmyGT in the presence of denaturants:

The activity of full length, wild-type, PfuAmyGT was observed in the presence of varying concentrations of urea by TLC (Figure 56). When only 1 % starch was used as a substrate significant activity was observed even in the presence of 2 M urea, and to a small extent in 3 M urea, after which it starts to diminish. When 20 mM glucose is also present along with starch, activity is observed up to 4 M urea but a small amount of it is also observed till almost 8 M urea.



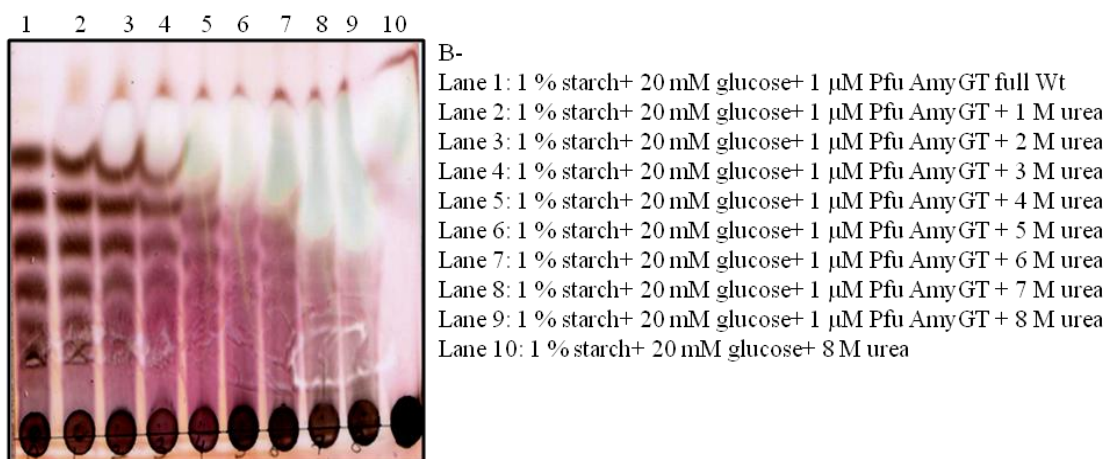


Figure 56: The TLCs show the presence of activity of PfuAmyGT in the presence of urea when only 1 % starch is used as a substrate and when 20 mM glucose is also added along with starch at the start of the reaction. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.5.11 Glucanotransferase activity of the domains of PfuAmyGT:

The glucanotransferase activity of individual domains was also assessed by using 20 mM maltotriose as substrate (since it has been already established to be self-sufficient in generating all oligosaccharide species) in the presence of 1 μ M as well as 5 μ M of each enzyme (Figure 57). In another set of reactions, 20 mM glucose was also added along with maltotriose to examine the activity of each domain. However, no significant activity was observed in any case. With 5 μ M D1, some hint of activity is observed.

Results and Discussion

Transferase activity	Domain
	<p>D1</p> <p>Lane 1: 20 mM maltotriose</p> <p>Lane 2: 20 mM maltotriose+ 20 mM glucose</p> <p>Lane 3: 20 mM maltotriose+ 1 μM D1</p> <p>Lane 4: 20 mM maltotriose+ 5 μM D1</p> <p>Lane 5: 20 mM maltotriose+ 20 mM glucose+ 1 μM D1</p> <p>Lane 6: 20 mM maltotriose+ 20 mM glucose+ 5 μM D1</p> <p>Lane 7: Marker</p>
	<p>D3</p> <p>Lane 1: 20 mM maltotriose</p> <p>Lane 2: 20 mM maltotriose+ 20 mM glucose</p> <p>Lane 3: 20 mM maltotriose+ 1 μM D 3</p> <p>Lane 4: 20 mM maltotriose+ 5 μM D 3</p> <p>Lane 5: 20 mM maltotriose+ 20 mM glucose+ 1 μM D 3</p> <p>Lane 6: 20 mM maltotriose+ 20 mM glucose+ 5 μM D 3</p> <p>Lane 7: Marker</p>
	<p>D1 + D2</p> <p>Lane 1: 20 mM maltotriose</p> <p>Lane 2: 20 mM maltotriose+ 20 mM glucose</p> <p>Lane 3: 20 mM maltotriose+ 1 μM D1+ D2</p> <p>Lane 4: 20 mM maltotriose+ 5 μM D1+ D2</p> <p>Lane 5: 20 mM maltotriose+ 20 mM glucose+ 1 μM D1+ D2</p> <p>Lane 6: 20 mM maltotriose+ 20 mM glucose+ 5 μM D1+ D2</p> <p>Lane 7: Marker</p>

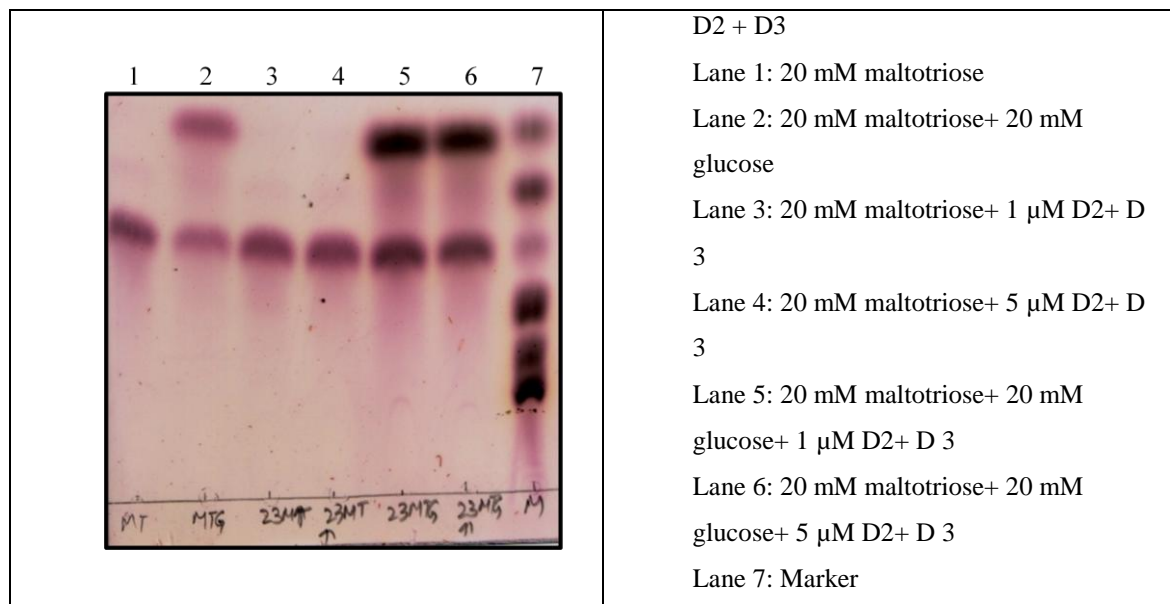


Figure 57: Examination of the glucoamylase activity of the different truncated domains of PfuAmyGT Wt protein. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.6 Identification of putative donor and acceptor sites through structural bio-informatic analysis:

To understand the mechanism by which the transfer of glucose from donor to acceptor molecule takes place, we first predicted (since the crystal structure is not available) the tertiary structure of PfuAmyGT using I-TASSER (Figure 58) and analyzed it using the PyMOL visualization tool.

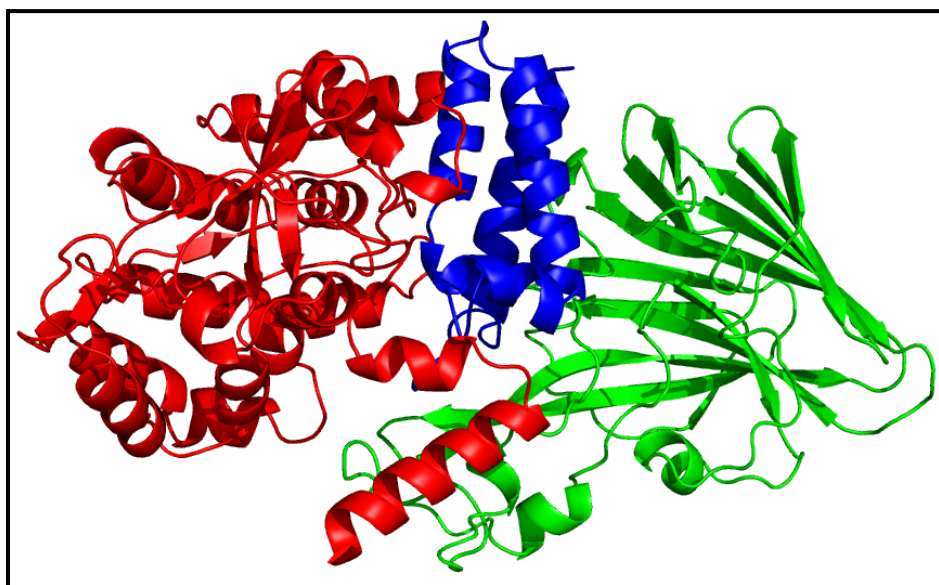


Figure 58: Cartoon representation of the predicted structure of PfuAmyGT.

Results and Discussion

Using the predicted structure, we carried out various docking studies to find out the putative donor and acceptor sites in PfuAmyGT.

For these studies, we utilized the available crystal structure of 4- α -glucanotransferase (GT) (PDB: 1K1Y) from *Thermococcus litoralis*. The amino acid sequences of PfuAmyGT and 4- α -glucanotransferase from *Thermococcus litoralis* share 65 % identity as shown in Figure 59, as assessed using ESript.

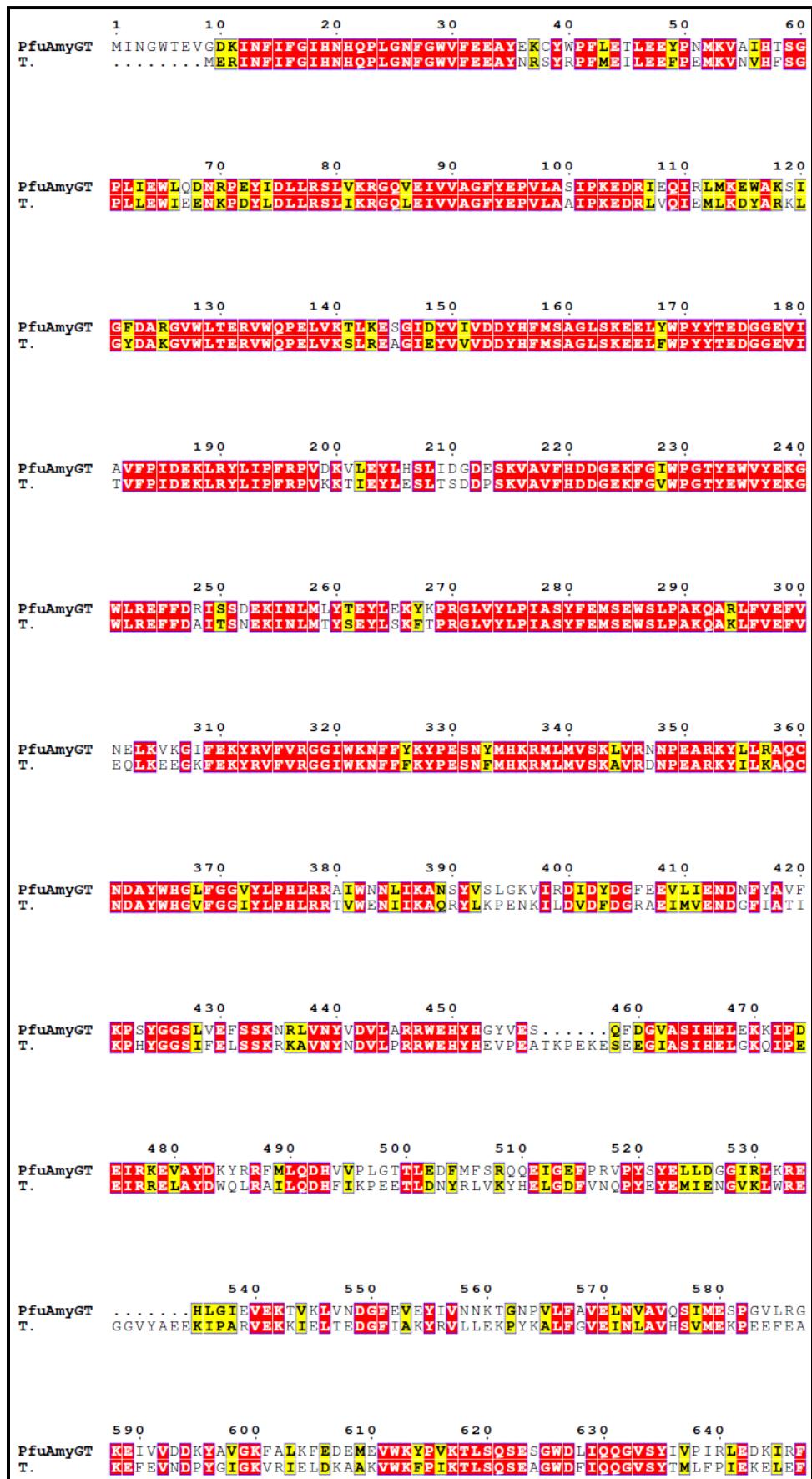


Figure 59: Alignment of PfuAmyGT and GT from *Thermococcus litoralis*.

Results and Discussion

The crystal structure of 4- α -glucanotransferase from *Thermococcus litoralis* was determined by Imamura *et al* in 2003. The overall structure of this protein consists of two domains: the N-terminal domain is a rare (β/α)₇ (pseudo TIM barrel) domain consisting of a cleft like active site, where they also found a bound molecule of acarbose (an inhibitor, discussed later), and the C-terminal domain consists of a β -sandwich fold. The crystal was obtained in the form of a homodimer and a maltose molecule was found bound on chain B, to a site other than the presumed active site at which the acarbose was found bound in chain A, as seen in Figure 60.

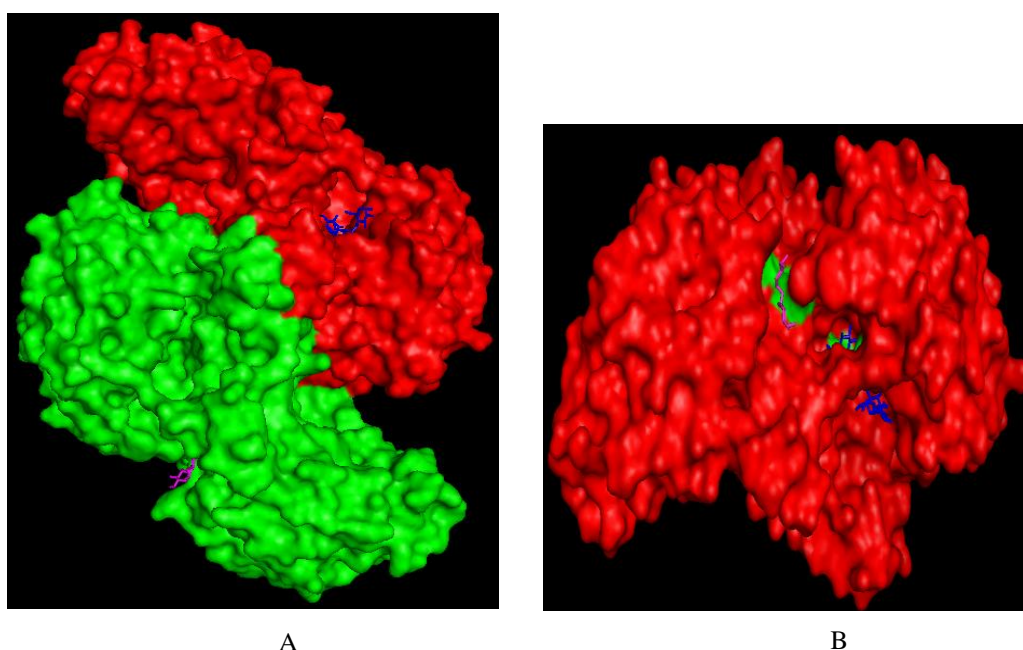


Figure 60: Panel A shows the dimeric (chain A shown in red and chain B shown in green) crystal structure representation of glucanotransferase of *Thermococcus litoralis* in PyMOL. Chain A shows bound acarbose molecule in blue color while chain B shows bound maltose in pink color. Panel B shows the two chains superimposed on each other and shows the putative donor and acceptor sites in a single chain of the molecule.

We superimposed the two chains A and B of *Thermococcus litoralis* upon each other using PyMOL and then also superimposed them on the predicted structure of PfuAmyGT (Figure 61). From the superimposition and careful examination of the structure using PyMOL, we hypothesize that the N-terminal domain of PfuAmyGT contains a tunnel-like donor site where the exo-amylase action takes place and it also contains a cleft like acceptor site (where the maltose was found

bound in the crystal structure from *Thermococcus litoralis*, presumably picked up by the enzyme from the cytoplasm of the organism in which it was made).

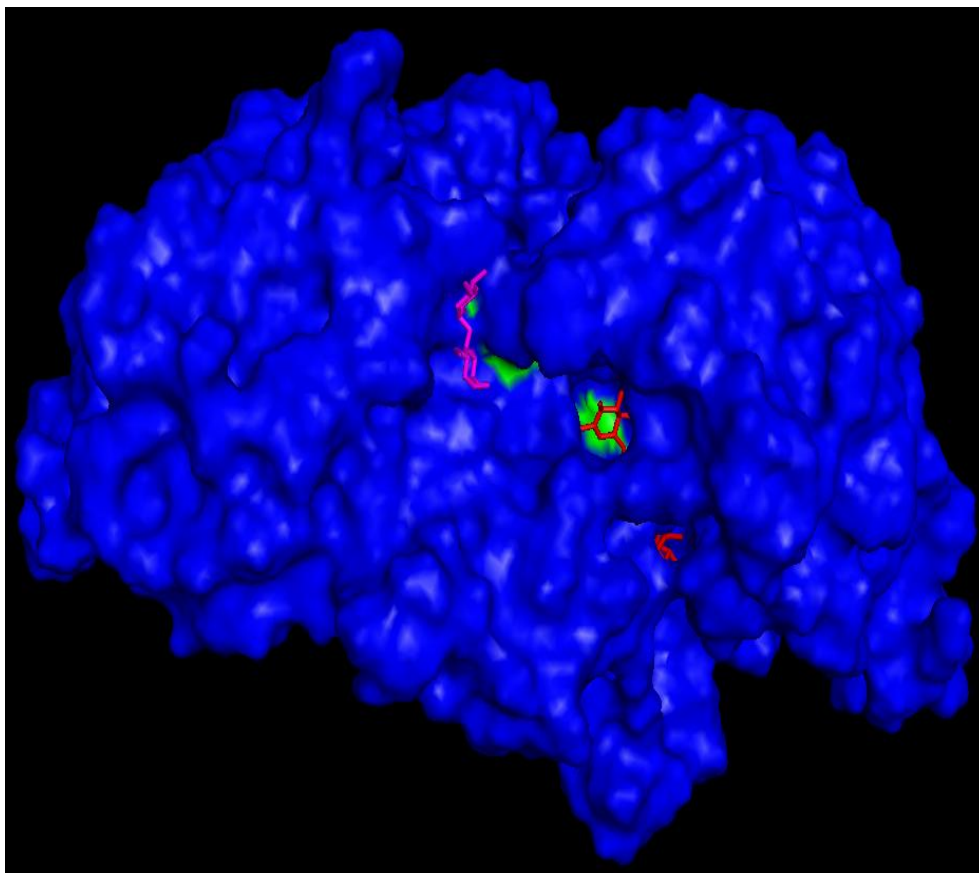


Figure 61: Shows the putative donor (acarbose represented in red) and acceptor (maltose represented in pink) sites of PfuAmyGT.

We also identified the residues interacting with acarbose in the donor site and maltose in the acceptor site of PfuAmyGT. We observed that acarbose was in close vicinity to (and mostly interacting with) glutamate and aspartate residues at positions 131 and 222 respectively while maltose interacts mainly with histidine at position 376, and arginine at 379.

3.7 Glucose is transferred from the head of the donor saccharide to the tail of the acceptor saccharide:

Using PyMOL we then labeled the carbon atoms in the acarbose and maltose. Assuming that maltotetraose is in place of acarbose, it is observed that the glucose is cleaved from the head of the maltotetraose/acarbose molecule, i.e., from the non-reducing end of the glucose polymer (in other words, the first glucose molecule of the four present in the maltotetraose chain is the one which is

cleaved). The transfer of this glucose then occurs to the tail of the maltose molecule, i.e., to its reducing end, as shown in Figure 62. The same would apply to any pair of donor and acceptor glucan, i.e., the single sugar unit of glucose would be cleaved from the non-reducing end of the donor glucan and attached to the reducing end of the acceptor glucan.

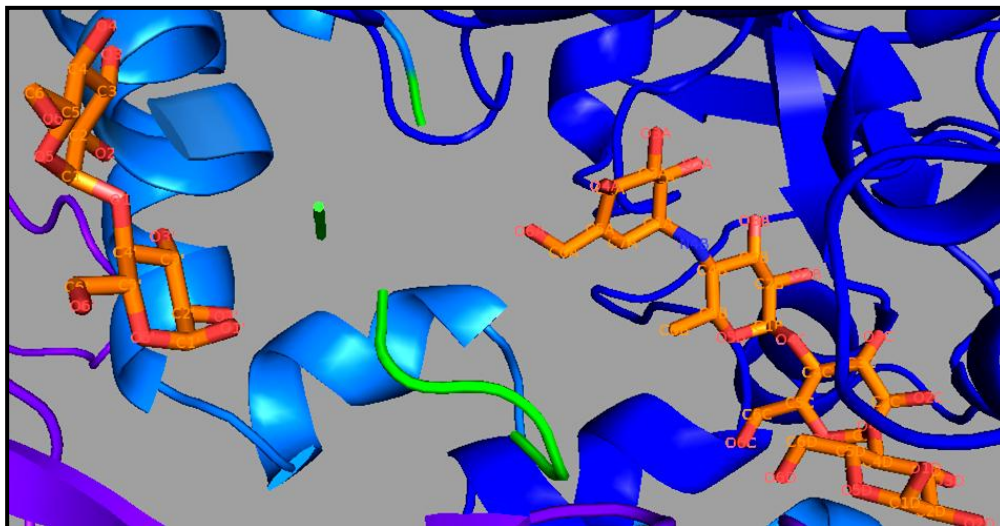


Figure 62: This cartoon representation shows labeled acarbose (a non-hydrolysable analog of maltotetraose) and maltose molecules, where the acarbose inhibits the enzyme by being unable to allow a cleaved glucose to be transferred. The orientation of the two leads us to predict that if there were a maltotetraose bound in place of acarbose, then the cleaved glucose would be transferred from head (non-reducing end) of the donor maltotetraose to the tail (reducing end) of the acceptor maltose.

3.8 The acceptor site is an open groove while the donor site is a buried tunnel:

Glycosidases can also be classified on the basis of their surface topography in the following ways (Davies and Williams 2016):

Endo acting enzymes have active sites in the form of open clefts. This allows the internal glycosidic bonds to be more accessible and susceptible to cleavage.

Exo acting enzymes have their active sites buried in pockets or shallow caves since hydrolysis of glycosidic bonds lying towards the end(s) of a bound chain is required.

Processive enzymes are those in which the polysaccharide chain undergoes multiple catalytic reactions before dissociating from the enzyme, with some relative movement occurring between the enzyme and its bound substrate (presumably involving a ratcheting mechanism, and the alternating breaking and re-making of various non-covalent interactions between the enzyme and the substrate at each state of binding followed by catalysis). The active site in such enzymes is mainly found to be in the form of a tunnel which is formed by extended loop regions that cover the catalytic centers.

Another topology exists, in which the active site/ center is located in a groove blocked at one end.

We aligned the structure of PfuAmyGT with one representative structure from each of the above-mentioned topologies using TM-align. We observed no significant similarity with any of these topologies since all the scores obtained were less than 0.5 even though the putative donor site looks like a tunnel and the putative acceptor site looks like an open cleft or groove (Figure 63).

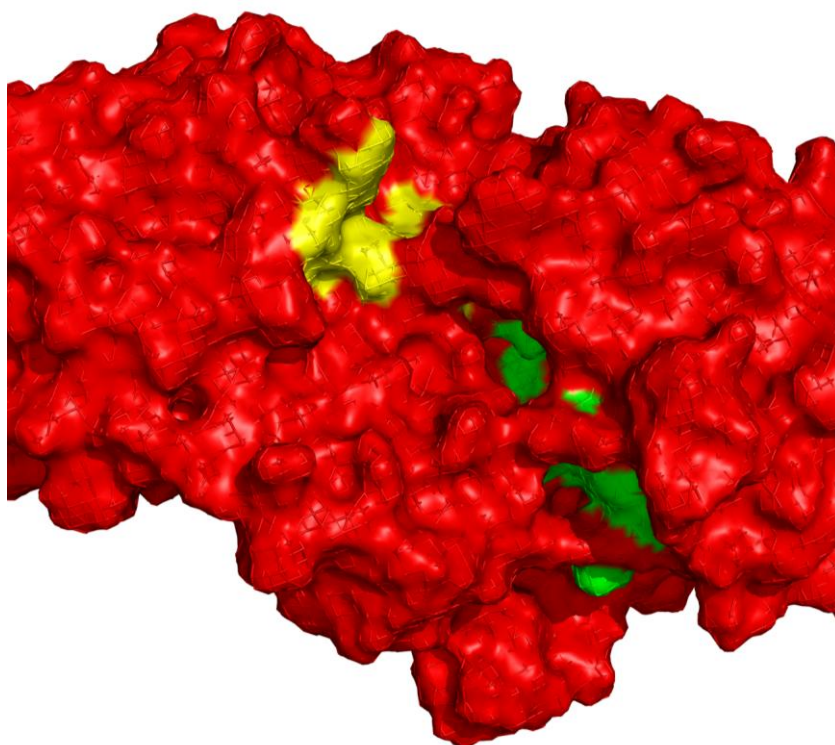


Figure 63: This representation of PfuAmyGT's structure shows the putative donor site to be tunnel-like (residues are shown in green) and putative acceptor site to be cleft-like (residues shown in yellow).

3.9 Acceptor site lies in domain 2 and donor site is in domain 1:

By docking studies, we found the residues interacting with acarbose to be Glutamate-131 and Aspartate-222 and with those of maltose to be Histidine-376 and Arginine-379. This indicates that the donor site mainly lies in domain 1 (D1) while the acceptor lies in domain 2 (D2).

3.10 A tryptophan bearing loop might transfer the glucose:

Leloir enzymes or glycosyltransferases which utilize activated sugar-nucleotide donor molecules to transfer the sugar moiety to other sugar, protein, lipid, nucleotide or even antibiotics show the presence of a flexible loop near the sugar-nucleotide binding site (Lairson *et al* 2008 and Qasba *et al* 2005).

This loop has been traced to be present in two conformations: open when the active site/donor site is not occupied by the substrate and closed when the substrate is bound and acts as a covering lid for the substrate.

We located a similar loop in PfuAmyGT and hypothesized that it might be helping in the transfer of excised glucose from the donor substrate molecule to the acceptor facilitated by a tryptophan residue present in it, as it lies in the close vicinity (within 6-7 Å) of the acarbose molecule as shown in Figure 64.

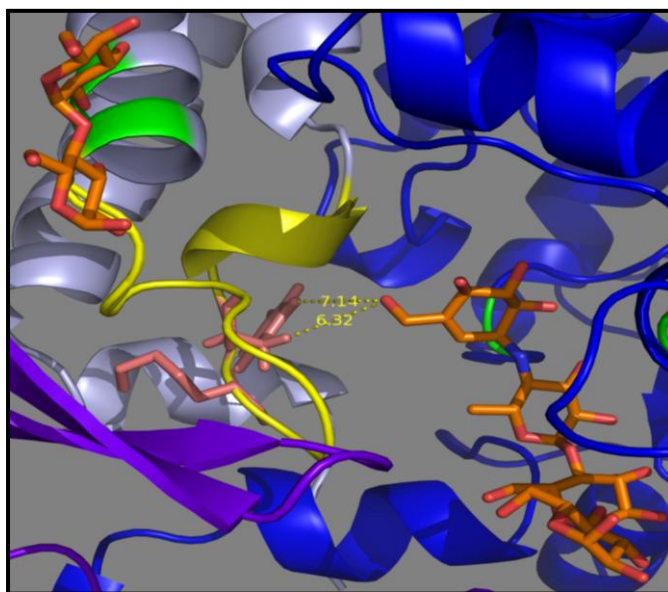


Figure 64: Shows a tryptophan (pink) bearing loop (yellow) with its distance from acarbose that might help in glucose transfer from donor to acceptor molecules.

3.11 Mutational analysis:

A few other residues in the loop were also identified (table 3) in the structure of PfuAmyGT in order to decipher the mechanism followed by it. These residues are mainly distributed in the loop region and also in close vicinity of the acceptor site and have been hypothesized to play important roles in the transfer of glucose from donor to an acceptor molecule, especially, the tryptophan residue which lies in close proximity of acarbose (Figure 65).

Residue identified	Position (aa)
Asp (D)- Loop region	362
Trp (W)- Loop region	365
His (H)- Loop region	366
Leu (L)- Loop region acarbose</td <td>368</td>	368
Pro (P)- Loop region	375

Table 3: It shows the residues that are considered to be important in the glucose transfer.

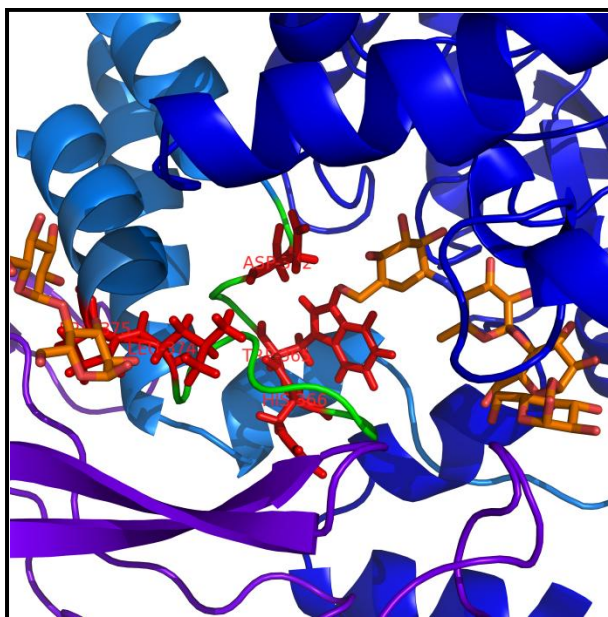


Figure 65: This representation shows that the residues mentioned in table 3 lie in close proximity of donor as well as the acceptor site.

Of these five residues shown in table 3, we mutated Asp, Trp and His to alanine, individually, using SOE-PCR, and cloned the mutated gene(s) into pET-23a between restriction sites, *NdeI* and *XhoI*. After confirmation of the mutations by

sequencing, the expression of the mutated proteins was also analyzed using SDS-PAGE (Figure 66).

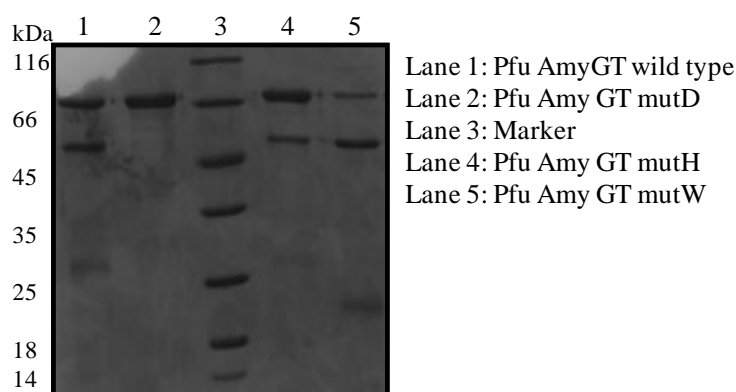


Figure 66: SDS PAGE analysis of PfuAmyGT wild type and the mutants.

All proteins observed were of a correct size and were further used for characterization, and activity assays.

3.11.1 Gel filtration chromatography of mutant forms of PfuAmyGT:

After IMAC purification, the mutants were chromatographically analyzed on a Superdex-200 Increase column pre-equilibrated with 50 mM Tris, pH-8.0. This was done to determine whether there was any change in the oligomeric status of the mutants compared to the wild-type PfuAmyGT. No major change was observed (Figure 67, the first column shows the gel filtration profile while the second column gives the observed elution volume in each case) in comparison with the wild-type PfuAmyGT.

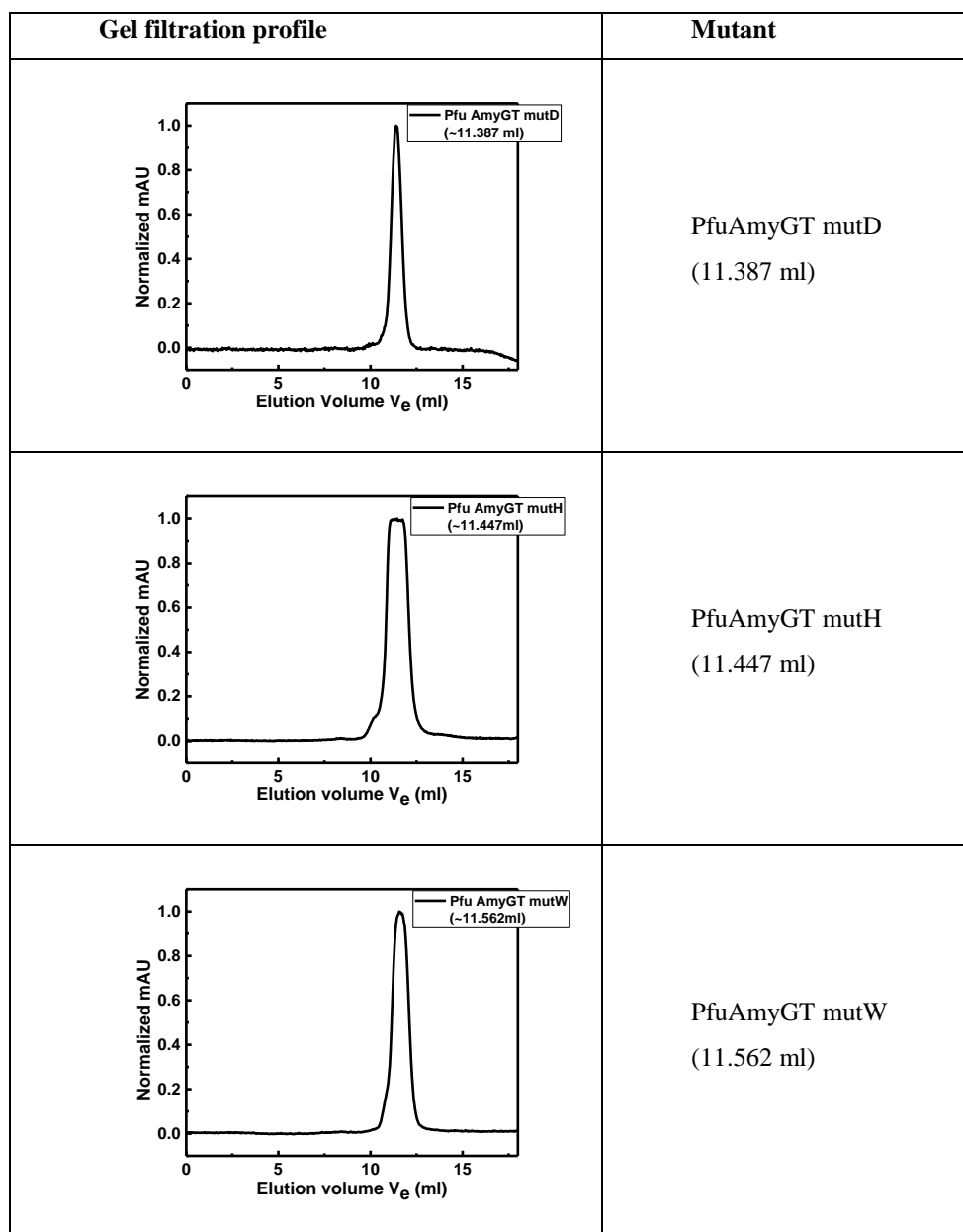


Figure 67: Gel filtration profiles of all mutants of PfuAmyGT.

3.11.2 Determination of the secondary and quaternary structure of the mutants:

The secondary structural contents of all mutants were examined by CD spectrometry and they were all found to have the mixed spectral signatures of proteins containing alpha helices and beta sheets, similar to wild-type PfuAmyGT. To determine the overall quaternary structure of the mutants, their

intrinsic tryptophan fluorescence was exploited and they were all observed to be well folded (Figure 68; column one gives the mutant name, column two represents the CD spectrum whereas column three represents fluorescence spectrum for each mutant).

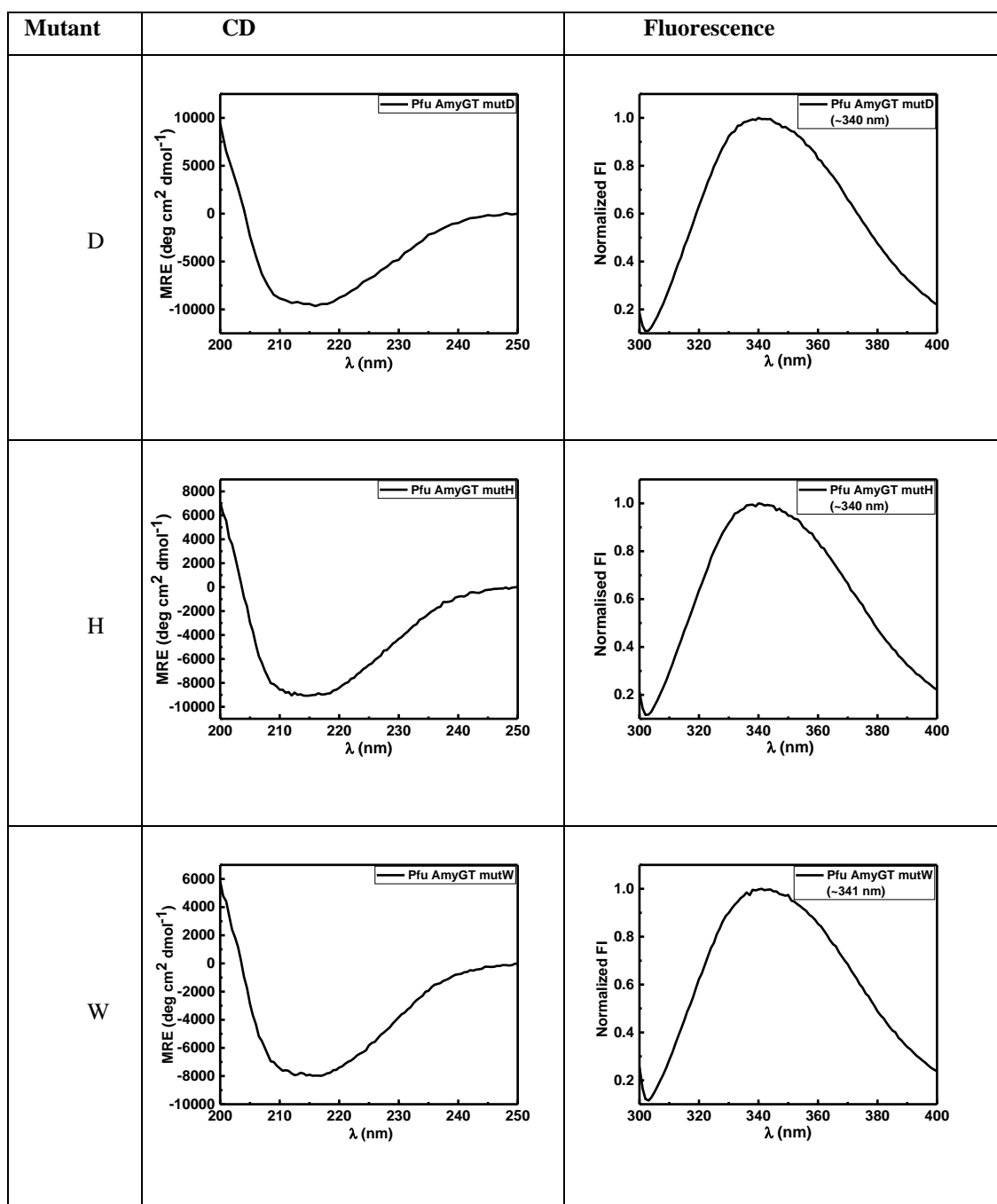


Figure 68: This Figure represents the CD spectra as well as fluorescence emission spectra of all mutants of PfuAmyGT.

3.11.3 Thermal stability of the mutants:

The thermal stability of the mutants was also examined using CD spectral observations by heating an appropriate amount and concentration of the sample from 25 °C to 90 °C at a rate of 5 °C/ minute in a sealed quartz cuvette. All the mutants were observed to be as thermostable as the wild-type PfuAmyGT (Figure 69), in that none of them underwent any significant unfolding, with all mutants showing little or no change in mean residue ellipticity (MRE) as a function of temperature.

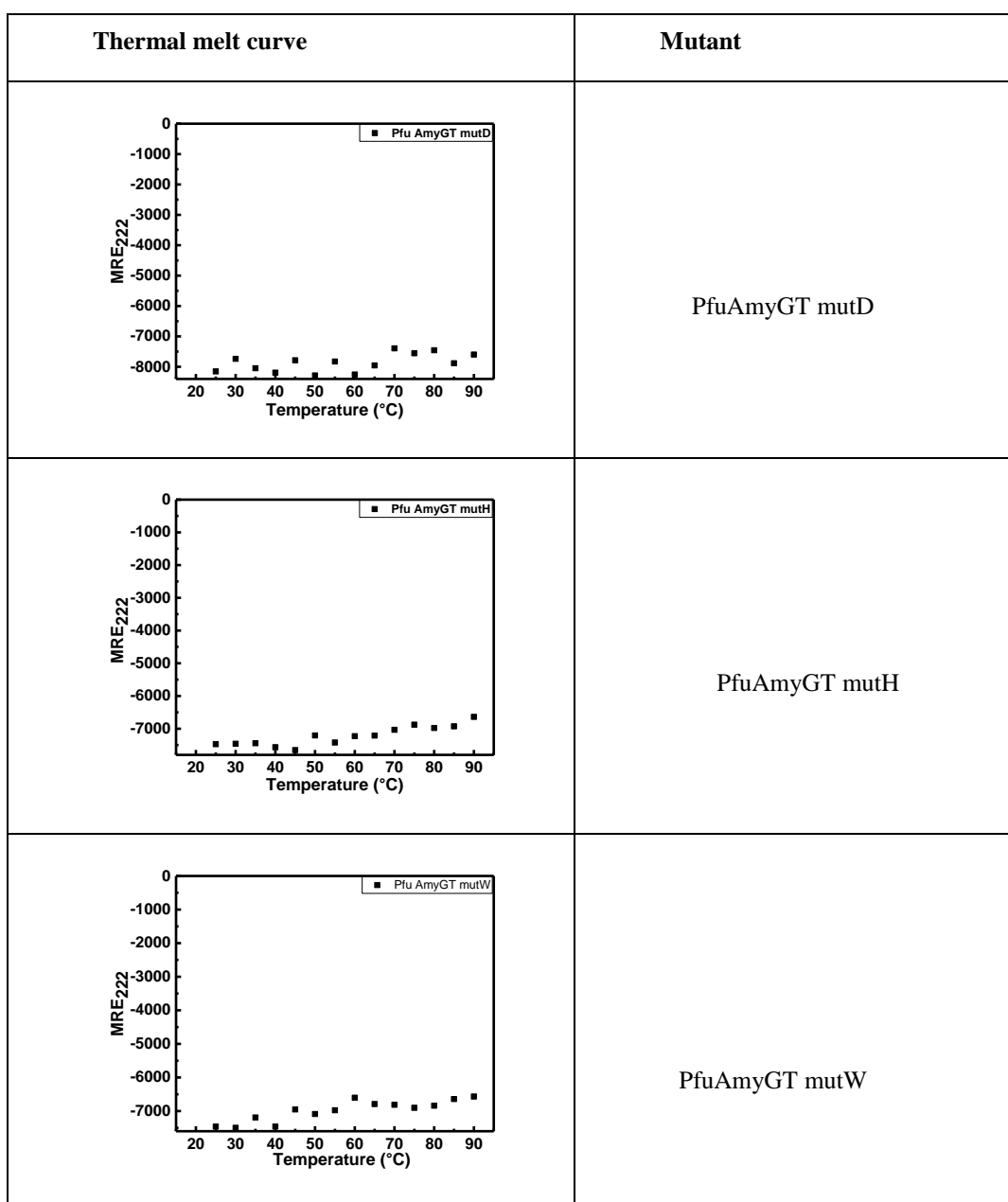


Figure 69: CD spectra showing the thermostable nature of the mutants of PfuAmyGT to be similar to the wild type.

3.11.4 Activity/behavior of the mutants:

The activity of the mutants was examined using 20 mM maltotriose as a substrate, by the TLC method (Figure 70). In the case of the mutD (D362A) and mutW (W365A) mutants, no activity was observed in terms of the formation of different oligosaccharides from maltotriose. In the case of mutD, the maltotriose remained unmodified over a range of enzyme concentrations, in comparison to the wild-type enzyme which generated all species of maltooligosaccharides. In the case of mutW, at higher concentrations of enzyme, there was evidence of the generation of some glucose and maltose from hydrolysis of the maltotriose, but no evidence of the formation of any longer maltooligosaccharides. This suggests that D362 is involved in the exo-amylase action of the enzyme but that W365 is involved in the glucose transfer. We have already seen from the observation that the exo-amylase activity is more efficient when the glucose transfer occurs more efficiently, as in the case of the increase in activity in the presence of maltose, that it may be necessary for the cleaved glucose to be transferred, for there to be more cleavage, and this can explain the poor hydrolysis shown by mutW. In the case of mutH (H366A), we find that there is no discernible effect of the mutation of this histidine residue to alanine, with the mutant being as active as wild-type PfuAmyGT. Thus we can conclude that His residue present in the loop region of PfuAmyGT does not have a significant role in the cleavage or transfer of the excised glucose unit, whereas the Asp residue could be responsible for hydrolysis, and the Trp residue could play a significant role in the transfer of the glucan, as its mutation lead to loss of transferase activity.

Results and Discussion

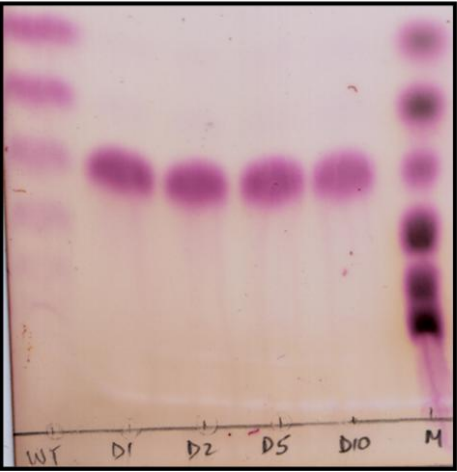
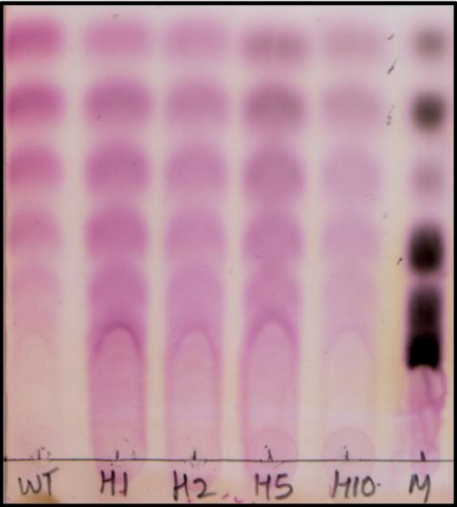
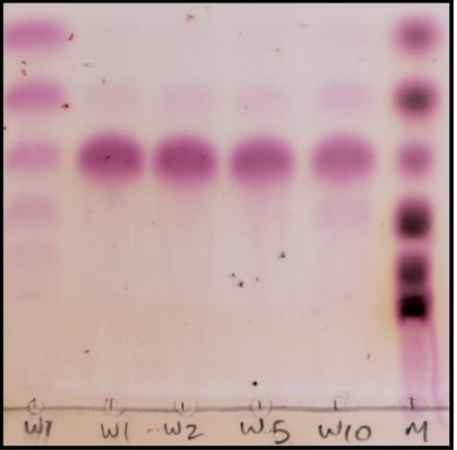
Activity	Mutant
 <p>1 2 3 4 5 6</p> <p>WT D1 D2 D5 D10 M</p>	<p>Lane 1: 20 mM maltotriose+ 1 μM PfuAmyGT wild type</p> <p>Lane 2: 20 mM maltotriose+ 1 μM PfuAmyGT mutD</p> <p>Lane 3: 20 mM maltotriose+ 2 μM PfuAmyGT mutD</p> <p>Lane 4: 20 mM maltotriose+ 5 μM PfuAmyGT mutD</p> <p>Lane 5: 20 mM maltotriose+ 10 μM PfuAmyGT mutD</p> <p>Lane 6: Marker (The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose)</p>
 <p>1 2 3 4 5 6</p> <p>WT H1 H2 H5 H10 M</p>	<p>Lane 1: 20 mM maltotriose+ 1 μM PfuAmyGT wild type</p> <p>Lane 2: 20 mM maltotriose+ 1 μM PfuAmyGT mutH</p> <p>Lane 3: 20 mM maltotriose+ 2 μM PfuAmyGT mutH</p> <p>Lane 4: 20 mM maltotriose+ 5 μM PfuAmyGT mutH</p> <p>Lane 5: 20 mM maltotriose+ 10 μM PfuAmyGT mutH</p> <p>Lane 6: Marker</p>
 <p>1 2 3 4 5 6</p> <p>WT W1 W2 W5 W10 M</p>	<p>Lane 1: 20 mM maltotriose+ 1 μM PfuAmyGT wild type</p> <p>Lane 2: 20 mM maltotriose+ 1 μM PfuAmyGT mutW</p> <p>Lane 3: 20 mM maltotriose+ 2 μM PfuAmyGT mutW</p> <p>Lane 4: 20 mM maltotriose+ 5 μM PfuAmyGT mutW</p> <p>Lane 5: 20 mM maltotriose+ 10 μM PfuAmyGT mutW</p> <p>Lane 6: Marker</p>

Figure 70: The TLCs in this Figure shows the activities of mutants of PfuAmyGT.

3.12 Acarbose inhibits the glucanotransferase activity of full-length PfuAmyGT:

Acarbose is a pseudotetrasaccharide which has a pseudo sugar ring at its non-reducing end [4,5,6-trihydroxy-3-(hydroxymethyl)-2-cyclohexene-1-yl] linked to the nitrogen of 4-amino-4,6-dideoxy-D-glucopyranose that is further linked via an α -(1,4)-glycosidic linkage to maltose. Acarbose was used as an inhibitor in this study, as it is known that acarbose acts as an inhibitor of some glycosyl hydrolases (Kim *et al* 1999). It is known to be a nanomolar inhibitor of the intestinal α -glucosidase domains and pancreatic α -amylase (Gloster and Vocadlo 2012). In the case of glucanotransferases, its effect is mainly reported on CGTases as a non-competitive inhibitor in some species (*Bacillus circulans* E192 CGTases) or as a competitive inhibitor (*Bacillus* sp.1011 CGTase) in others (Kim M *et al* 1999).

From previous studies, acarbose is known to be bound at the active site of the enzyme (Imamura *et al* 2003) and it is known that it cannot be hydrolyzed by 4- α -glucanotransferases. So, in our case acarbose seems to be binding to the hydrolytic center (catalytic center responsible for exo-amyolytic activity) of the enzyme where the donor glucan binds, and inhibiting hydrolysis competitively, i.e. competing with maltotriose, in the experiment shown in Figure 71A, for binding at the donor site, thus reducing the amount of glucose produced which can then be transferred. When only 20 mM maltotriose is used as a substrate, inhibition of activity is seen in presence of 10 mM acarbose. In lanes 4 and 5 of Figure 71A the TLC band of only acarbose is seen, as it is a non-hydrolyzable analog of maltotetraose. The band of no other oligosaccharide species is seen. When glucose is already added to the reaction mixture we understand that acarbose may also be binding to the acceptor site in addition to the donor site and preventing the formation of an oligosaccharide species (Figure 71B).

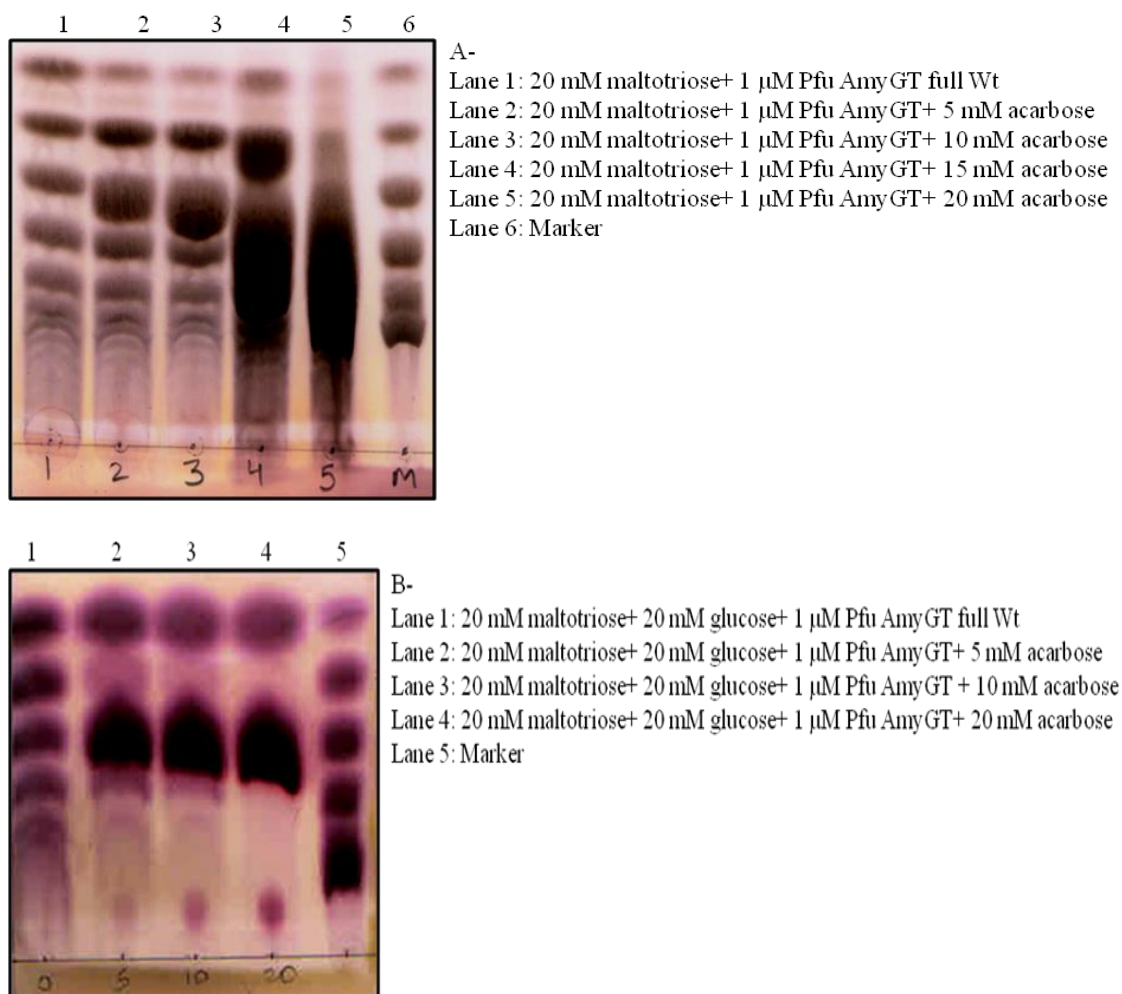


Figure 71: The above TLCs show the effect of acarbose on the action of PfuAmyGT. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.13 Versatility of full-length PfuAmyGT's glucanotransferase activity:

3.13.1 PfuAmyGT works on sucrose as an acceptor:

As mentioned above, we have tested the enzymatic action of PfuAmyGT on starch which is an α -1,4 linked glucan. We have tried several other donor molecules composed of monosaccharides other than glucose, and which are linked by different glycosidic bonds.

In the case of sucrose, which is composed of glucose and fructose, we see the formation of oligosaccharides when 1 μ M enzyme acts upon 1 % sucrose in the presence of 20 mM maltose as well as 20 mM maltotriose. These oligosaccharide

Results and Discussion

species do not run alongside the markers we have used in the TLC (Figure 72), but rather run with a phasing-out of mobility, as bands with mobilities intermediate to those of the marker oligosaccharides. We assume that this is because the products formed are through progressive addition of glucose units to a fructose/glucose-fructose anchor. Since fructose has a different polarity (but the same molecular weight) as compared to glucose, the shift is seen in the mixture of product oligosaccharides in terms of how they move in the solvent phase used for TLC.

When sucrose is used in the presence of starch (here sucrose is assumed to be acting as an acceptor molecule) formation of fructo-oligosaccharides is observed.

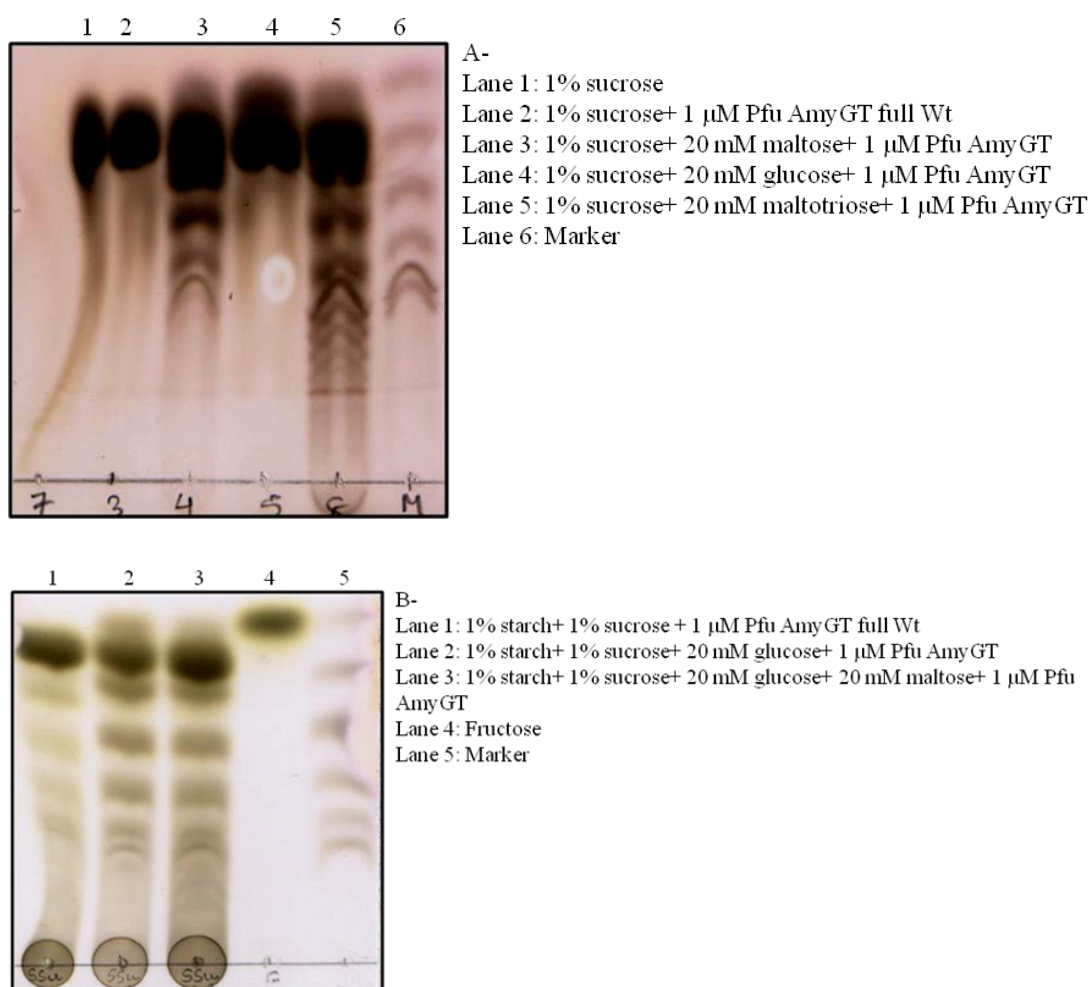


Figure 72: The TLCs in this Figure depicts that sucrose can be used both as donor and acceptor. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

These results were further confirmed by mass spectrometry which will be described in a section, later.

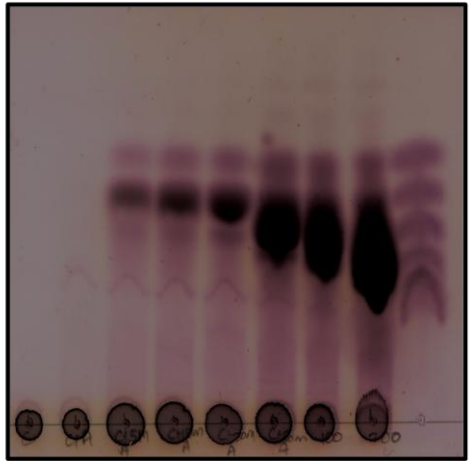
3.13.2 Cellulose, Pectin, and Xylan also act as glucose/xylose donors:

Cellulose is the most abundant component of biomass, composed of β -1,4 linked glucose units. When Carboxy Methyl Cellulose (CMC) is used as a substrate for PfuAmyGT it acts as a donor in the presence of maltose and formation of few oligosaccharide species is seen. However, at higher concentrations of maltose, species smaller than glucose are also observed which are assumed to be sugar alcohols of greater polarity than sugars like glucose, in that they have additional hydroxyl groups.

Pectin is a heteropolysaccharide rich in galacturonic acid present in plant cell walls. When this is used as a donor molecule formation of different oligosaccharide species is seen. When 5 mM maltose is used along with it as acceptor the same species are observed.

Xylan belongs to the class of hemicelluloses composed of xylose units linked via β -1,4 glycosidic bonds. Formation of oligosaccharide species is seen in presence of higher concentration of the acceptor molecule, i.e. where 50 mM maltose is present in the reaction mixture (Figure 73; column one shows the TLC and column two gives the legends).

All the reactions were carried out at 90 °C for 12 hours.

Activity	Donor molecule
<div style="text-align: center;"> <p>1 2 3 4 5 6 7 8 9</p>  </div>	<p>Cellulose</p> <p>Lane 1: 1 % CMC</p> <p>Lane 2: 1 % CMC+ 1 μM PfuAmyGT full Wt</p> <p>Lane 3: 1 % CMC+ 1 μM PfuAmyGT+ 5 mM maltose</p> <p>Lane 4: 1 % CMC+ 1 μM PfuAmyGT+ 10 mM maltose</p> <p>Lane 5: 1 % CMC+ 1 μM PfuAmyGT+ 20 mM maltose</p> <p>Lane 6: 1 % CMC+ 1 μM PfuAmyGT+ 50</p>

Results and Discussion

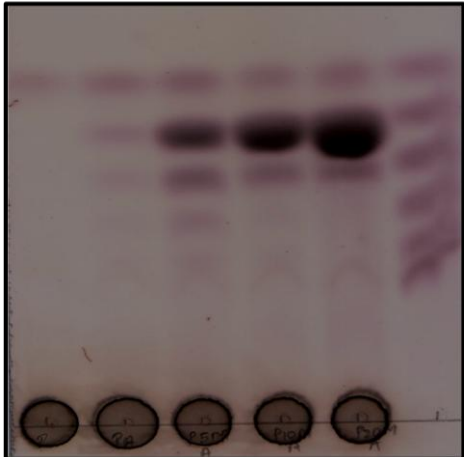
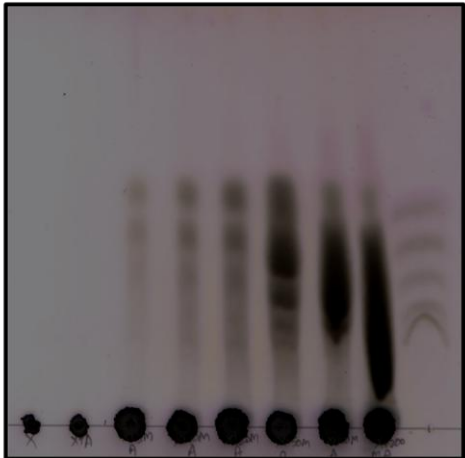
	<p>mM maltose</p> <p>Lane 7: 1 % CMC+ 1 μM PfuAmyGT+ 100 mM maltose</p> <p>Lane 8: 1 % CMC+ 1 μM PfuAmyGT+ 200 mM maltose</p> <p>Lane 9: Marker</p>
	<p>Pectin</p> <p>Lane 1: 1 % pectin</p> <p>Lane 2: 1 % pectin+ 1 μM PfuAmyGT full Wt</p> <p>Lane 3: 1 % pectin+ 1 μM PfuAmyGT+ 5 mM maltose</p> <p>Lane 4: 1 % pectin+ 1 μM PfuAmyGT+ 10 mM maltose</p> <p>Lane 5: 1 % pectin+ 1 μM PfuAmyGT+ 20 mM maltose</p> <p>Lane 6: Marker</p>
	<p>Xylan</p> <p>Lane 1: 1 % xylan</p> <p>Lane 2: 1 % xylan+1 μM PfuAmyGT full Wt</p> <p>Lane 3: 1 % xylan +1 μM PfuAmyGT+ 5 mM maltose</p> <p>Lane 4: 1 % xylan +1 μM PfuAmyGT+ 10 mM maltose</p> <p>Lane 5: 1 % xylan +1 μM PfuAmyGT+ 20 mM maltose</p> <p>Lane 6: 1 % xylan +1 μM PfuAmyGT+ 50 mM maltose</p> <p>Lane 7: 1 % xylan +1 μM PfuAmyGT+ 100 mM maltose</p> <p>Lane 8: 1 % xylan +1 μM PfuAmyGT+ 200 mM maltose</p> <p>Lane 9: Marker</p>

Figure 73: The TLCs in this figure shows the action of full-length and wild-type PfuAmyGT on various donor molecules in the presence of maltose.

3.13.3 Mass spectrometric analysis of products formed:

After completion of the intended duration of the reaction, all the reaction mixtures mentioned above were diluted 10 to 100 times using mass spectrometry grade water and subjected to MALDI-TOF analysis. Since MALDI is a soft ionization technique, the reactions were carried out in the sodium-phosphate buffer, pH 8.0. This helps in the formation of sodium adducts of the products formed which can be easily detected. The mass spectrometric analysis was done to confirm the nature of the products visualized on TLCs. Table 4 shows the calculated and observed masses of products that can be observed in any reaction by the action of the PfuAmyGT full-length, wild-type as well as all its variants.

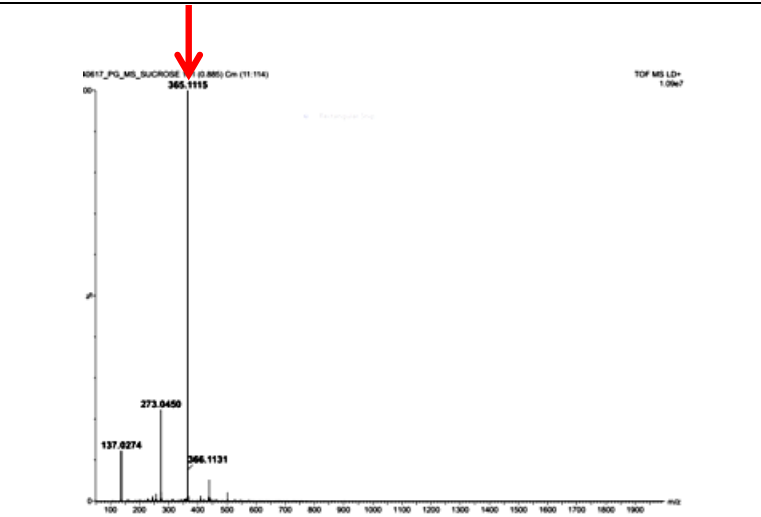
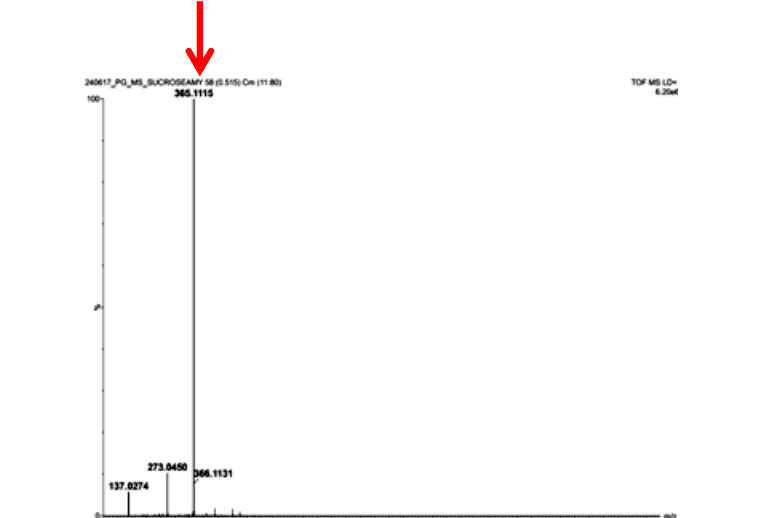
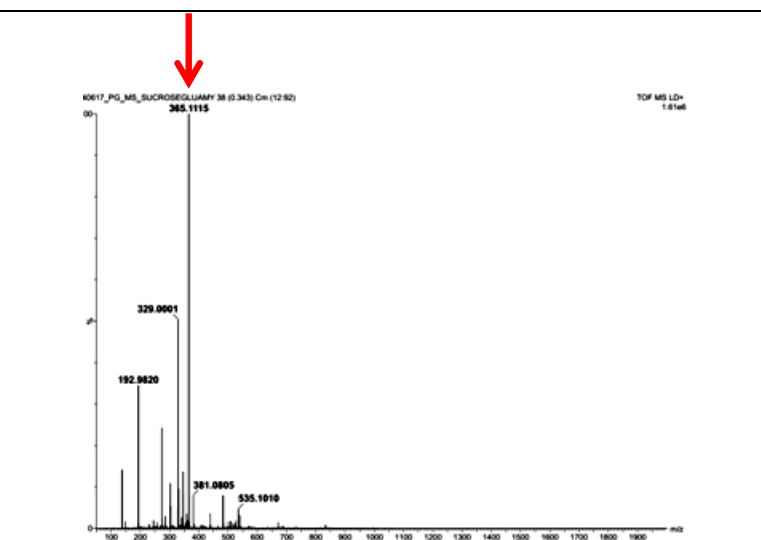
Product formed	Calculated molecular weight	The observed molecular weight of sodium adduct
Glucose	180.00	203.0
Maltose/ Sucrose	342.00	365.0
Maltotriose	504.00	527.0
Maltotetraose	666.50	689.5
Maltopentaose	828.70	851.7
Maltohexaose	990.80	1013.8
Maltoheptaose	1153.00	1176.0
Maltooctaose	1477.00	1500.0

Table 4: Masses of products obtained/observed after the action of PfuAmyGT.

Figure 74 (column one shows the mass spectrum and column two details the contents of the reaction mixture) below shows few representative mass spectra showing the molecular weight of sodium adducts of products formed by the action of PfuAmyGT. The red arrows in the mass spectra point at the observed (sodium adducts) mass peaks in each reaction.

Mass Spectrum	Reaction
	<p>20 mM glucose+ 20 mM maltotriose+1 μM PfuAmyGT</p>
	<p>1 % starch+ 20 mM glucose+1 μM PfuAmyGT</p>

Results and Discussion

 <p>Mass spectrum showing relative intensity versus m/z for 1% sucrose. The base peak is at m/z 365.1115, indicated by a red arrow. Other labeled peaks are at m/z 137.0274, 273.0450, and 366.1131. The plot title is 'K0617_PG_MS_SUCROSE 1 (0.885) Ch (11.154)' and 'TOF MS LD= 1.00e7'.</p>	<p>1 % sucrose</p>
 <p>Mass spectrum showing relative intensity versus m/z for 1% sucrose + 1 μM PfuAmyGT full Wt. The base peak is at m/z 365.1115, indicated by a red arrow. Other labeled peaks are at m/z 137.0274, 273.0450, and 366.1131. The plot title is '240617_PG_MS_SUCROSEAMY 58 (0.515) Ch (11.80)' and 'TOF MS LD= 6.25e4'.</p>	<p>1 % sucrose+ 1 μM PfuAmyGT full Wt</p>
 <p>Mass spectrum showing relative intensity versus m/z for 1% sucrose + 20 mM glucose + 1 μM PfuAmyGT full Wt. The base peak is at m/z 365.1115, indicated by a red arrow. Other labeled peaks are at m/z 192.0820, 329.0001, 361.0805, and 535.1010. The plot title is 'K0617_PG_MS_SUCROSEGLUAMY 38 (0.342) Ch (12.92)' and 'TOF MS LD= 1.61e6'.</p>	<p>1 % sucrose+ 20 mM glucose+ 1 μM PfuAmyGT full Wt</p>

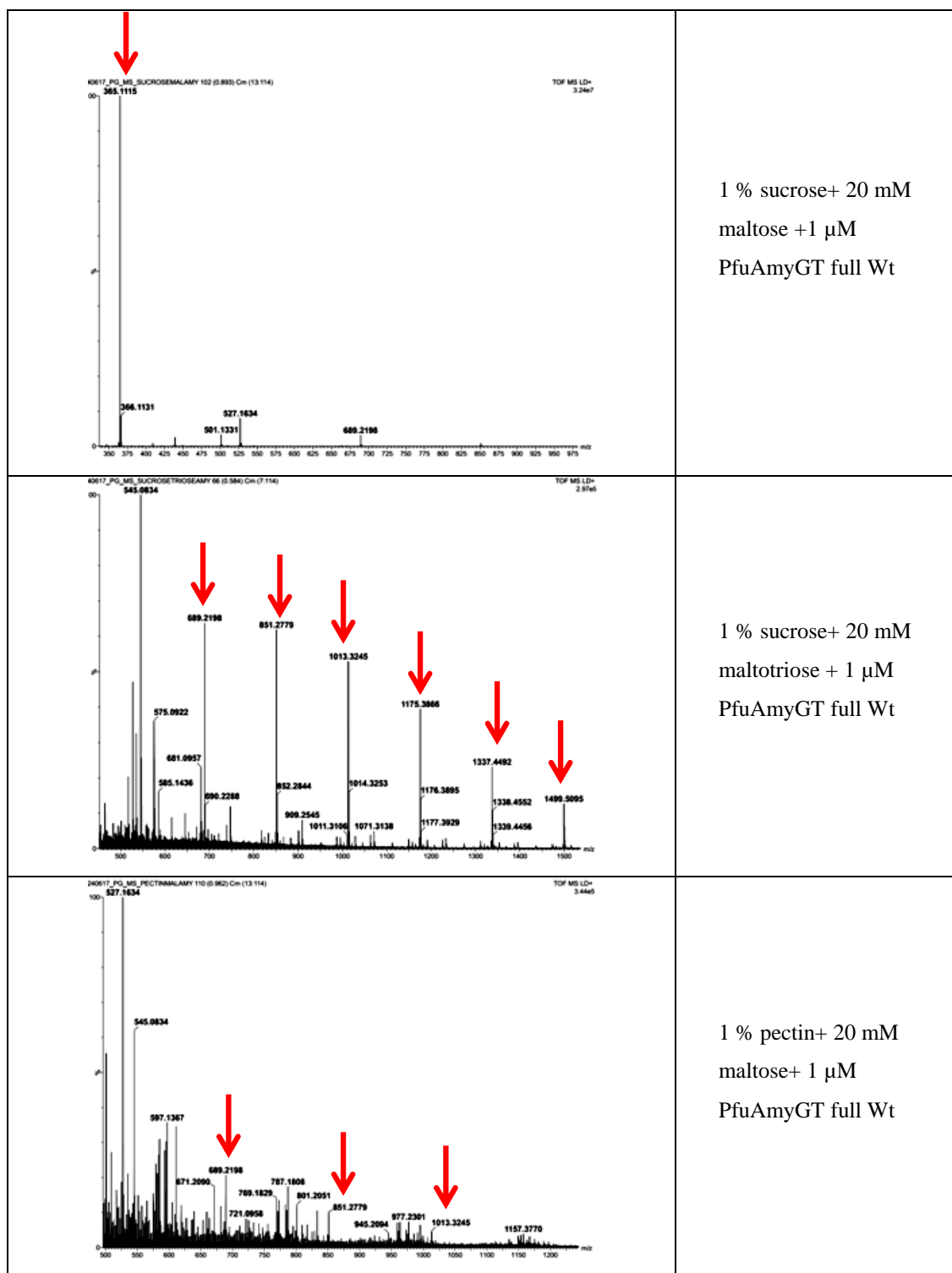


Figure 74: Mass spectra showing the observed molecular weights of various products obtained by the action of PfuAmyGT on different donor molecules.

3.14 Complementation studies

On the basis of sequence and/or structure, proteins are often divided into domains in order to accurately predict their molecular function(s). A typical view of a multidomain protein is that the overall function of the protein is the sum of its constituent domains (Ponting and Russell 2002). A protein domain is generally defined as a spatially distinct unit in the structure; it is often associated with a specific experimentally known function and from an evolutionary perspective is also present in other protein sequences and/or structure.

In order to predict domains in PfuAmyGT, we used the Pfam database. It predicted three domains in our protein as shown already in Figure 13 and now summarized in Table 5.



Source	Domain	Start	End	Gathering threshold (bits)		Score (bits)		E-value	
				Sequence	Domain	Sequence	Domain	Sequence	Domain
Pfam	GH57	8	292	20.60	20.60	228.90	227.70	1.4e-64	3.4e-64
Pfam	DUF1925	312	385	21.10	21.10	129.90	129.90	4.7e-35	9.3e-35
Pfam	DUF1926	393	644	19.90	19.90	245.40	244.00	1.2e-69	3.3e-69

Table 5: Summary of the analysis of different domains of full-length PfuAmyGT as done using the Pfam database.

The Pfam database predicted three domains in our protein; (I) belonging to Glycosyl Hydrolase family 57, (II) belonging to Domain of Unknown Function 1925 having an immunoglobulin/albumin-binding domain-like fold, with a bundle of three alpha-helices but its function is yet unknown, (III) DUF 1926 having a beta-sandwich fold, in which two layers of anti-parallel beta-sheets are arranged in a nearly parallel fashion.

The scores of matching are also significant: E value is the significance of the match and should be less than 1, bit score of zero implies that the match has occurred by chance and hence the greater the bit score the more significant is the match.

In light of this information, we cloned the individual domains separately as well as in fusion with each other and checked for activity in each case as discussed earlier. Taking a step ahead in this direction, we attempted domain complementation studies for which we tried co-expression as well as co-refolding:

3.14.1 Co-expression:

We co-transformed the vectors containing PfuAmyGT domain one (D1) and recombinantly fused domain 2 + 3 (D2+D3) into a single *E. coli* cell.

For this, we also cloned domain 1 (D1) in vector pET-28a which would yield an N terminally 6X-His-tagged protein and then co-transformed it with fused domains D2 and D3 (cloned in pET-23a to yield a C-terminally 6X-His-tagged protein) into BL21 Star(DE3)pLysS cells. This was done to achieve double antibiotic-based selection since pET-28a has kanamycin based selection requirements and pET-23a has ampicillin based selection requirements, to maintain selection pressure in order to prevent plasmid curing. We could see the expression of both proteins in the elution fraction in SDS PAGE (Figure 75) but on running a native PAGE (Figure 75B) gel, we could not observe a single band corresponding to full-length PfuAmyGT, which would have implied interaction between the separately-produced D1 and D2 + D3, and hence a successful complementation of domains.

Thus, we could successfully express both the domains within a single cell, but they did not seem to interact with each other (which could be due to a difference in folding, in comparison with full-length PfuAmyGT).

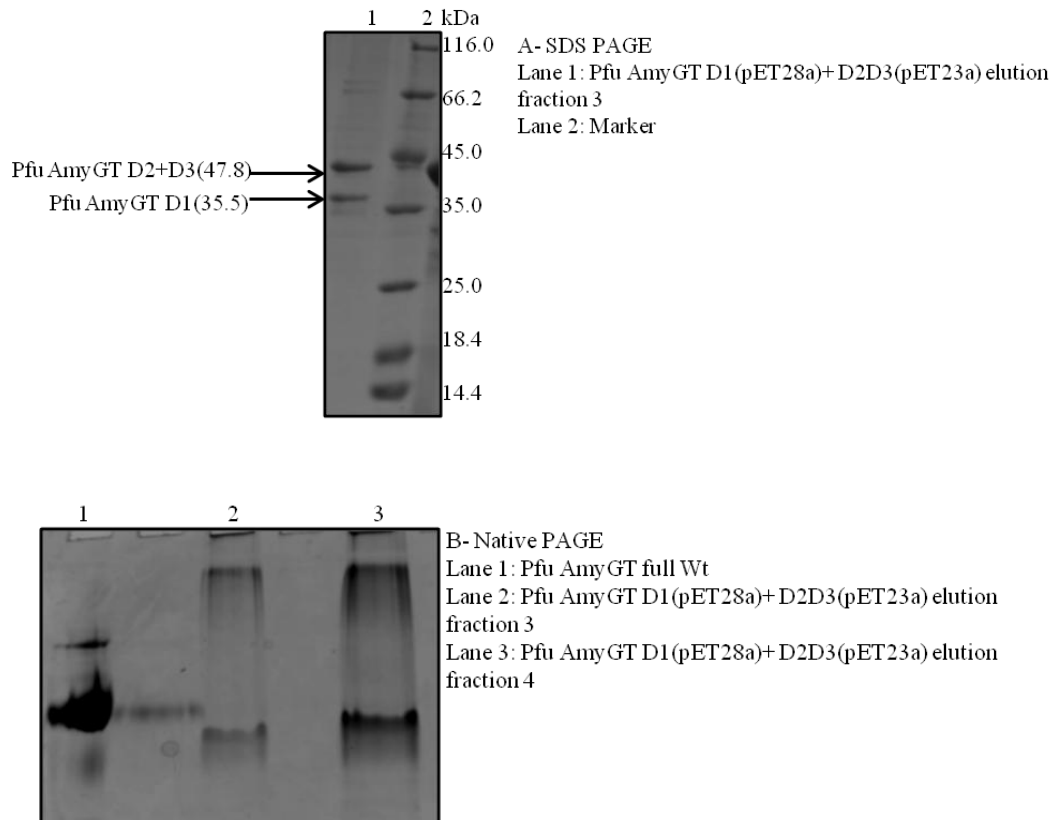


Figure 75: Panel A shows the expression of both domains on SDS PAGE. Panel B shows native PAGE gel of the same protein.

We also created an untagged version of domain 1 (D1) and co-purified it using Ni-NTA chromatography with the two domain fusion, D2 + D3, using the rationale that since the fused protein D2+ D3 contains a 6X-His tag it will bind to the column but domain 1 will only bind (and be seen in elution fraction) if it interacts non-covalently with D2 + D3, since D1 has no affinity tag of its own. After SDS PAGE analysis of this purification, we observed bands for D1 and D2 + D3 in the elution fraction which indicated an interaction between the domains. However, on native PAGE we did not observe a single band.

3.14.2 Co-refolding:

We also tried co-refolding of domain 1 + 2 (D1+ D2) with domain 3 (D3) by completely unfolding the proteins separately. Then they were mixed in equimolar fractions and given conditions to refold together. We thought that this method would give the domains a chance to fold in a similar manner as they are folded in the full-length PfuAmyGT. Similar results (Figure 76) were observed in this case

also where bands of individual domains were seen on SDS PAGE but no band corresponding to full-length PfuAmyGT was observed on native PAGE.

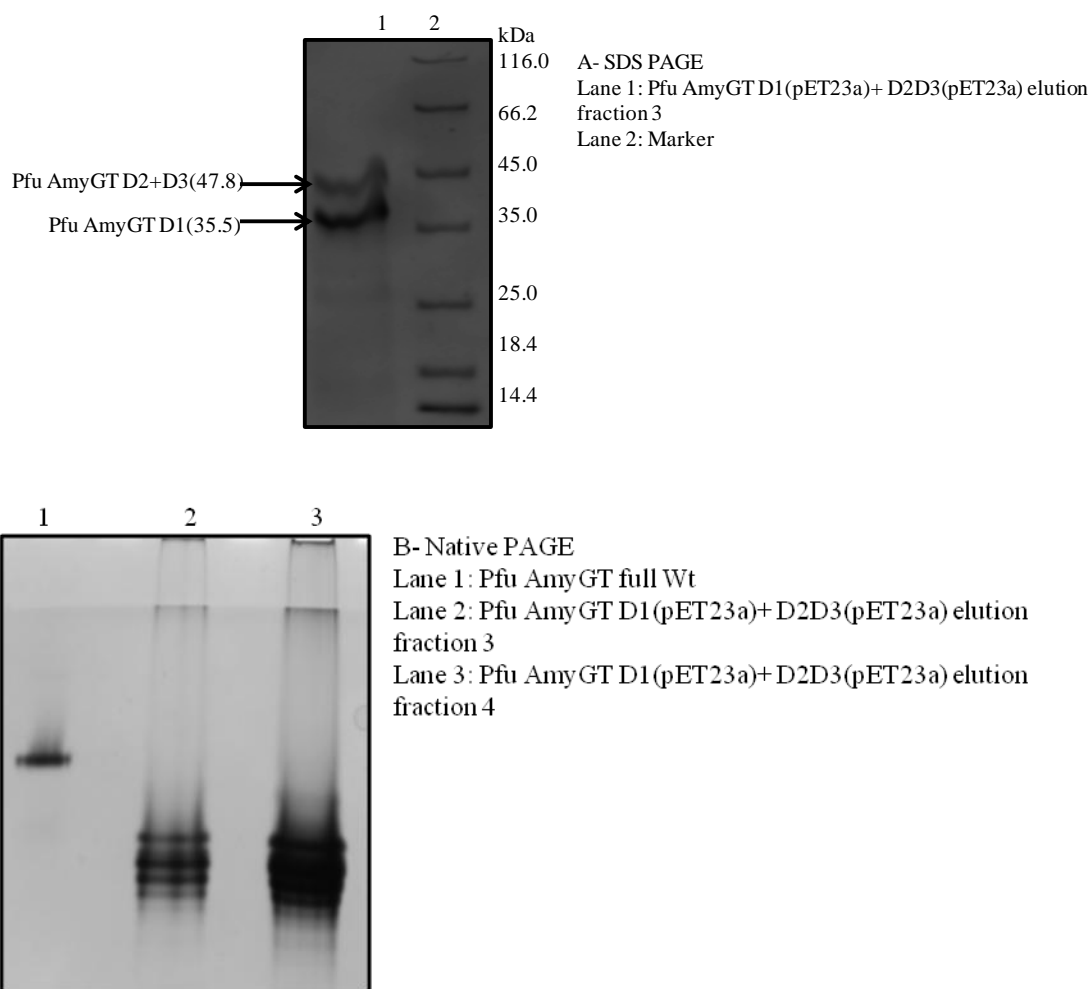


Figure 76: Panel A shows the expression of both domains on SDS PAGE. Panel B shows native PAGE gel of the co-refolded domains of PfuAmyGT.

3.15 Cloning, Expression and Purification of a glucanotransferase from *Thermococcus onnurineus* (TonAmyGT/TonGT):

Thermococcus onnurineus (strain NA1) is a hyperthermophilic archaeon (archaebacterium) ranging in size from 0.6-2.0 μm in diameter isolated from hydrothermal vents in the PACMANUS field of the East Manus Basin. This organism thrives at temperatures ranging from 70 to 100 $^{\circ}\text{C}$ and at pH of 5.6-8.5 (UniProt). The genome of this organism was recently sequenced by H S Lee *et al* in 2008. The genomic DNA of this organism was provided to us as a gift by Dr. Sanjeev Chandrayan (Institute of Chemical Technology, Mumbai). This was used

as a template to amplify the glucanotransferase gene (Figure 77) using the primers as mentioned in material and methods.

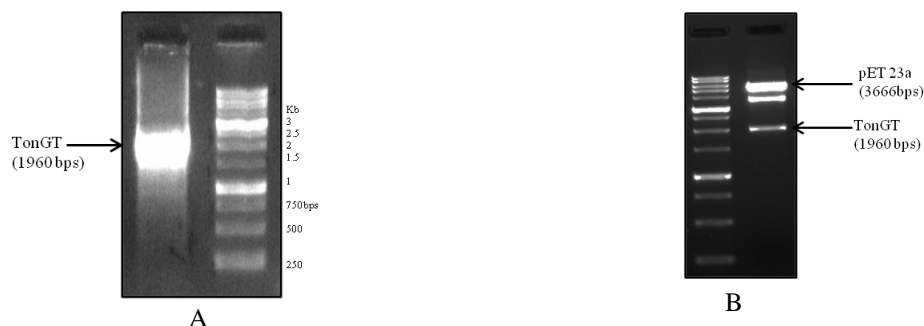


Figure 77: Panel A: It shows the amplified TonAmyGT gene from its genomic DNA. Panel B: It shows the digested plasmid obtained from a positive colony showing the insert of the desired size.

The obtained amplicon (Figure 77A) was then digested using enzymes *NheI* and *NotI* along with the vector pET-23a. The digested fragments were then ligated and transformed in XL1-Blue *E. coli* cells. After screening the obtained colonies and further confirmation by digestion (Figure 77B) and sequencing this plasmid containing the TonAmyGT as the insert was transformed into BL21 Star(DE3)pLysS for expression. These cells were then grown on a large scale (secondary culture) and induced with 1 mM IPTG at O.D₆₀₀ 0.6 for 6-8 hours at 37 °C, 220 rpm. The culture was then pelleted at 8,000 rpm and the cells obtained were re-suspended in native lysis buffer supplemented with 0.2 % sarkosyl and allowed to incubate overnight. For purification, these cells were then sonicated briefly and subjected to heating at 90 °C for 20-30 minutes and then centrifuged at 16,000 rpm and the supernatant was subjected to Ni-NTA affinity chromatography. The fractions collected during purification were analyzed on SDS PAGE (Figure 78), revealing the expression of the protein, TonAmyGT.

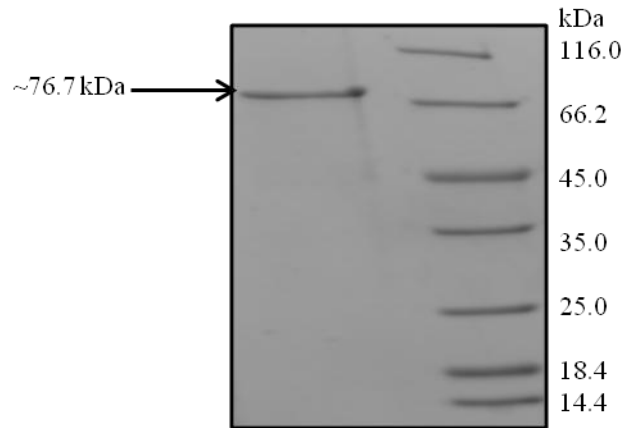


Figure 78: SDS-PAGE shows a reasonably purified, desired size band of Ni-NTA purified TonAmyGT.

3.16 Confirmation of identity by mass spectrometry:

The band observed on SDS PAGE was excised out of the gel and used for preparing samples for mass spectrometry. After in-gel tryptic digestion, the samples were subjected to MALDI TOF analysis and compared with fragments obtained from *in-silico* digestion of the protein sequence (Figure 79).

A

Mass	Position	Peptide Sequence
3903.8996	563-598	VLFGAELNLAVHSVMEEPDE FEATAIEVNDPYGIGK
3222.5366	1-27	MVNFIFGIHNHQP LGNFGWV FEDAYNR
3214.5942	616-644	TL SQSEAGWDFIQQGVSYTV LFPVDGELK
3173.4446	438-466	WEHYHEVPEAATPEEGDGEG VASIHEIGK
3155.5611	397-424	EEVFIESENF FAVFKPAYGGALFELSSK
2905.5457	44-68	VAVHISGPLLEWLDENKPDY IDLLR
2865.3617	505-529	YEELGDFLTGAYDYSPLGG IILWR
2846.3446	155-178	EELFWPYYTEDGGEVI AVFP IDEK
2417.1607	348-368	AQCNDAYWHGFGGVYLP HL R
2401.2148	261-281	GLVYLP IAS YFEMSEWSLPAK
2330.1009	134-154	EAGIDYVIVDDYHFMSAGLS K
2114.1895	74-93	GQLEIVVAGFYEPVLA AIPK
2090.0189	484-501	AILQDHF LSPETGLDDYR
2046.9663	28-43	SYPPFMEILEEYPNMK
1721.7864	191-205	TLEYLHSLDDGDESK
1617.7736	216-228	FGVWP GTYDWWYK
1542.7985	245-256	INLMLYSEYLQR
1485.7471	546-558	LTDDGFIVDYTVK
1347.6691	378-389	ANSYVFTGSFVR
1261.7303	181-190	YLPF RPVEK
1168.5092	318-326	YPESNYMHK
1134.5902	427-436	AVNYNDVLAR
1116.5320	206-215	VAVFHDDG EK
1087.5167	475-483	ELAYDSHPR
1016.5775	370-377	TIWENIHK
1010.5557	285-292	LFVEFVEK
1005.5583	328-336	MLMVSGLVR
998.5669	123-130	VWQPELVK
973.5061	530-539	DGSVSGKPAR

874.4053	295-301	EHGQFEK
860.4624	116-122	GVWLTER
837.3737	390-396	DIDFDGR
756.4250	468-473	IPFEER
713.3253	234-238	EFFDR
702.3609	313-317	NFFFK
701.3213	337-342	DNPEAR
692.3209	239-244	VSSDER
666.3457	110-115	LGYDAK
657.4082	343-347	HFLK
630.3821	97-101	IEQK
560.3191	308-312	GGWK
547.3351	257-260	FRPK
533.3293	69-73	SLVSK
520.3242	304-307	VFVR
504.3180	612-615	FPIK
503.2976	230-233	GWLK

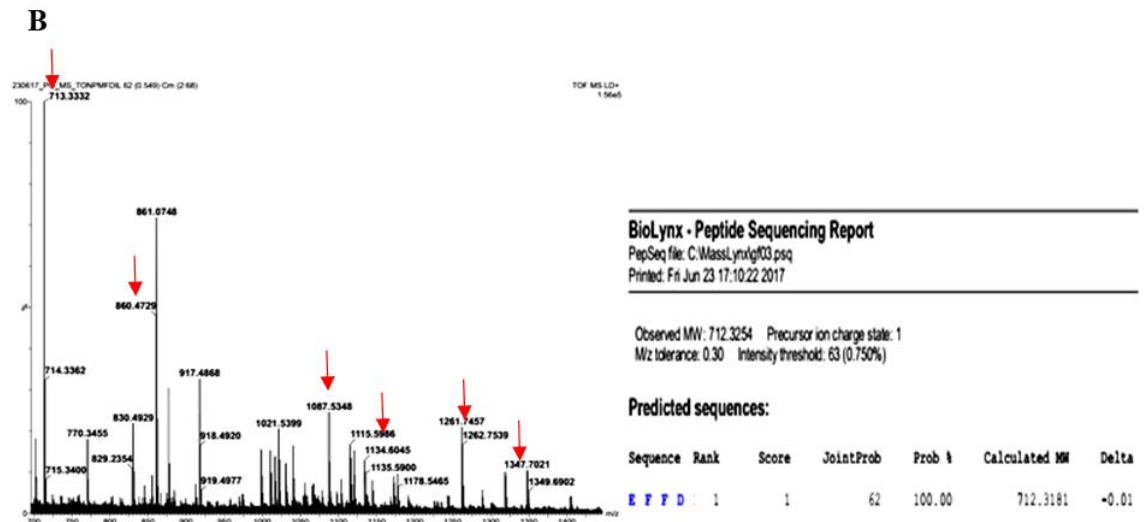


Figure 79: Panel A shows the *in silico* tryptic peptides obtained from TonAmyGT sequence and the masses we observed (marked in red in the table and shown by red arrows in the spectrum) which were further confirmed by *de novo* sequencing. Panel B shows the sequences/ peptide masses that were experimentally observed in the mass spectrum.

3.17 Oligomeric status of TonAmyGT:

3.17.1 Gel filtration chromatography:

The elution fractions obtained from Ni-NTA purification were subjected to chromatography on a Superdex-200 Increase column for determining the oligomeric status of the protein by size exclusion chromatography using AKTApurifier (GE Healthcare). The column was pre-equilibrated with 50 mM Tris, pH 8.0. We obtained a small fraction of the population at ~8 ml (Figure 80) which is the void volume of the column hence it can be interpreted as aggregated protein population, whereas the major fraction was obtained at ~11.84 ml which again corresponds to a dimeric population of the protein, as well as a fraction at ~14 ml which shows the same size as the 11.84 ml fraction on SDS-PAGE.

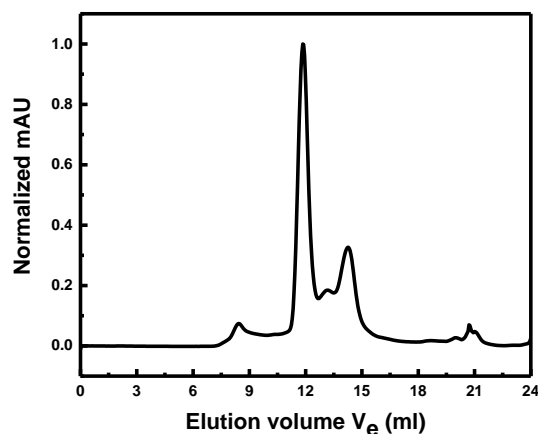


Figure 80: Gel filtration profile of TonAmyGT with a major population at ~11.8 ml.

3.17.2 Dynamic Light Scattering (DLS):

To further confirm the oligomeric status/ size of TonAmyGT, we performed Dynamic Light Scattering experiments and observe a hydrodynamic radius of ~5.01 nm (Figure 81), similar to PfuAmyGT.

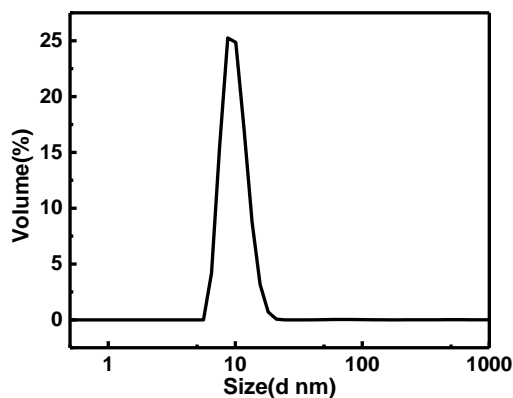


Figure 81: DLS profile of TonAmyGT.

3.18 Secondary and tertiary structure of TonAmyGT:

The major elution fraction from gel filtration chromatography was subjected to CD spectroscopy to determine the secondary structure of TonAmyGT. The CD spectrum shows a signature of mixed α helices and β sheets which correspond to the secondary structure obtained from I-TASSER (Figure 82A). To determine the overall tertiary structure of TonAmyGT its intrinsic fluorescence was exploited.

We observed peak fluorescence at ~340 nm which corresponds to a reasonably well-folded protein (Figure 82B).

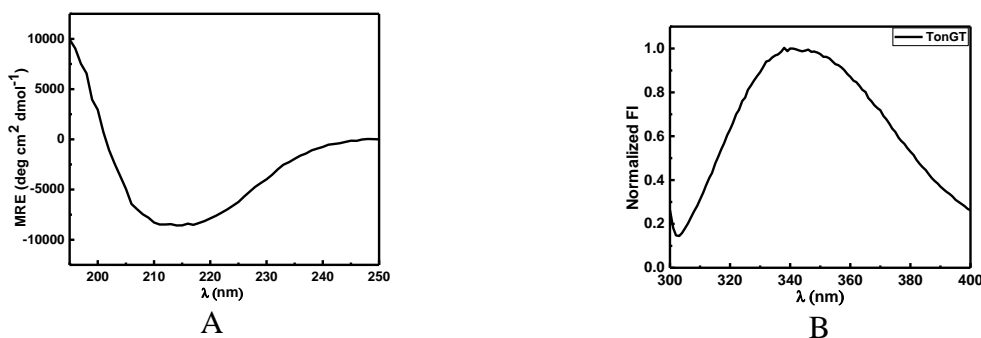


Figure 82: Panel A: Shows the CD spectrum of TonAmyGT and B shows the fluorescence emission spectrum of TonAmyGT.

3.19 Thermal and chemical stability of TonAmyGT:

The thermal stability of TonAmyGT was determined by CD spectroscopy and Differential Scanning Calorimetry (DSC). For CD, a suitable concentration (0.15mg/ml) of the protein was taken in 0.1 cm sealed CD cuvette and heated from 20 °C to 90 °C at a rate of 1 °C/minute.

We observed no substantial unfolding in the CD spectra owing to the hyperthermophilic nature of the protein (Figure 83).

We also examined the thermal stability of this protein using DSC at two different concentrations of 0.25 mg/ml and 0.5 mg/ml. The DSC data could be fitted in two ways; one showing two T_m of 90.5 °C and 103.2 °C, respectively while the other fitting showed three T_m values of 87.5 °C, 100.1 °C and 104 °C respectively. The high values of melting temperatures are in accordance with the hyperthermophilic nature of the protein.

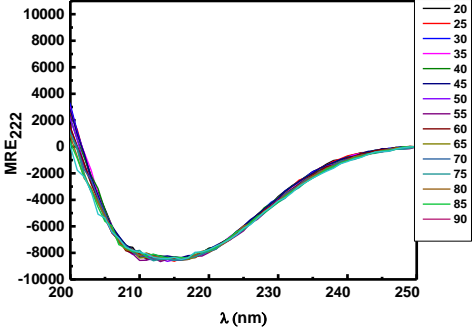
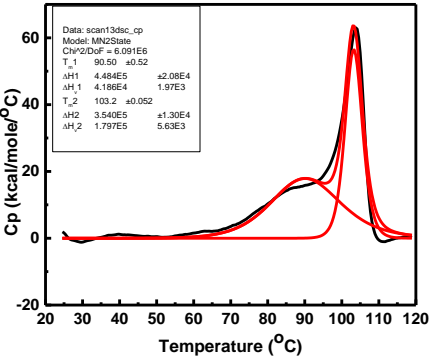
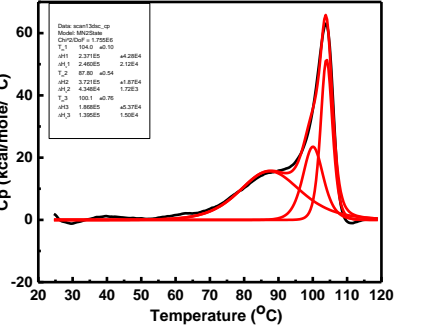
Spectrum	Thermal stability of TonAmyGT
	<p>Assessed using CD spectroscopy</p>
	<p>Assessed using DSC</p> <p>T_{m1}: 90.5 °C</p> <p>T_{m2}: 103.2 °C</p>
	<p>T_{m1}: 87.5 °C</p> <p>T_{m2}: 100.1 °C</p> <p>T_{m3}: 104 °C</p>

Figure 83: Assessment of the thermostable nature of TonAmyGT.

Using CD spectroscopy, the stability of TonAmyGT towards common chemical denaturants; urea and Gdm.HCl was also examined in a similar experimentation set up as used for full-length, wild-type PfuAmyGT. The results are summarized in Figure 84.

We observed that urea was unable to completely unfold TonAmyGT (as the MRE did not reach a value of zero) so the apparent C_m calculated is ~5.5M. On the

other hand, Gdm.HCl was able to completely denature TonAmyGT with a C_m of $\sim 3M$.

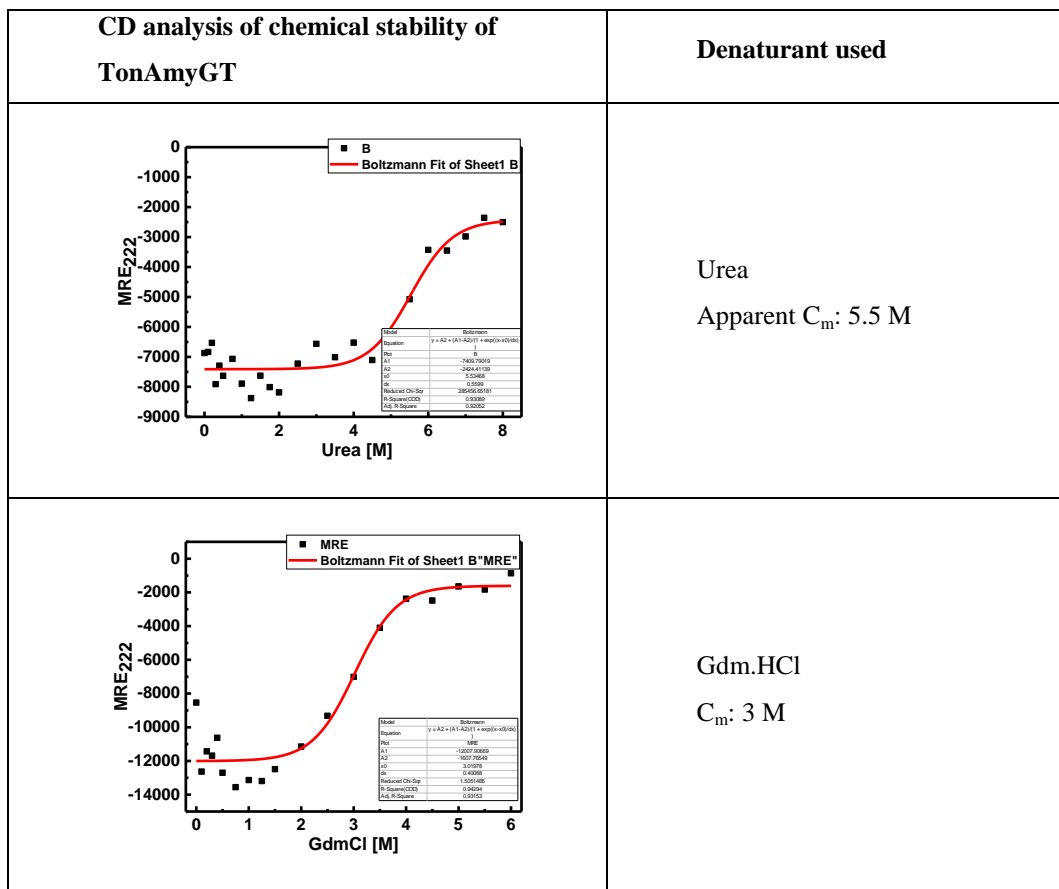


Figure 84: Assessment of chemical denaturation of TonAmyGT and calculation of C_m s for urea and Gdm.HCl for TonAmyGT.

3.20 Activity of TonAmyGT:

3.20.1 Determining the smallest donor and acceptor molecule:

The activity of TonAmyGT was examined in a manner similar to that of PfuAmyGT using TLC. Similar to studies with PfuAmyGT, a combination of donor and acceptor molecules was used to observe activity as well as determine the smallest length necessary for each case. Each reaction mixture contained a specified amount of donor and/or acceptor molecule (20 mM each) with 2 μM TonAmyGT which was incubated at 90 °C for 12 hours (Figure 85).

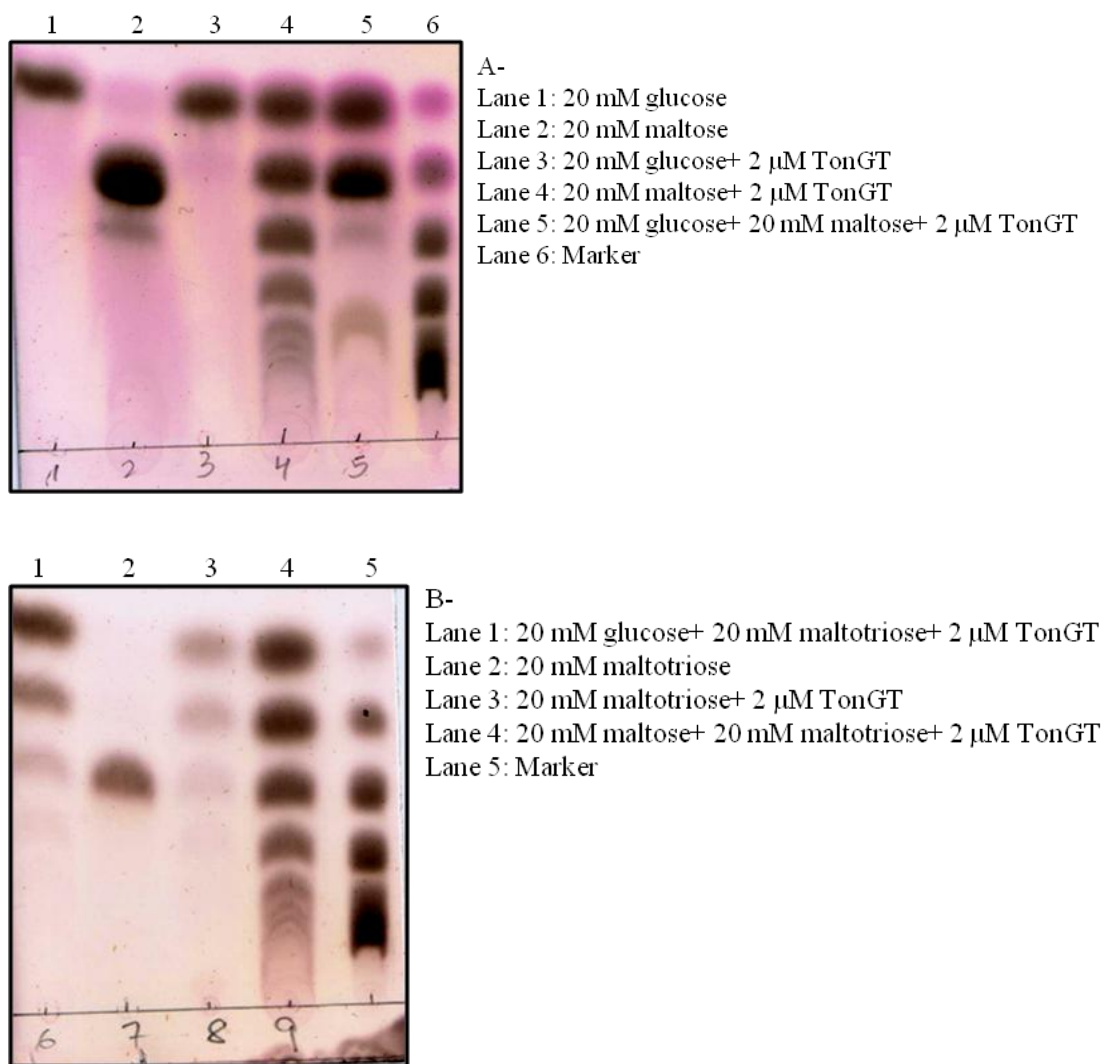


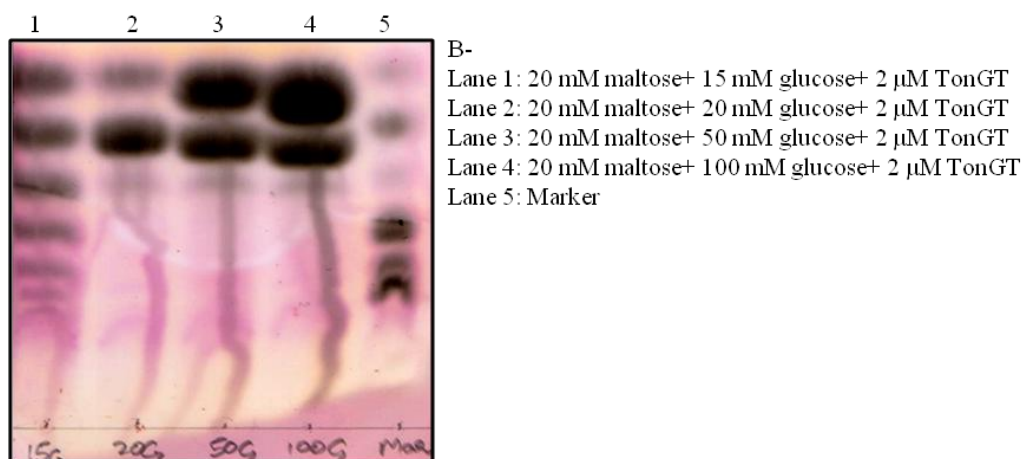
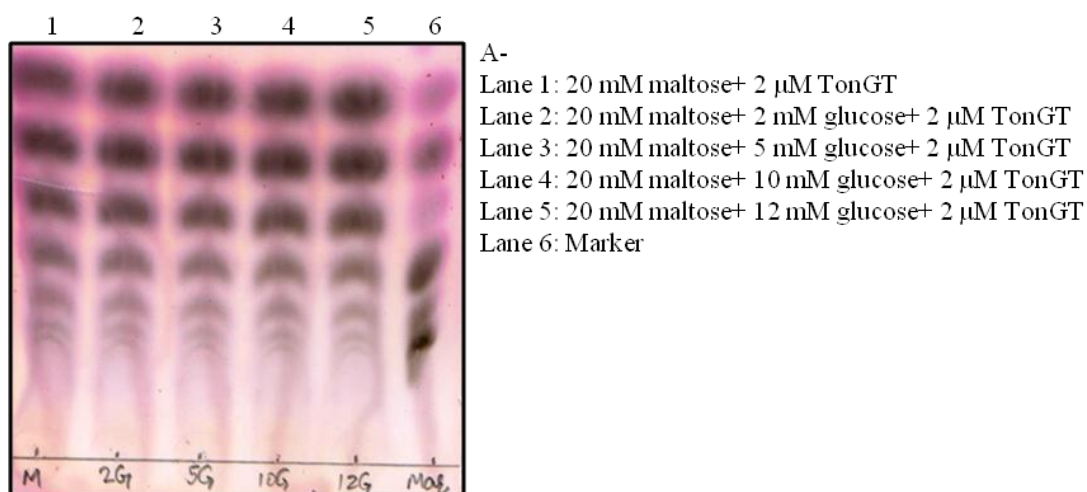
Figure 85: The above TLCs show the activity of TonAmyGT on various oligosaccharides. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

We observed the formation of different oligosaccharides when maltose was present alone with the enzyme as well as when it was present along with equimolar concentrations of maltotriose. This suggests that the smallest donor, as well as acceptor, is maltose, as the enzyme is capable of working with maltose alone, whereas PfuAmyGT was not capable of working with maltose alone but capable of working with maltotriose alone.

We also observed that when glucose was present in the reaction mixture, very few oligomeric species were observed. The presence of glucose seems to inhibit the activity of TonAmyGT when present with a fixed concentration of maltose, as seen in Figure 86. Till a glucose concentration of 15 mM glucose along with 20

Results and Discussion

mM maltose, significant activity is observed. After this, TonAmyGT activity starts to decrease with increasing glucose concentration and is completely abolished in the presence of 20 mM glucose. This observation suggests that there is a form of feedback inhibition caused by the increasing concentration of glucose. The simplest explanation is that glucose competitively inhibits the binding of maltose to the donor and acceptor sites, and activity is not seen since the smallest donor is maltose.



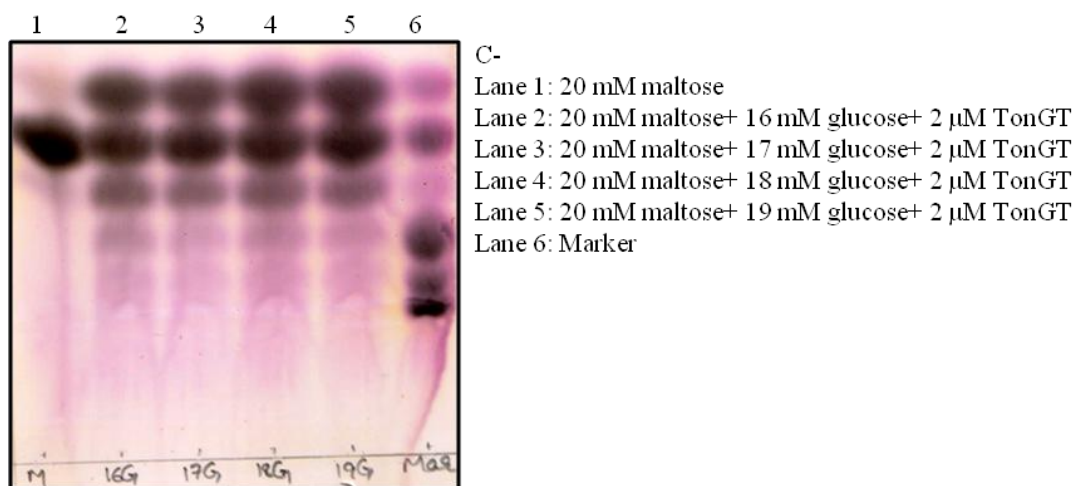
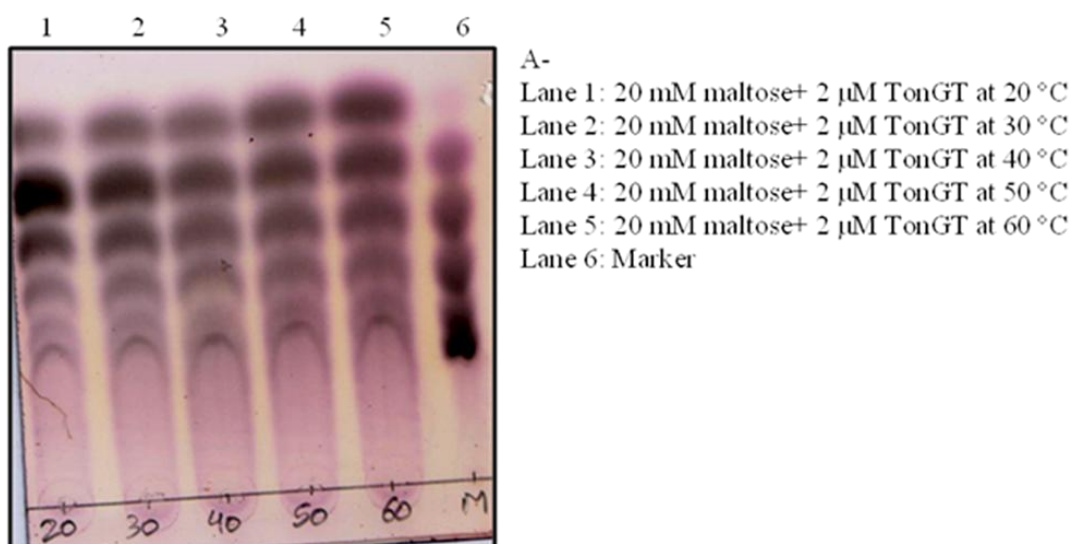


Figure 86: The effect of increasing glucose concentration on TonAmyGT activity when maltose is used as the substrate at 90 °C for 12 hrs. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.20.2 Temperature-dependent activity of TonAmyGT:

Using maltose as the only substrate for TonAmyGT, its temperature-dependence was examined and the enzyme was found to be functional from 20 to 100 °C (Figure 87). Increase in activity is observed after 70 °C.



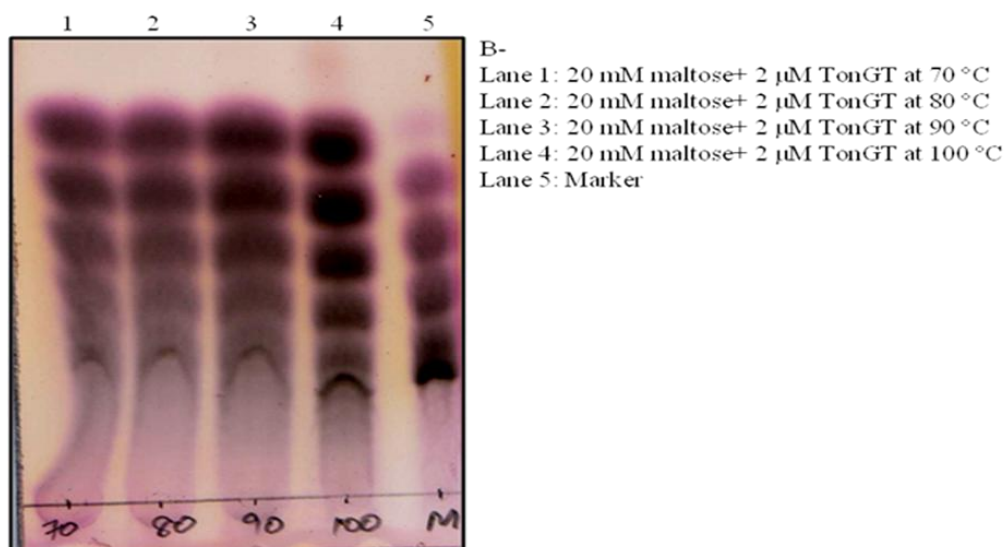


Figure 87: TLCs showing the temperature-dependence profile of TonAmyGT activity on maltose. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.20.3 Time-dependent activity of TonAmyGT:

Similar to the above set of reactions, the time-dependent activity of TonAmyGT on 20 mM maltose was observed at 90 °C. The enzyme becomes active after 2 hours of incubation and remains significantly active even after incubation for 12 hours at 90 °C (Figure 88).

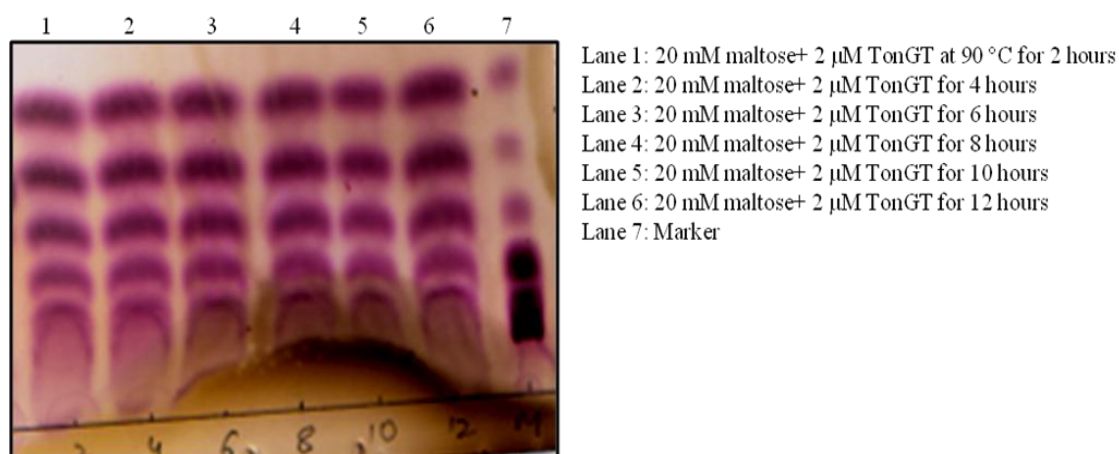


Figure 88: This TLC shows the time-dependent reaction of TonAmyGT. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

3.20.4 pH-dependent activity of TonAmyGT:

A reaction mixture containing 20 mM maltose with TonAmyGT was incubated at different pH at 90 °C for 12 hours to obtain a working pH range for the enzyme. We observed that this enzyme is reasonably active between the pH range from 3.0 to 9.0 (Figure 89).

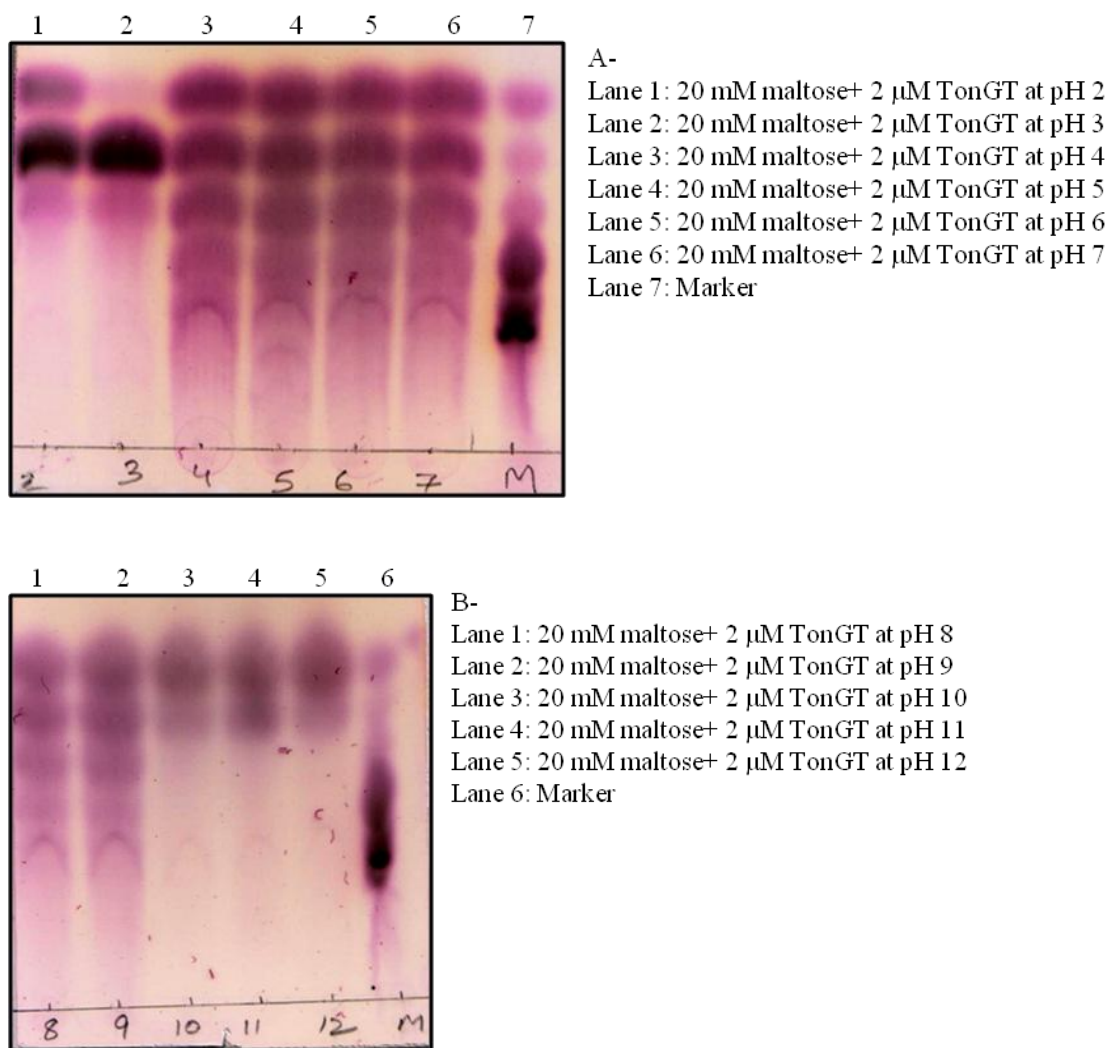


Figure 89: The above TLCs show the working pH range of TonAmyGT to lie between pH 3.0 to 12.0. The markers from top to bottom are glucose, maltose, maltotriose, maltotetraose, maltopentaose, and maltohexaose.

References:

- Chris P. Ponting and Robert R. Russell. The Natural History of Protein Domains. *Annu. Rev. Biophys. Biomol. Struct.* 2002, 31, 45-71.
- Fuwa H. A new method for microdetermination of amylase activity by the use of amylose as the substrate. *Journal of Biochemistry.* 1954, 41(5), 583-603.
- Gideon J Davies and Spencer J Williams. Carbohydrate-active enzymes: sequences, shapes, contortions and cells. *Biochem. Soc. Trans.* 2016, 44, 79-87.
- Gregory L Cote and Bernard Y Tao. Oligosaccharide Synthesis by Enzymatic Transglycosylation. *Glycoconjugate.* 1990, 7, 145-162.
- Hiromi Imamura H, Shinya Fushinobu S, Masaki Yamamoto M, Kumasaka T, Jeon B S, Wakagi T and Matsuzawa H. Crystal Structures of 4- α -Glucanotransferases from *Thermococcus litoralis* and its Complex with an Inhibitor. *The Journal Of Biological Chemistry.* 2003, 278 (21), 19378-86.
- Imamura H, Fushinobu S, Yamamoto M, Kumasaka T, Jeon B S, Wakagi T, and Matsuzawa H. Crystal Structures of 4- α -Glucanotransferase from *Thermococcus litoralis* and Its Complex with an Inhibitor. *Journal of Biological Chemistry.* 2003, 278(21), 19378-19386.
- Kim M J, Lee S B, Lee H B, Lee S Y, Baek J S, Kim D, Moon T W, Robyt J F and Park K H. Comparative Study of the Inhibition of α -Glucosidase, α -Amylase, and Cyclomaltodextrin Glucanosyltransferase by Acarbose, Isoacarbose, and Acarviosine–Glucose. *Archives of Biochemistry and Biophysics.* 1999. 371, 2, 277–283.
- Kim T J, Kim M J, Kim B C, Kim J C, Cheong T K, Kim J W and Park K H. Modes of Action of Acarbose Hydrolysis and Transglycosylation catalyzed by a thermostable maltogenic amylase, the gene for which was cloned from *Thermus* strain. *Applied and Environmental Microbiology.* 1999, 65(4), 1644-51.
- Kuharski RA, Rossky PJ. Solvation of hydrophobic species in aqueous urea solution: A molecular dynamics study. *J Am Chem Soc.* 1984, 106, 5794–5800.
- Laderman K A, Asada K, Uemori T, Mukai H, Taguchi Y, Kato I and Anfinsen C B. α -Amylase from the Hyperthermophilic Archaeobacterium *Pyrococcus*

furiosus Cloning and Sequencing of the gene and Expression in *Escherichia coli*. *J. Biol. Chem.* 1993, 268(32), 24402-24407.

- Lairson L L, Henrissat B, Davies G J, Withers S G. Glycosyltransferases: Structures, Functions, and Mechanisms. *Annu. Rev. Biochem.* 2008, 77, 521–55.
- Lee H S, Kang S G, Bae S S, Lim J K, Cho Y, Kim Y J, Jeon J H, Cha S S, Kwon K K, Kim H T, Park C J, Lee H W, Kim S, Chun J, Colwell R R, Kim S J and Lee H J. The Complete Genome Sequence of *Thermococcus onnurineus* NA1 Reveals a Mixed Heterotrophic and Carboxydrotrophic Metabolism. *Journal of Bacteriology.* 2008, 190(22), 7491-7499.
- Lee H S, Shockley K R, Schut G J, Connors S B, and Montero C I. Transcriptional and Biochemical Analysis of Starch Metabolism in the Hyperthermophilic Archaeon *Pyrococcus furiosus*. *Journal of Bacteriology.* 2006, 188(6), 2115-2125.
- Monera O D, Kay C M and Hodges R S. Protein denaturation with guanidine hydrochloride or urea provides a different estimate of stability depending on the contributions of electrostatic interactions. *Protein Science.* 1994, 3, 1984-1991.
- Palmer T N, Ryman B E and Whelan W J. The Action Pattern of Amylomaltase from *Escherichia coli*. *Eur. J. Biochem.* 1976, 69, 105-115.
- Qasba PK, Ramakrishnan B, Boeggeman E. Substrate-induced conformational changes in glycosyltransferases. *TRENDS in biochemical sciences.* 2005, 30(1), 53-62.
- Raymond F. Greene JR. and C. Nick Pace. Urea and Guanidine Hydrochloride Denaturation of Ribonuclease, Lysozyme, Zhymotrypsin, and Lactoglobulin. *J. Biol. Chem.* 1974, 249(17), 5388-5393.
- Soper AK, Castner EW, Luzar A. Impact of urea on water structure: A clue to its properties as a denaturant. *Biophys Chem.* 2003, 105, 649–666.
- Tao H, Liu W, Simmons B N, Harris H K, Cox T C, and Massiah M A. Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. *BioTechniques.* 2010, 48(1).
- Tracey M Gloster and David J Vocadlo. Developing inhibitors of glycan processing enzymes as tools for enabling glycobiology. *Nature Chemical Biology.* 2012, 8, 683-694.

4. Conclusions

During the course of evolution, many enzymes take up new functions while retaining a residual amount of ancestral functions as well. A lot of such examples are found in amylolytic and other related enzymes mainly belonging to the $(\beta/\alpha)_8$ -barrel protein family.

Hydrolytic enzymes often exhibit the property of bond formation either on the change of reaction conditions or intrinsically to a small extent. These molecules can transfer the cleaved moiety from a substrate molecule to water, another molecule and/or form cyclic products.

Literature is full of reports at attempts trying to convert naturally occurring enzymes from various sources to one of the three above mentioned categories, depending upon the importance and utilization of products in different industrial processes. This course has led to the discovery and characterization of many new enzymes as well as the determination of the structure of many existing ones. Different reports talk about the presence of one main active site in the enzyme and another secondary site or regions around the active site where a putative substrate molecule has been found to be bound, from X-ray crystallography studies. Mutations in these auxiliary regions have also helped to increase the transferase activity in many enzymes eliminating the hydrolytic activity of the enzyme to a great extent.

In the present thesis, we have characterized one such enzyme PfuAmyGT from *Pyrococcus furiosus* in terms of its amylase as well as glucanotransferase activity. We have successfully demonstrated this enzyme to have an exo-amylase function to a small extent while its major activity is in the form of a glucanotransferase. This enzyme can use various substrate molecules as donors and acceptors. The smallest donor molecule required by PfuAmyGT is maltotriose and the smallest acceptor is glucose making this enzyme unique in the class of glucanotransferases. In terms of polysaccharide substrates, it can convert starch, cellulose, pectin, and xylan into varying lengths of the oligosaccharides. An added advantage of this enzyme is its thermal stability and stability towards common denaturants making it a suitable candidate for many industrial applications.

Conclusions

Combining structural analysis, comparison to known homologs and mutational analysis we deciphered the mechanism of action of PfuAmyGT. It uses a tryptophan-containing and aspartate-containing loop, to transfer the excised glucose molecule from the donor to the acceptor molecule.

We have also cloned and characterized another 4- α -glucanotransferase from *Thermococcus onnurineus* (TonAmyGT/TonGT). The genome of this organism has been recently sequenced. This enzyme is similar to PfuAmyGT in terms of its thermal and chemical stability but differs from it, in respect of its smallest donor, as well as acceptor molecule, which happen to be maltose, although the transferred glucan is still glucose as in the case of PfuAmyGT. Also, unlike PfuAmyGT, TonAmyGT appears not to be able to handle diverse substrates and works only with starch and maltooligosaccharides. The reasons for the differences in the actions of this enzyme remain to be deciphered, which may open new doors for future investigations.