Isolation, Characterisation and Application of Novel Microbial Protein Cross-Linking Enzymes

Richard Leonard Parsons

Doctor of Philosophy

Aston University

December 2006

Acknowledgements

I would like to thank Prof. M. Griffin for providing me with the opportunity to carry out this study. I would also like to thank the EPSRC and BLC Leather Technology Centre Ltd. for providing financial support and the European Project HIPERMAX for additional funding.

I am indebted to Dr. Russell Collighan for his advice and for rescuing my laptop, which decided to crash at a very inopportune moment. There are many other people that I would like to thank for the unforgettable moments over the past few years and perhaps mostly for keeping me sane, well almost. There are far too many to list but you know who you are.

I would finally like to thank my parents, Gillian and Peter, for their constant support throughout the many highs and lows and Lisa Holland for putting up with me, because I know how moody I can get !

.

Isolation, Characterisation and Application of Novel Microbial Protein Cross-Linking Enzymes

Richard Leonard Parsons

Doctor of Philosophy

December 2006

Microbial transglutaminase is favoured for use in industry over the mammalian isoform, and hence has been utilized, to great effect, as an applied biocatalyst in many industrial areas including the food and textiles industries. There are currently only a limited number of microbial TGase sources known.

A number of organisms have been screened for transglutaminase activity using biochemical assays directed towards TGase catalyzed reactions (amine incorporation and peptide cross-linking assay). Of those organisms screened, TGase was identified in a number of isolates including members of the *Bacillus* and *Streptomyces* families. In addition, a protein capable of performing a TGase-like reaction was identified in the organism *Pseudomonas putida* that was deemed immunologically distinct from previously described TGase isoforms, though further work would be required to purify the protein responsible.

The genuses *Streptoverticillium* and *Streptomyces* are known to be closely related. A number of micro-organisms relating to *Streptomyces mobaraensis* (formerly *Streptoverticillium mobaraensis*) have been identified as harboring a TGase enzyme. The exact biological role of *Streptomyces* TGase is not well understood, though from work undertaken here it would appear to be involved in cell wall growth.

Comparison of the purified *Streptomyces* TGase proteins showed them to exhibit marginally different characteristics in relation to enzymatic activity and pH dependency upon comparison with *Streptomyces mobaraensis* TGase. In addition, TGase was identified in the organism *Saccharomonospora viridis* that was found to be genetically identical to that from *S. mobaraensis* raising questions about the enzymes dissemination in nature. TGase from *S. baldaccii* was found to be most diverse with respect to enzymatic characteristics whilst still retaining comparable ε(γ-glutamyl) lysine bond formation to *S. mobaraensis* TGase. As such *S. baldaccii* TGase was cloned into an expression vector enabling mass production of the enzyme thereby providing a viable alternative to *S. mobaraensis* TGase for many industrial processes.

Keywords: bacterial, transglutaminase, *Streptomyces*

Publications

Articles

R. L. Parsons. Microbial Transglutaminase: Current Knowledge and Commercial Applications. 2003, *BLC Journal* p 12-16.

Abstracts

R.L.Parsons, R. Collighan and M. Griffin. (2003) Selection, Characterisation and Application of Novel Microbial Protein Cross-Linking Enzymes. 1st Pro-Bio Faraday Annual Conference. 9-11 November 2003, Edinburgh, UK.

R.L.Parsons, R. Collighan and M. Griffin. (2004) Selection, Characterisation and Application of Novel Microbial Protein Cross-Linking Enzymes. Nottingham Trent University Poster Day, Nottingham. 15th July 2004, Awarded 2nd Prize

R. Parsons, R. Collighan, M. Griffin & G. Szakacs. (2005) Biochemical analyses of transglutaminase from strains of *Streptomyces*. 8th International Conference on Protein Crosslinking and Transglutaminases, Luebeck, Germany. 1-4 September 2005.

R. Parsons, R. Collighan, M. Griffin & G. Szakacs. (2005) Biochemical analyses of transglutaminase from strains of *Streptomyces* Pro-Bio Annual Conference - Biocatalysis: Enzymes, Mechanisms and Bioprocesses, Manchester, UK. 21-22 November.

Abbreviations

°C	degrees Centigrade
λ	wavelength
aa	amino acid
ATCC	American Type Culture Collection, USA
ATP	adenosine-5'-triphosphate
Au	Absorbance units
B4.2	band 4.2 protein
BSA	bovine serum albumin
Biotin-X-cadaverine	5((N-biotinoyl) amino) hexanoyl) amino) pentylamine tifluoroacetate salt
bp	base pairs
BTLCP	bacterial transglutaminase-like cysteine protease
CBZ-Gln-Gly	N-Carbobenzoxy-Glutaminylglycine
Da	Daltons
dH_2O	distilled water
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECL	enhanced chemiluminescence
EDC	N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride
EDTA	ethylene diamine tetra acetic acid
ELISA	enzyme linked immunosorbant assay
EXAP	extravidin peroxidase conjugate
FXIII	factor XIII
GC	gas chromatography
gpl	guinea pig liver
GST	glutathione S transferase
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
IgG	immunoglobulin
IPTG	isopropyl-β-D-thiogalactopyranoside
kDa	kilodaltons
kb	kilo bases
L	litre
LB	Luria bertani
M	molar
MAb	monoclonal antibody
M_W	molecular weight
Min	minutes
ml	millilitres
mM	millimolar
mm^2	millimetres squared
mTGase	microbial transglutaminase
NEM	N-ethylmaleimide
µl	microlitre

μM	micromolar
nM	nanomolar
nm	nanometres
NRRL	Northern Regional Research Centre, Illinois, USA
NTUCC	Nottingham Trent University Culture Collection
OPA	O-phthaldialdehyde
PAb	polyclonal antibody
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
pH	negative log of hydrogen ion concentration
PMSF	Phenyl methyl sulfonyl fluoride
R281	N-Benzyloxycarbonyl-L-phenylalanyl-6-dimethylsulphonium-5-oxo-L-norleucine
R283	1,3-dimethyl-2-[(2-oxopropyl)thio]imidazolium chloride
Rpm	revolutions per minute
SD	standard deviation
SDS	sodium dodecyl sulphate
TAE	tris-Acetate-EDTA
TBS	tris buffered saline
TCA	trichloroacetic acid
TE	tris-EDTA
TEMED	N, N, N, N'-tetramethylene diamine
TFA	triflouroacetic Acid
TGase	transglutaminase
TMB	3, 3', 5, 5'- tetramethylbenzidine
TNF	tumour necrosis factor
TNTU	Nottingham Trent University
TUB	Technical University of Budapest, Hungary
U	units
UV	ultra violet
IFO	Institute for Fermentation, Japan
NCIB	National Collection of Industrial Bacteria, Scotland

Table of Contents

List of Figures

Chapter 1; Introduction

Chapter 2; Materials and Methods

Chapter 3; Method Development

Chapter 4; Screening for Microbial TGase Enzymes

Chapter 5; Studies on TGases of Streptomyces and Related Genera

List of Tables

Chapter 1; Introduction

Chapter 2; Materials and Methods

Chapter 5; Studies on TGases of Streptomyces and Related Genera

Chapter 1

Introduction

1. Introduction

The term transglutaminase was first introduced by Clarke and co-workers to describe the transamidating activity of an enzyme found within guinea pig liver (Mycek 1959). Since their initial discovery, many types of transglutaminases "R-glutaminyl-peptide-amine-γ-glutamyl transferase" [EC 2.3.2.13] have been discovered, all of which mediate (either by a calcium dependent or independent manner) an acyl transfer reaction between the γ-carboxamide group of a peptide-bound glutamine residue and the ε-amino group of a peptide-bound lysine (Figure 1.1a). The resulting covalent bonds, first observed in the stabilisation of fibrin monomers (Pisano 1968), are deemed resistant to most proteases and chemical degradation. The specificity of TGase for peptide-bound glutamine residues distinguishes them from similar enzymes involved in glutamine metabolism.

1.1. TGase Enzymology

TGase catalysed reactions, for most enzyme isoforms, are dependent on the presence of Ca^{2+} for enzymatic activity, though TGases found in plants and most notably in microbes (*Streptoverticillium sp.* and *Bacillus sp.*) may post translationally modify proteins in a Ca^{2+} independent manner. The formation of the isopeptide linkage proceeds as follows; in the first step, a glutamine-containing acyl donor substrate binds to the cysteine residue in the active site of TGase (Folk 1969). Nucleophilic attack of the γ-carboxamide group of a glutamine residue, of the substrate protein, by the active site thiol of the enzyme leads to the formation of an acyl-enzyme intermediate. The second step involves the amine donor substrate binding to the acyl-enzyme intermediate whereby the acyl group is transferred to the acyl acceptor substrate which results in the formation of an isopeptide bond and synchronized release of the reactivated enzyme. The reaction is the reverse of the proteolysis reaction catalysed by the thiol proteases, enzymes that possess an identical catalytic triad and to which TGases are structurally related. The resulting cross-link may be intramolecular to modify the properties of single polypeptides/proteins, or more frequently, intermolecular

to form polymers. The TG-catalysed reaction may therefore add new properties to the protein substrates, resulting in an altered function.

In addition to peptide bound lysines, TGases may also incorporate unbranched aliphatic primary amines or primary amines that satisfy the enzymes optimal structural requirements (Figure 1.1b) (Lorand 1979). Low molecular weight amines, especially polyamines, taking part in the transamidating reaction resulting in the formation of an N- mono (glutamyl) polyamine. In the presence of a second reactive glutamine residue, the reaction may proceed to covalent cross-linking between two polypeptide chains via a N,N-bis (glutamyl)polyamine bridge (Figure 1.1c). Alternatively, in the absence of suitable amine donors and in relevant conditions (the presence of water and a slightly acidic pH), TGases may deamidate polypeptide glutamines to glutamate in a deamidation reaction (Folk 1977).

Transglutaminases may also catalyse certain other reactions, though this depends on the particular enzyme or the substrate availability (Lorand 1979). TGases have been shown to play a role in the aminolysis of esters (i.e. p-nitrophenyl-acetate) by the insertion of polyamine groups within the ester chain (Folk 1966; 1977). Additionally, it has been proposed that transglutaminase type 1 (involved in cell envelope formation during the terminal differentiation of keratinocytes) may catalyze an ester linkage between involucrin glutamyl residues and ω-hydroxyceramides, which may be important in epidermal lipid envelope formation (Nemes 1999).

The $\varepsilon(\gamma$-glutamyl)lysine isopeptide bond generated by TGase catalysed reactions were initially thought to be resistant to proteolytic degradation (Lorand 1984). However it is possible to hydrolyse the bonds formed as demonstrated by enzymes found in the glandular secretions of *Hirudo medicinalis* (Zavalova 1996). Additionally Factor XIIIa and TGase 2 have been shown to possess isopeptidase activity leading to the suggestion that these enzymes play a more dynamic role in the cell cycle than previously thought (Parameswaran 1997).

Figure 1.1; Reactions Catalysed by Transglutaminases.

Formation of the acyl-enzyme intermediate by nucleophilic attack of the active site thiol, A. When the acyl acceptor is lysine the formation of a ε(γ-glutamyl) lysine bond is catalysed, B. Other reactions catalysed include the incorporation of amines into the glutamine residue of the acceptor protein, C, which may lead to the formation of a N',N'-bis (γ-glutamyl) polyamine linkage if the amine is a diamine, D. Deamidation of a protein bound glutamine may also occur, E.
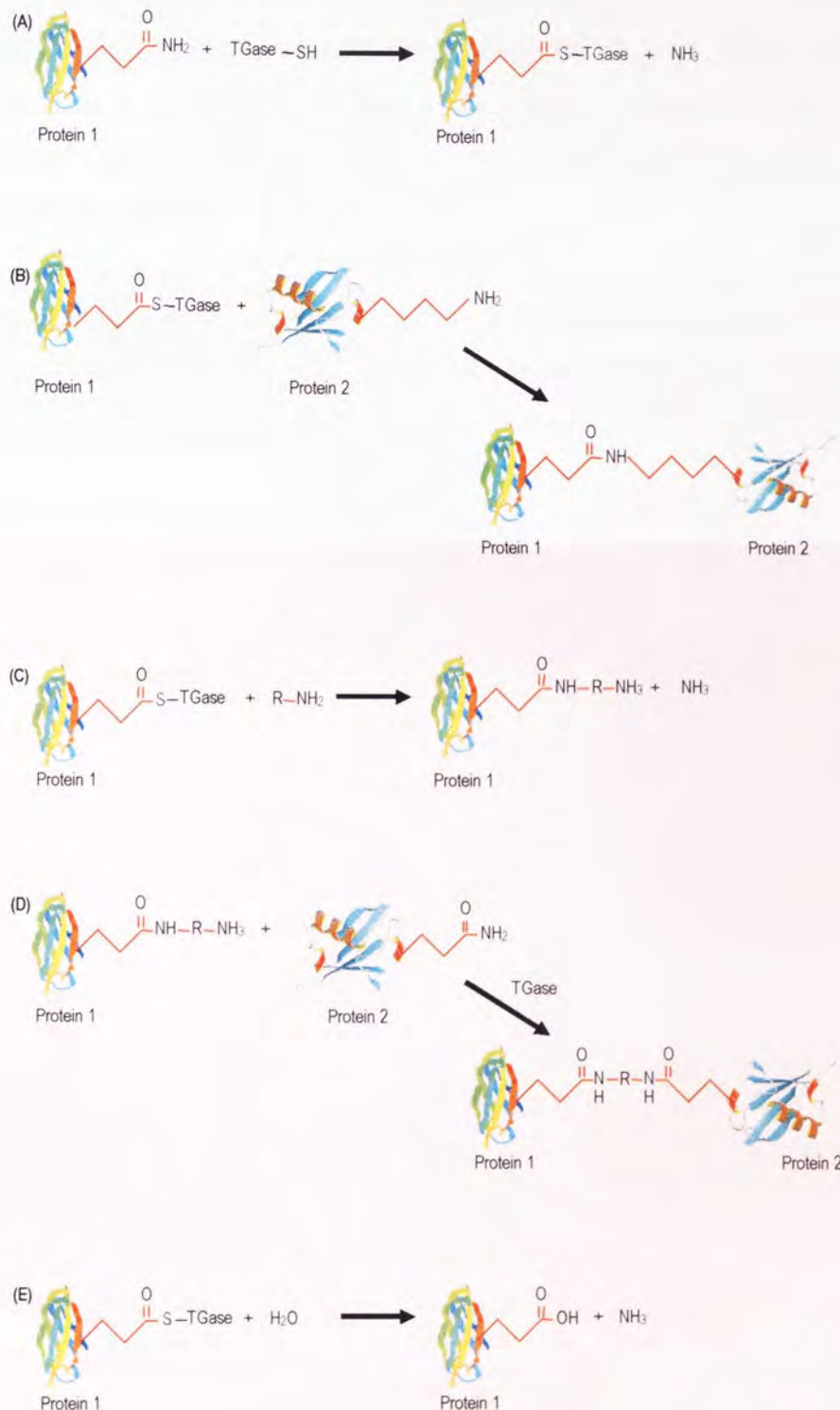
1.2. TGase Superfamily

TGase isoforms have been identified in a wide variety of organisms implying a functional necessity and a wide distribution. Enzyme activity has been described in a number of vertebrates and invertebrates (Bures 1978; Yasueda 1994; Nozawa 1997), (Mehta 1990; 1992) , plants (Serafini-Fracassini 1995) and micro-organisms (Kanaji 1993; Kobayashi et al. 1998; Kobayashi 1998b). With the exception of bacteria and plants, enzyme activity is dependant on the presence of extraneous Ca^{2+}. Figure 1.2 shows the distribution of TGases in a number of organisms and their degree of relatedness. It is apparent that TGases of bacterial origin are clear outliers from the main group, sharing little or no sequence homology to their eukaryotic counterparts.

Figure 1.2; A Phylogenetic Tree of the TGase Protein Family.

Closely related members are grouped. Sequence alignment constructed using ClustalW (Higgins 1994) and phylogenetic tree visualised using TreeView software (Page 1996). The bar represents 0.1 substitutions per site.
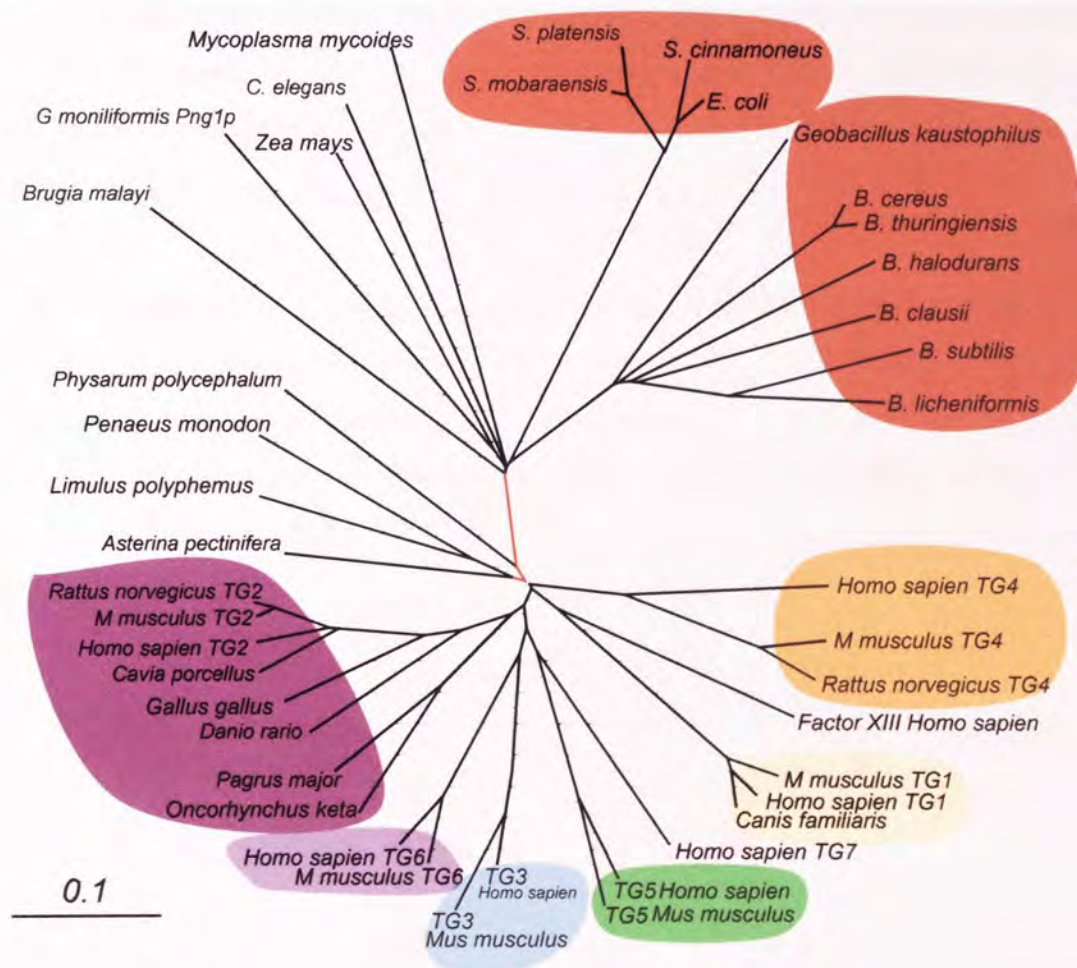
1.2.1. Enzymes Exhibiting TGase Activity

In certain instances, toxins produced by the bacteria *Escherichia coli* and *Bordetella pertussis,* cytotoxic necrotizing factor (CNF) and dermo necrotizing toxin (DNT) respectively, have been shown to possess *in vitro* TGase activity in the presence of primary amines (Lerm 1999) with the epsilon-(gamma-glutamyl)lysine bond having been described in the organism *E. coli* previously (Matacic 1979). Both CNF and DNT comprise a family of bacterial toxins that target eukaryotic Rho proteins and modify their function via deamination that occurs at a specific glutamine residue (Gln63) in mammalian RhoGTPase thereby inhibiting the GTPase activity of Rho (Sugai 1999). In cultured cells, both toxins induce actin polymerization and inhibit cytokinesis that results in the formation of multinucleated cells (Fiorentini 1995). The toxins CNF and DNT are;115 and 165 kDa proteins, respectively, which share a region of homology encompassing their enzymatic domain (Walker 1994). Both possess similar biological properties; they are homologous in a localised C-terminal region, showing 35% identity to *Streptomyces* TGase with the C-terminal regions comprising the active domain for deamination (Horiguchi 2001). Both CNF and DNT have been shown to possess the residues necessary for TGase activity, sequence analysis reveals the presence of the cysteine residue, deemed enzymatically essential (Figure 1.3).

This suggests that TGases are widely distributed throughout the prokaryotic kingdom, carrying out a myriad of functions. Although they are not as well conserved as their eukaryotic counterparts, they do possess a cysteine residue reminiscent of the active site region, considered enzymatically essential (Kanaji 1993).

Figure 1.3; Comparison of Active Site Regions of known TGases and the Bacterial Toxins CNF1, CNF2 and DNT.

Active site Cysteine underlined. Conserved eukaryotic active site sequence shaded.
Where DNT is dermonecrotising toxin: CNF is cytotoxic necrotising factor; S TG is *Streptomyces* derived TGase; TG1 is keratinocytes TGase and TG2 is tissue TGase, TG3 is epidermal TGase, TG4 is prostate TGase with TG5, 6 and 7 representing TGase x, y and z respectively.

```
DNT           G G S L S - G C T T M V G V K E G I L 1316
CNF1          S G N L S - G C T T I V A R K E G I I  877
CNF2          S G N L S - G C T T I V A R K G G I I  877

S TG          R E W L S Y G C V G V T W V N S G Q Y   75

Factor XIII   Q P V R Y G Q C W V F A G V F N T F L  325
TG1           Q S V P Y G Q C W V F A G V T T T V L  384
TG2           Q R V K Y G Q C W V F A A V A C T V L  287
TG3           Q P V K Y G Q C W V F A A V A C T V L  297
TG4           M P V R F G Q C W V F S G V L T T A L  265
TG5           Q P V R Y G Q C W V F A A V M C T V L  289
TG6           K P V K Y G Q C W V F A G V L C T V L  298
TG7           Q P V K Y G Q C W V F A S V M C T V L  289
```

1.3. Mammalian-Type TGase

Currently eight distinct TGase isoforms have been identified, at the genomic level, in mammals (Grenard 2001), however only six have been isolated and characterised at the protein level. The mammalian TGase family briefly comprises: the intracellular isoforms TG1, TG3, and TG5; TG2, expressed in a wide variety of tissues which possesses both an intra and extracellular form; TG4, localised in the prostate gland and secreted into the seminal fluid; TG6 and TG7, whose tissue localisation is unknown, Factor XIII, which exists as a zymogen that is proteolytically processed to release active factor XIII (termed Factor XIII A subunit), it is expressed in a variety of cells including blood, megakaryocytes, hepatocytes and fibroblasts (Fear 1984); and erythrocyte band 4.2, a TGase-like protein, encoding a component of the erythrocyte membrane that exhibits no enzymatic activity (for a summary see Table 1.1). TGase isoforms have been shown to be widely distributed amongst tissues (epithelium, endothelium, stratum corneum, dermis, liver, spleen, bone marrow, CNS) and physiological fluids (platelets, lymphatic system) (Griffin 2002). There is accumulating evidence to support the notion that several

TGases can be expressed in the same tissue and perform different functions (Grenard 2001; Griffin 2002).

Though all mammalian-type TGases posses very similar physical attributes, catalytic activity is not conserved throughout. TG2 and TG4 show inherent activity with FXIIIA subunit, TG1 and TG3 possessing a degree of latent catalytic activity. However, erythrocyte band 4.2, has lost its enzymatic activity completely, and only appears to serve a membrane integrity function (Lorand 2003). All mammalian isoforms share appreciable structural homology and are members of the papain-like superfamily of cysteine proteases (Makarova 1999). All mammalian isoforms also possess a certain degree of amino acid similarity and share a common active site sequence [commonly Y-G-Q-C-W-V] (Gentile 1991; Greenberg 1991; Ikura 1998). Members of this superfamily possess the conserved catalytic triad of Cys-His-Asp or Cys-His-Asn. The different primary structures and enzymatic requirements of each isoform would appear to be responsible for the diverse biological functions that have been identified for this class of enzymes.

Although these TGase isoforms differ in molecular weight and biochemical properties they all share the requirement of calcium in order to exhibit their biological effect. Deregulation of enzyme activity is generally associated with major disruption in cellular homeostatic mechanisms resulting in these enzymes contributing to a number of human disease states, including neurodegeneration, autoimmune disease, infectious diseases, progressive tissue fibrosis and diseases related to the assembly of the stratum corneum or the skin epidermis (Kim 2002; DiGiovanna 2003).

Table 1.1; Transglutaminase Properties and Biological Functions

With the exception of TGase from *B. subtilis* (regulated at the transcriptional level) TGases are post-translationally regulated.

Enzyme	Synonym	Residues (Molecular weight in kDa)	Regulation	Function
TG1	Keratinocyte TGase	814 (90)	Proteolytically activated, Ca2+ activated, Reducing agent required	Cell envelope formation during terminal differentiation of keratinocytes
TG2	Tissue TGase	686 (80)	Ca 2+ activated Reducing agent required GTP	Adhesion protein, matrix stabilisation, cell death, cell signalling
TG3	Epidermal TGase	692 (77)	Proteolytically activated Ca 2+ activated Reducing agent required	Cell envelope formation during differentiation of keratinocytes Formation of hair shaft
TG4	Prostate TGase	683 (77)	Ca 2+ activated Reducing agent required	Fertility in rodents, constituent of seminal fluid
TG5	TGase X	719 (81)	Ca 2+ activated	Important in the epidermis
TG6	TGase Y	706 (79)	Unknown	Unknown
TG7	TGase Z	710 (80)	Unknown	Unknown
Factor XIII	Fibrin stabilising factor A subunit	732 (83)	Proteolytically activated, Ca2+ activated, Reducing agent required	Blood clotting and wound healing
Band 4.2	Erythrocyte protein band 4.2	691 (77)	No enzymatic activity	Postulated important for erythrocyte membrane integrity
Bacillus TGase		245 (28)	σ factor regulated	Spore coat formation
Streptomcyes TGase		331 (37)	Proteolytically activated	Unknown

1.3.1. Factor XIII

Factor XIII is one of the most extensively studied members of the TGase family. The FXIII zymogen is comprised of AB protomers, in which the A subunits possess the catalytic potential. The molecular weight of the purified plasma protein indicates an A2B2 heterotetrameric assembly, where each A /B subunit is about 80 kDa in molecular weight. Platelets and some cells express only the A subunits in the form of an A2 homodimer with a molecular weight of about 160 kDa. The A subunits share sequence homologies with the transglutaminases, whereas the B subunits are related to the family of proteins containing small consensus or sushi domains (Lorand 2005). Factor XIII is expressed by a wide variety of cells including; macrophages, megakaryocytes and monocytes, and is found in a variety of tissues including; placenta, uterus, liver and prostate.

Analysis of the primary structure reveals the enzyme shares high similarity with other acyl transferases (Ichinose 1986). X-ray crystallography revealed that the protein is folded in 5 distinct domains: an activation peptide, an N-terminus β-sandwich domain, a core domain and two C-terminus β-barrels (Yee 1994). The presence of the common TGase family catalytic triad (Cys^{314} – His^{373} – Asp^{396}) was confirmed in the active site of factor XIIIA (Yee 1994). But, unlike many of the other TGases, factor XIIIA is a pro-enzyme, requiring cleavage by thrombin at the Arg^{37}-Gly^{38} peptide bond, prior to its activation during the final stage of the blood coagulation cascade (Takagi 1974).

The overriding function of factor XIII is in the wound healing response, it is the last enzyme generated in the blood coagulation pathway. Upon activation, Factor XIII (the active enzyme is notated as Factor XIIIa) catalyses the formation of intermolecular $\varepsilon(\gamma$-glutamyl)lysine cross-links between fibrin molecules (Lorand 1984), serving to mechanically strengthen the blood clot formed (Doolittle 1979). Factor XIII may also cross-link other coagulants (actin, von Willebrand factor, and factor V) which further reinforces the clot (Cohen 1985; Francis 1986). Factor XIII also cross-links α-2 antiplasmin to fibrin, thus increasing the resistance of clots to plasmin degradation (Tamaki 1981). Moreover, FXIIIA cross-links fibrin to thrombospondin and fibronectin,

hence anchoring the blood clot to the targeted site of injury (Bale 1986; Lynch 1987). Though the mechanism of factor XIIIA release into the plasma has not yet been determined, its co-externalisation with lactate dehydrogenase suggests that this process might be triggered by cell insult (Kaetsu 1996).

1.3.2. Keratinocyte TGase

Keratinocyte TGase (TG1, kTGase) is the largest of the nine human TGases and is synthesized as an 817-residue polypeptide giving rise to a protein with a molecular weight of 90 kDa. The full length TGase 1 apoprotein is inactive until activation via proteolysis at two sites during maturation with enzymatic activity being upregulated by terminal differentiation (Boeshans 2007). Keratinocyte TGase is induced during the terminal differentiation of keratinocytes and is capable of cross-linking specific intra-cellular proteins that contribute to the formation of the cell envelope. In the granular layer of the epidermis its major function appears to be the cross-linking a number of proteins including: keratins (Yaffe 1992); involucrin (Simon 1988); loricrin (Hohl 1993); and elafin (Nonomura 1994). Keratinocyte TGase exists in keratinocytes as multiple forms, either intact or proteolytically processed at conserved sites. Each of the isoforms have varying specific activities and quite possibly carry out different functions (Rice 1977). It has also been hypothesised that the cross-linking activity of intercellular TG1 may also play an important role in the stabilisation of the vascular endothelium barrier (Baumgartner 2004).

A three dimensional model of TGase 1 has recently been proposed, based on the crystal structure of TGase 3 (Boeshans 2007). The predicted model shows TGase 1 to comprise five domains; the N-terminal 92 amino acids (10 kDa), unique to TGase 1, shown to encode the membrane anchorage region. The remaining β-sandwich, core domains and β-barrels 1 and 2 show high similarity to the TGase 3 enzyme. It is well known that both TGases 2 and 3 are capable of hydrolysing GTP to GDP (Liu 2002). For example the binding of GTP to TGase 3 causes the loss of Ca^{2+} ions from binding site 3 with Mg^{2+} taking its place (Ahvazi 2004). The opening and closing of the active site may

be governed by the presence of different metal ions bound at site 3. On the basis of sequence conservation it has been predicted that TGase 1 may be able to shift between active and inactive states when bound to GDP (Boeshans 2007).

1.3.3. Tissue TGase

Tissue TGase (tTG,TG2) has been characterised in a wide variety of cells (e.g. vascular endothelial cells, smooth muscles cells of any origin, renomedullary interstitial cells, mesangial cells in the kidney and colonic pericryptal fibroblasts) and tissues (e.g. liver, lung, brain, kidney, adrenal glands, testis, pancreas, erythrocyte, macrophage, uterus and muscle) (Thomazy 1989).

Tissue TGase exists as a monomeric protein comprising 685-691 amino acids with a molecular weight between 77-85 kDa, depending on the species of origin (Aeschilmann 1994). Tissue TGases from mouse, guinea pig and human share 80% identity, though most notably 49 of the 51 residues that comprise the active site region are conserved. The three-dimensional structure of tTGase has been solved (Noguchi 2001; Liu 2002). The tTGase protein belongs to the α/β folding class and encodes four distinct regions; an N-terminal β-sandwich, an α/β catalytic core and two C-terminal barrel domains (Figure 1.4). This organisation into four distinct domains is echoed throughout the mammalian TGase family.

Studies performed by Lismaa and colleagues have revealed that the N-terminal domain is required for the enzymes cross-linking ability and that the core domain is essential for the hydrolysis of ATP and GTP (Lismaa 1997). The GTP binding site is located in a 15-residue hydrophobic pocket between the core and barrel 1 (Liu 2002). Guanine nucleotide binding and Ca^{2+} concentration reciprocally regulate the extent of TGase 2 transamidation activity, with guanine nucleotide binding having a negative regulatory effect. The catalytic triad is located at the base of the cavity bound by the core and barrel 1 domain. The catalytic Cys^{277} in tTGase is critical for the formation of the thiolester bond with a substrate bound glutamine (Lee 1993).

Hence, it has been concluded that the catalytic core alone is not sufficient enough for the TG activity; hence the flanking domains also play a role in the modulation of the transamidating activity of TGase. Interestingly, the GTP-binding domain of tTGase contains almost all of the conserved tryptophans of the enzyme and is thought to be responsible for the stabilisation of the transition state (Murthy 2002). It has also been shown, by the latest site-directed mutagenesis analyses, that Trp[241] is critical for tTGase cross-linking activity and is highly conserved in all of the TG members with the exception of the catalytically inactive erythrocyte band 4.2 (Murthy 2002).

Figure 1.4; Schematic Structure of Tissue Transglutaminase.

Backbone structure of tTG. Domains I-IV are coloured green, magenta, purple and yellow respectively. The area encompassing the active site residues Cys^{277}, His^{335} and Asp^{358}), has be enlarged.

Image, based on the crystal structure of tTGase (Liu 2002) in the GDP bound form, drawn using PyMol Software (DeLano 2002).
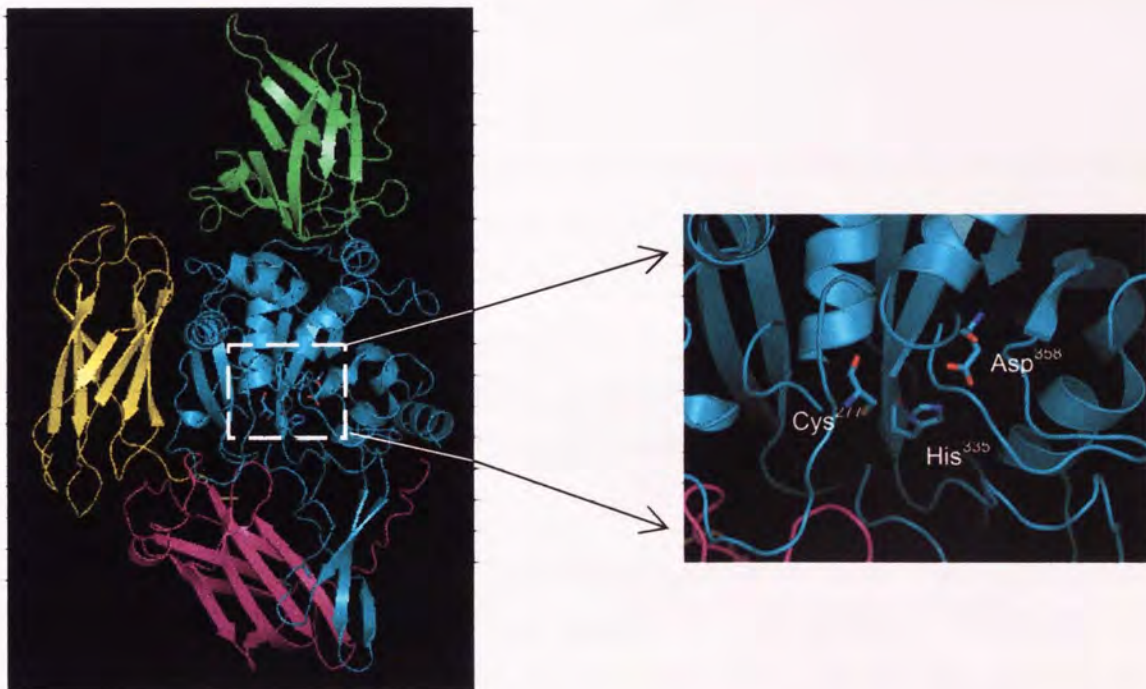
1.3.4. Epidermal TGase

Epidermal TGase (TG3, eTGase), initially isolated from human callus, bovine stratum corneum and snout epidermis (Buxman 1975), is localised in the cytosol and exists in a latent conformation (77 kDa) requiring proteolytic cleavage for activation. Following cleavage, conformational changes facilitate Ca^{2+} binding thereby increasing the enzymes specific activity 2-fold (Kim 1995). The crystal structure of TGase 3 has been solved (Ahvazi 2004) which has mapped the locations of the calcium and GTP binding sites, indicating that GTP hydrolysing activity is not restricted to TGase 2. Both TGase 2 and 3 bind GTP in the same region though they utilise different amino acid residues. As with TGase 2, in TGase 3, GTP displaces one of the three calcium binding sites causing a conformational change that results in the closing of the central cavity, preventing access to the active site. Some researchers believe TGase 3 to play a role in hair shaft formation, through the cross-linking of structural proteins (trichohylain and keratin intermediate filaments) (Lee 1993) whilst others believe it to be involved in the formation of the cornified cell envelope precursor proteins of the epidermis during terminal differentiation of keratinocytes (Rice 1977).

1.3.5. Prostate TGase

The prostate TGase (TG4,pTGase) gene encodes a 684 amino acid protein, which exists as a 150 kDa homodimer, that is strictly expressed in huminal epithelial cells of the prostate (Dubbink 1996). TGase 4 has been shown to circulate in the seminal fluid following secretion by the dorsal prostate and coagulating gland in rats, where it is believed to be involved in the formation of the copulatory plug by cross-linking seminal vesicle proteins SVP-1 and SVP-IV (Williams-Ashman 1984; Seitz 1991). The enzyme is GTP regulated with the GTP activation site being located in the N-terminal 103 amino acids (Marinello 2003). This N-terminal region is also deemed essential for enzymatic activity, even though it does not encompass the active site residues. It has been proposed that TGase 4 may play a role in suppressing the immune response elicited by immuno-competent cells in the female tract against sperm cells (Paonessa 1984).

1.3.6. Transglutaminase 5, 6 and 7

Transglutaminases 5, 6 and 7 have only recently been discovered. Though their structures, gene localisation and organisations have been revealed only the function of TG5 has been determined (Griffin 2002). TGase 5 has been shown to be a unique member of the TGase family (Aeschilmann 1998). TGase 5 exists in four different isoforms including active full length TGase 5, active splice variant $\Delta13$ and inactive splice variants $\Delta3$ and $\Delta3\Delta13$ (Candi 2001). TGase 5 has recently been shown to be a dual-function enzyme capable of both calcium-dependant transamidation activity as well as GTP binding and calcium-dependent GTP hydrolysis (Candi 2004). Under physiological conditions TGase5-transamidating activity is regulated by localised concentrations of both calcium and nucleotides, as is the case for TGase 2 (Haroon 1999) TGase 5 appears to be localised within the upper layers of the epidermis and has been implicated in keratinocyte differentiation (Candi 2001).

1.3.7. Erythrocyte Band 4.2

Erythrocyte band 4.2 (pallidin) is a 77kDa protein, encoded by a 20kb gene. It is a major structural component of the cytoskeletal network underlying the red blood cell membrane and is often expressed at high levels in erythroid cells (Cohen 1993). Erythrocyte band 4.2 shares close sequence similarity with other TGase family members, though it remains the sole catalytically-inactive member, possessing an alanine in place of the catalytically essential cysteine residue in the active site (Korsgren 1990). As such the protein is believed to play a structural role in red blood cell integrity (Sung 1992).

1.4. *Eukaryotic TGase Enzymes*

1.4.1. *Fish Derived TGase*

TGase has been described in a number of marine vertebrates including in the liver of *Pagrus major* (red sea bream) (FTG) though its exact biological function remains unknown. The enzyme is 78 kDa in size shares 33% identity, Ca^{2+} dependency and the catalytic triad in common with the mammalian-type TGases mentioned previously (Nielsen 1995). The catalytic resides of FTG adopt a similar conformation when compared with their mammalian counterpart (Noguchi 2001). However unlike TG2, which works as a homodimer, FTG functions as a monomer and does not require initial proteolysis for its activation. In certain fish marine species, namely scallop, botan shrimp and squid, sodium chloride has been found to enhance TGase activity 11-fold, 2-fold and 6-fold respectively leading to the hypothesis that there exists novel TGase isoforms in marine invertebrates that are capable of enzymatic activity in high osmotic pressures (sea water) (Nozawa 1997).

1.4.2. *Lower Eukaryote Derived TGase*

Both unicellular and small multicellular organisms have derived mechanisms whereby they are able to protect themselves from the harsh environmental conditions, as well as withstanding complex life cycles. An important protection mechanism involves the formation of rigid and inert structures that are impervious to mechanical disruption and chemical attack. Such structures are composed of complex carbohydrates and cross-linked protein aggregates (Chandrashejar 2000). It would be feasible to assume that the protective structures synthesized are as a result of a protein cross-linking reaction catalysed by TGase.

TGase has been described in nematodes where it is thought to be involved in development, growth and maturation of embryos and other larval stages. Evidence for this involvement is mounting based on the observation that membranous structures, for example the cuticle, epicuticle and sheath are highly enriched with TGase-catalysed $\varepsilon(\gamma$-glutamyl) lysine isopeptides.

Additionally, inhibition of TGase was found to completely block the differentiation from the larval stage (Chandrashejar 2000).

The slime mold, *Physarum polycephalum,* exhibits a unique life cycle that displays both ameboid and plasmodial forms. The plasmodia are giant, multinucleated cells with a veined structure and no internal cell walls. Plasmodia transferred to non-nutrient media differentiate into dormant spherules, with hard outer walls, also starvation of plasmodia along with light exposure induces the formation of haploid spores (Wada 2002). When plasmodia differentiate into spherules there is a 10-fold increase in TGase activity, suggesting a role in spherule formation. Additionally, upon spherule damage TGase may serve to breach the lesion and protect the organism (Wada 2002). The active enzyme is shown to be present in both microplasmodia and spherules. It is a thiol enzyme, dependent on calcium and is 101kDa in size (Klein 1992). Its molecular mass is slightly higher than those for the organisms previously mentioned, perhaps suggesting a degree of species specialisation. Likewise it is calcium dependent and is inhibited by GTP, like the tTGases of vertebrates. To date there have been limited discoveries surrounding novel TGases from such organisms.

1.4.3. Plant Derived TGase

At present, several TGase activities have been detected both in higher and lower plants, supporting the presence of this type of enzyme. This type of activity has been found in a range or organs, including; chloroplasts; thylakoids; mitochondria; cytoplasm and the cell membrane of green leaves, from both higher and lower plants. Calcium is deemed essential for catalytic activity though the exact concentration is not known since there is a high availability of Ca^{2+} in plant cell-free extracts. In fact in higher concentrations it was found to be inhibitory (Serafini-Fracassini 1995). Though the exact biological role remains to be determined it is though that they contribute to structural or conformational processes. They may also play a role in organelle-specific metabolism and plant TGase activities have also been related to growth, differentiation, programmed cell death and stress (Serafini-Fracassini 2002).

Research on plant TGases has been greatly delayed so far because the identity of the responsible proteins remains unknown. Also no DNA sequence, with significant homology to animal TGases, has been identified from the genomes of Arabidopsis, *Zea mays*, *Oryza sativa*, and *Solanum tuberosum* making sequence comparison analysis for plant derived TGases difficult. Recent computational analysis has concluded that the *AtPng1p* gene, originating in Arabidopsis, encodes a putative *N*-glycanase. AtPng1p contains the Cys-His-Asp triad present in the TGase catalytic domain. The recombinant protein could be immuno-detected using animal TGase antibodies. Furthermore the recombinant protein has been shown to link various polyamine substrates into dimethylcasein in a calcium-dependent manner confirming that the *AtPng1p* gene product acts as a TGase (Della Mea 2004).

1.5. Bacterial Transglutaminase

TGase has only been purified, characterised and cloned from a small number of related microbes including members of the *Streptoverticillium sp.* and *Bacillus sp.*. These enzymes are distinct from their eukaryotic counterparts in that they do not require calcium for activity. *Streptoverticillium sp.* TGase is an extracellular enzyme whereas *Bacillus sp.* TGase acts on an intracellular environment. The DNA sequences of the two TGases from *Stv. mobaraensis* and *B. subtilis* share little similarity except for a highly conserved cysteine residue in the active site region necessary for enzymatic activity (Kanaji 1993; Kobayashi 1998b).

1.5.1. Bacillus subtilis *Derived TGase*

During sporulation, members of the *Bacillus* family differentiate into two cell types. As a result the spores produced are known to protect against a number of environmental stresses including heat (Fox 1969), dessication (Setlow 1999), chemicals (Bloomfield 1994; McDonnell 1999) and UV radiation (Setlow 1988). Spore formation is initiated and controlled by a complex gene expression program described below.

External stimuli present in the appropriate amount, including carbon, nitrogen and phosphate deprivation, serve to initiate a complex phosphorelay cascade (Dempfle 1992) controlled by the sporulation global regulator (*SpoOA*), which directly regulates the transcription of 121 genes and indirectly controls a further 400 genes (approx) (Fawcett 2000). A series of transcription factors, are transcribed during distinct stages of sporulation and serve to direct the transcription of genes from specific promoters. After the daughter chromosome has been safely packaged inside the forespore, and cell-wall like material has been deposited between its membrane layers, the spore coat is constructed. The spore coat, visible around five hours after initiation, is synthesized from proteins manufactured within the mother cell which are subsequently assembled around the forespore prior to release of the mature spore into the surrounding environment (Figure 1.4).

Figure 1.4; Morphogenesis and Gene Regulation during Sporulation

Coat protein genes active at each stage. Repressive actions of GerE and SpoIIID not shown
Adapted from (Errington 1993)

Vegetative Cell	Stage II	Stage III	Stage IV	Stage V	Stage VI
	Septation and Engulfment	Prespore Protoplast	True Coretex	Coat Formation	Maturation and Release

cotE	cotE	cotA	cotB
spoIVA		cotD	cotC
spoVID		cotF	cotG
cotJ		cotH	cotS
		cotM	cotV
		cotT	cotW
		cotV	cotX
		cotW	cotY
		cotX	cotZ
		cotY	
		cotZ	
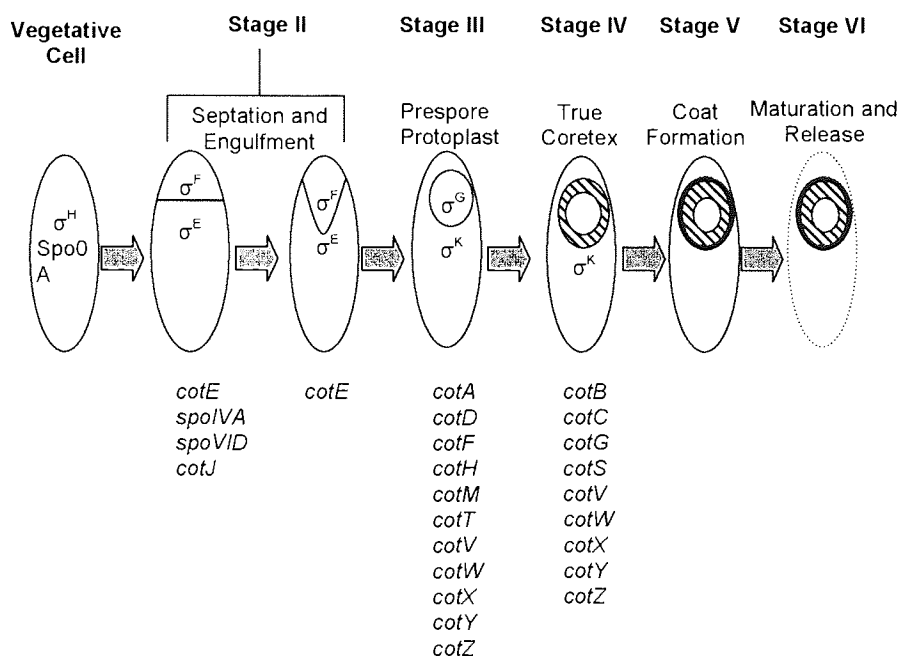
It is during the stages concerning spore coat formation, where ε-(γ-glutamyl) lysine bonds have been identified (Santangelo 1998) suggesting the involvement of a transglutaminase. It has been deduced that the spore coat fraction and spore coat protein possess one cross-link for every 6.7×10^6 and 8.3×10^6 Da protein respectively (Kobayashi 1998b). This amount of ε-(γ-glutamyl) lysine cross-link is similar to the amount found in *Escherichia coli* and *Paramecium aurelia* (Matacic 1979). The characteristic isopeptide bond is only detectable during this late phase of sporulation, not during growth, indicating that transglutaminase transcription is isolated to sporulation and hence is under the control of one of the sigma factors responsible for directing gene expression at that particular stage. Spore coat assembly predominantly takes place in the mother cell where gene expression is controlled by the transcription factors σ^E, SpoIIID, σ^k and GerE (Errington 1993). It is the latter of these sigma factors (σ^k and GerE) that are responsible for transcription and ultimately regulation of cross-linking, since TGase is only produced after the chromosome has been safely packaged and the spore coat is ready for cross-linking. *Bacillus subtilis* is the first microbial transglutaminase shown to act on an intracellular environment.

Transglutaminase from *B. subtilis* is encoded by a gene, 735 bp in length. Analysis of the sequence directly upstream of the gene reveals the presence of a GerE binding site highlighting that TGase is expressed solely during the later stages of sporulation (Kobayashi 1998b). GerE is a small DNA-binding protein that acts in conjunction with σ^k to either positively or negatively regulate gene expression and has been shown to direct the transcription of genes encoding structural components of the protein coat which encases the mature spore (Zheng 1992). *Bacillus sp.* TGase has a molecular mass of 28.3kDa, a pI of 7.12 and its catalytic activity is centred on cysteine, histidine and asparagine as are all TGases. Unlike its eukaryotic counterparts, however, the enzyme is calcium independent and exhibits maximal enzymatic activity at pH 8.2 and at a temperature of 60 °C (Suzuki 2000). Since TGase from *Bacillus subtilis* bears little resemblance to either *Streptomyces mobaraensis* or the eukaryotic TGases, it would appear that it has arisen from an alternative evolutionary pathway.

Within the Bacillus family, two other enzyme isotypes have been identified, in *B. halodurans* (a halophile), to which it shares a high degree of homology and *B. circulans* (de Barros Soares L 2003). The latter of which has been isolated recently from an aquatic Amazonian environment, is 45 kDa in size, larger than previously reported forms of microbial TGase and as such may present different exploitable characteristics, provided sufficient protein may be obtained. To date TGase from *Bacillus sp.* has yet to be applied on an industrial scale.

1.5.2. TGase from Streptomyces S-8112

Initial studies on the *Streptomyces* genus recognised *Streptomyces* and *Streptoverticillium* as two distinct genera, both comprising; a type 1 cell-wall typified by the presence of L-diaminopimelic acid and glycine with an absence of characteristic cell wall sugar (Lechevalier 1970), that were lysed by the same phages (Wellington 1981) and were phylogenetically closely related (Stackebrandt 1981). Additional studies highlighted similar physiological traits (Kampfer 1991) with differences only being observed for DNA-RNA pairing (Gladek 1985) and the ability of the *Streptovericillia* to form whorls. On the basis of 16S and 23S rRNA comparisons the genus *Streptoverticillium* has been unified with the *Streptomyces* genus (Witt 1990).

Streptomyces belong to a large group of Gram positive, aerobic, filamentous soil bacteria that exhibit a complex life cycle, characterised by at least three distinct morphological stages involving the formation of a system of elongated, branching filaments which, after a period of vegetative growth, respond to specific signals by producing specialized spore-bearing structures. The spores generated are relatively resistant to desiccation, enzymatic digestion and moderately high temperatures (Ensign 1978; McBride 1987), allowing the organisms to persist for a limited amount of time in an unfavourable environment.

Culture on; agar slopes; media containing glucose, yeast and malt extracts allows the development of substrate and aerial mycelia leading to the formation of spores (Zotzel 2003a). At this stage, thought to be brought about

by nutrient limitation (Karandikar 1997), many vegetative cells die with the accumulated materials being used to provide the nutrients necessary for hyphal growth (Chater 1989). After the completion of hyphal growth the development of spores ensues by the sequential insertion of cell walls and cell-wall thickening. The onset of aerial hyphal growth is closely associated with the secretion and activation of a number of hydrolases, the functions of which are not well understood. It would appear that they have more important roles in regulating cellular differentiation as well as the digestion of substrate mycelium to supply nutrients to aerial hyphae development. Indeed, recent results suggest that mycelium differentiation may be comparable to the events of programmed cell death in eukaryotes (Nicieza 1999; Fernandez 2002).

Members of the *Streptomyces* have been shown to secrete large amounts of TGase into the culture medium ($\approx$20 mg/L) (Ando 1989a).Transglutaminase from *Streptomyces mobaraensis* has been described as a Ca^{2+} independent enzyme of 38 kDa that is secreted as an inactive precursor comprising an activation peptide 45 amino acids in length which serves to suppress activity and to increase the thermostability of the zymogen (Kanaji 1993).

Though the exact role of TGase in the *Streptomyces* is unclear, it is feasible to assume that the enzyme plays a role in mycelial growth and morphological differentiation by strengthening the cell wall through the introduction of protein cross-links.

1.5.2.1. Characteristics of Microbial TGase

The gene encoding transglutaminase from *Streptomyces mobaraensis* is 1221bp in length, which is preceded by a potential ribosome binding site (Washizu 1994). Analysis of the complete sequence reveals that the TGase is synthesized as a precursor protein of 406 amino acid residues. Of these, 75 aa comprise a prepro region whilst the remaining 331 aa encode the mature protein. Further analysis of the prepro region reveals that the 57 amino acids comprising the pro-region are deemed important for efficient protein secretion and folding. The remaining 18 amino acids, located at the

N-terminal (pre-region), may function as a signal peptide, which are subsequently cleaved during membrane translocation. Among the *Streptomyces* these signal peptides show great diversity, but do however, share a similar organisation. Each signal peptide comprises a typical tripartite structure; a basic amino-terminus, a central polar region and a carboxy-terminal region containing the signal peptidase recognition site. The zymogen is secreted into the surrounding medium whereby it is proteolytically processed to release the active enzyme (Pasternack 1998). mTGase has a pI of 8.9 and a molecular mass of 37.9 kDa (Kanaji 1993). In contrast mammalian TGase (Factor XIII) has a molecular weight of 83 kDa and a pI of 4.5. An enzyme, such as microbial TGase, that is calcium independent, possesses a higher reaction rate, broader substrate specificity for the acyl donor, lower activity for deamination and a smaller molecular size have enabled the enzyme to be utilised in a wider variety of industrial applications than mammalian type TGase isoforms (Nielsen 1995).

1.5.2.2. Overall Structure of mTGase

The crystal structure of mTGase from *S. mobaraensis* has been solved (Kashiwagi 2002a) and has been shown to consist of novel overall and active site structures. MTGase exists as a monomeric protein comprising a single polypeptide chain that adopts a disc-like conformation (dimensions 6 x 59 x 41 Å) with a deep cleft on one side housing the active site cysteine at the base (Figure 1.4). MTGase belongs to the α/ß folding class of proteins, as do eukaryotic TGases, containing 11 α-helices and 8 β-sheets. The α-helices and β-sheets are mainly concentrated at the amino and carboxy terminals of the polypeptide respectively.

Electrostatically, the outer surface of mTGase is predominantly covered by positively charged amino acids whilst negatively charged amino acids mostly occupy the active site cleft (Figure 1.5). Within the active site cleft there exist a locally high number of acidic residues with a number of aromatic residues around the top of the active site cleft. Such characteristics would serve to govern the substrate specificity of mTGase.

Figure 1.5; Van der Waals Model of mTGase

Front and rear views of *Stv. mobaraensis* transglutaminase (left and right respectively). The active site cysteine (Cys64) is coloured red, aromatic and hydrophobic residues are coloured blue and green respectively. Diagram drawn using PyMol imaging software (DeLano 2002).
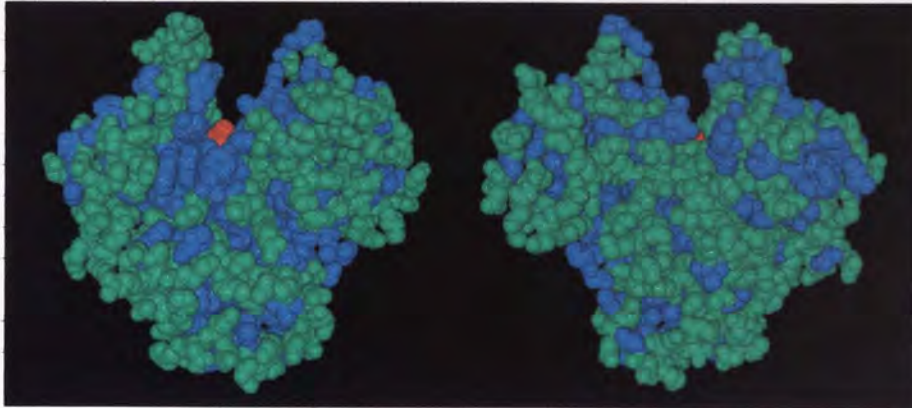
Sequence analysis of the mature protein identifies the presence of a single cysteine residue. Activity assays in the presence of thiol inhibiting compounds (Iodoacetamide, N-ethyl-maleimide) have confirmed the presence of a sulfhydryl group involved in enzymatic activity (Ando 1989a). Indeed mature TGase contains a sole cysteine residue at position 64 essential for activity that is predicted to reside in a β turn connecting the α helix and β sheet structures as found within mammalian TGases. Indeed the secondary structures of both mTGase and mammalian TGases share a high degree of similarity.

1.5.2.3. Activation of mTGase

TGases of eukaryotic origin have evolved mechanisms to control the extent of enzymatic activity, to include GTP (in the case of TG2 and to some extent TG5 and 3) and calcium dependence. In the *Streptomyces* however, it would appear that post secretion, the inactive zymogen is post-translationally modified to release the active, mature protein.

TGase produced by *S. mobaraensis* is a 42.5kDa zymogen which is activated by the removal of the 45 amino acid N-terminal pro-peptide. During the course of cultivation the inactive zymogen is processed to its active form

via a distinct pathway (Figure 1.6). The zymogen is co-secreted into the surrounding medium with its activator, P1`-endoprotease TAMEP, and is subsequently processed by hydrolysis of the peptide bond between Phe(4) and Ser(5), (Zotzel 2003a) resulting in the removal of 41 amino acids at the N-terminus generating the intermediate FRAP-TGase, which exhibits full enzymatic activity. This process is strictly regulated by a 14 kDa protein (P_{14}), related to the Streptomyces subtilisin inhibitors, that elicits its control indirectly by negatively regulating the activating protease (TAMEP) (Zotzel 2003a). Cleavage of pro-TGase by TAMEP leaves a tetrapeptide at the N-terminus of the protein. Since the intermediate exhibits full enzymatic activity it is reasonable to assume that removal of this tetrapeptide is a relic of an aminopeptidase secreted with TGase. The remaining tetrapeptide, however, can be successfully removed by a tripeptidyl aminopeptidase (TAP), isolated from the culture broth, which generates the mature N-terminus of TGase in a single step reaction that is not sensitive towards P_{14} (Zotzel 2003b).

It has been proposed that the function of the tetrapeptide may be to regulate amounts of active TGase by retaining the partially processed enzyme (FRAP-TGase intermediate) in the murein layer by means of ionic interactions. Such interactions would be overcome once the tetrapeptide had either been processed or salt concentration had risen high enough (Zotzel 2003b). It should be noted that this hypothesis has yet to be substantiated.

After activation by proteolytic cleavage. It has been shown that the precursor protein, secreted into the medium, from S. mobaraensis actually inhibits the activity of the mature enzyme (Pfleiderer 2005) until it has been completely digested by TAMEP. Also detailed in a different Streptomyces sp., namely Streptomyces lavendulae, is a high molecular weight, melanin-like, TGase inhibitor that has been shown to competitively inhibit a number of TGases, both prokaryotic and eukaryotic (Ikura 2000).

Figure 1.6; Schematic Representation of mTGase Processing

(Adapted from (Zotzel 2003b)). Pro-TGase is cleaved by TAMEP releasing the 41 aa pro-peptide. P_{14} may negatively regulate the reaction by binding to TAMEP in equimolar amounts (Zotzel 2003a). FRAP-TGase exhibits full enzymatic activity with the final processing step resulting in release of the true mature form of TGase. This final reaction is promoted by low concentrations of Ca^{2+}.
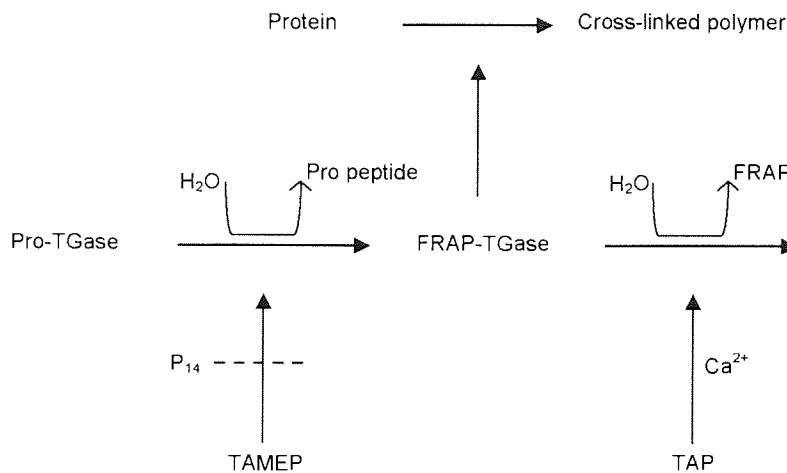
1.5.2.4. Structure of the Active Site Cleft

The similarity of the catalytic triad and reaction mechanism indicates that TGases share the core structural fold with the thiol proteases. It is therefore likely that microbial homologues of TGase may act as proteases and that mammalian TGases may have evolved from an ancestral protease. It is well known that eukaryotic TGases possess an active site reminiscent of the thiol proteases. However the active mechanism of mTGase is centred around a sole cysteine (Cys 64) necessary for catalytic activity (Kanaji 1993).

The N-terminus of mTGase occupies the entrance to the active site suggesting that it serves to prevent substrates entering until the N-terminal region responsible has been proteolytically cleaved to release the mature enzyme. The hydrophobic environment of the active site region, is similar to those of other TGases as well as the thiol proteinases (e.g. papain and cathepsin) suggesting some degree of evolutionary relationship. The active site cysteine residue is considered to play an essential role in the acyl transfer reaction, since erythrocyte band 4.2 protein (Nakanishi 1991), has

this cysteine residue replaced by an alanine, and consequently shows no TGase activity.

1.5.2.5. Comparison of TGase Active Sites

To date the crystal structures of Factor XIII A subunit (Yee 1994) *Pagrus major* (Pedersen 1994), TG2 (Liu 2002), TG3 (Ahvazi 2004) and *S. mobaraensis* TGase (Kashiwagi 2002a) have been determined. Whereas the overall structures of Factor XIII, FTG and TG3 resemble one another the structure of mTGase is completely different, though this is not entirely unexpected due to the differences in sequence similarity and molecular size. Despite this however, the active site arrangements share a high degree of similarity (Figure 1.7).

The active site cysteines, Cys64 in mTGase and Cys272 in FTG (Taken here as representative of the eukaryotic TGases), both reside near the N-terminus of α-helices. Such an arrangement is also seen in cysteine proteases, subtilisin proteases and the α/β hydrolases (Pedersen 1994). This α-helix is flanked by a 4-stranded β-sheet in each enzyme. In *P. major* the remaining residues of the catalytic triad (His332 and Asp355) occupy locations on the central strands of this β-sheet, an arrangement which is shared by factor XIII and some cysteine proteases, most notably the papain (Drenth 1976) and actinidin proteases (Baker 1980).

Figure 1.7; TGase Active Site Comparison

The active site regions of *S. mobaraensis*. TGase (left) and TGase from *Pagrus major* (right). Image visualised using PyMol imaging software (DeLano 2002).

The arrangement of a cysteine protease-like catalytic triad (Cys-His-Asp (Asn)) is not seen in mTGase, which constitutes a major difference in structure between mTGase and its eukaryotic counterparts. In mTGase Asp255 and His274 occupy positions relative to those of His332 and Asp355 in FTG, respectively (Figure 1.6) suggesting that they possess both an altered substrate specificity (de Jong 2001) and a difference in substrate recognition (Taguchi 2000).

1.6. Biotechnological Applications of TGase

TGases post translationally modify proteins by the formation of ε(γ-glutamyl) lysine bonds resulting in protein polymers that are chemically stable, mechanically strong and enzymatically resistant to proteases. Such characteristics make them attractive biocatalysts with applications in a number of industries. Due to the scarce source, complicated purification procedure and high cost, mammalian TGases (gpl-TGase and FXIIIA) are yet to be applied to large scale industrial procedures (Berovici 1987) such as the modification of food proteins. Microbial TGase provides a commercially viable alternative to mammalian TGase. As such there has been considerable patent activity in the area (Motoki 1992; Andou 1993; Bech 1996; Kobayashi 1998c). MTGase is commercially available from the Ajinomoto Co. Ltd. (Japan), Novo Nordisk (Denmark) and Yiming Fine Chemicals (China).Of those companies, Ajinomoto market TGase produced by *S. mobaraensis* whilst Novo Nordisk market TGase from *S. platensis*, the organism used by Yiming Fine Chemicals has yet to be disclosed though it is likely to be closely related to both *S. mobaraensis* and *S. platensis*.

1.6.1. Applications of Mammalian TGases

TGase has been extensively applied to the medical area. It has been used as an effective form of substitutive therapy in patients with rare genetic defects of blood clotting, relating to the loss of Factor XIII (Gootenberg 1998). TGase, predominantly Factor XIII, is an essential component of a 'Biological Glue' used extensively by surgeons to aid the post-operative repair of surgical wounds, fractures and cartilage lesions (Jurgensen 1997). This

methodology exploits the large-scale production of the recombinant enzyme and is being explored in surgical practice as a possible treatment for intestinal disease (D'Argenio 2000). An alternative to the direct application of tTGase involves the potential modulation/induction of tTGase expression via specific inducers such as the retinoids. Such an approach has been used *in vivo* to treat selected malignancies (Jetten 1990; Lentini 1998) and has also been considered as an alternative treatment for certain dermatological disorders (Bershad 2001).

Eukaryotic-type TGases have also been applied to the food industry where they have been found to cross-link a variety of food proteins, including α-casein, β-casein, κ-casein and β-lactoglobulin. However the biochemical application of mammalian-type TGases for use in the food industry has been limited since the costs associated with producing sufficient amounts of enzyme are too high. In attempts to produce sufficient amounts of enzyme mammalian TGase was expressed in *E. coli* (Ikura. et al 1990). However, the enzyme produced was of moderate stability, as a function of the large amount of cysteine residues, none of which comprised disulfide bonds (Ikura et al 1988). In addition mammalian TGase has also been expressed in *Saccharomyces cerevisiae,* with the expression of enzyme being dependent on the carbon source. Medium optimisation has resulted in a yield of up to 100mg/L Factor XIIIa (Broker 1991). Factor XIII may be obtained from the blood of swine and cattle at slaughter (Wilson 1992) though the blood enzyme is rarely utilized in food manufacture, since a specific protease (thrombin), is required to activate the enzyme and the red pigmentation is often detrimental to product appearance. Another unfavourable characteristic is that Ca^{2+} is required to activate eukaryotic TGases, bringing about bitterness in the end product.

1.6.2. Applications of mTGase

The possibility of modification of functional properties in milk caseins and soybean globulins has been demonstrated using TGases derived from guinea pig liver (Ikura 1992) or bovine plasma (Kurth and Rogers 1984). It was found that whey proteins, actomyosin from beef, pork, chicken or fish

could be effectively gelled by TGase giving rise to improvements in solubility, water holding capacity and thermal stability, indicating the potential of using TGase to create proteins with unique functional properties (Motoki and Seguro 1998). Unfortunately, TGases of eukaryotic origin have never been a commercially viable option for large industrial exploitation and so far have only found applications in niche markets in the medical arena. In order to facilitate mass production at a more reasonable cost a large scale "hunt" for microbial sources of transglutaminase with the potential to be used as applied biocatalysts was undertaken by Ajinomoto Co. Inc. (Japan) and Amano Pharmaceutical Co. (Nagoya, Japan) (Motoki 1992).

A variant of *Streptomyces mobaraensis* (formerly classified as *Streptoverticillium mobaraensis)* which exhibited high cross-linking activity (Washizu 1994) was identified and subjected to further study. Microbial TGase from *Streptomyces mobaraensis* was found to be totally independent of Ca^{2+} (Ando 1989b) and as such is quite unique from other mammalian enzymes. It also exhibits a wider pH and temperature range (stable between pH 6-9 and retains activity at 50°C as well as a proportion of activity at temperatures just above freezing), properties that make it more appealing for commercial/industrial usage.

1.6.2.1. mTGase and Food

Within the global marketplace there is increasing pressure on manufacturers to produce highly nutritious food products. This often involves the functional modification of proteins to produce products with greater consumer appeal. Enzymatic modification of food proteins exhibits greater specificity and increased safety over non-enzymatic methods (Ikura 1980). Microbial TGase is capable of gelling proteins such as soybean proteins, milk proteins, beef, pork, chicken and fish gelatin and myosin in a similar way to that of the guinea pig liver enzyme (Motoki and Seguro 1998). The initial commercial application of mTGase centred on the production of fish paste (surimi) products for the Japanese market. After completion of the setting process, gels produced were deemed to have greater breaking stresses and showed improved functionality (Zhu 1995). The effects generated were shown to be

as a result of TGase catalysed crosslinking of fish proteins, with up to 3 μmol of ε-(γ-glutamyl)lysine/100 g gel (Kumazawa 1995).

Microbial TGase has also been applied to a large variety of food processes and shown to shown to cross-link other meat and fish products and has the ability to effectively bind meat pieces and particles, important in the production of restructured and processed meats (Eligula 1998). Reformed and restructured meat products are traditionally produced using salts and phosphates that promote the extraction of myofibrillar proteins, after cooking these form stable protein matrices (Carballo 2006), though these matrices are not stable in the raw product leading to consumers selecting fresh products. Using mTGase to gel meat pieces reduces the need to add salt and phosphate adding to health benefits (Wijngaards and Paardekooper 1988). Mixing TGase with caseinate forms a viscous gel which can be used to bind restructured meat pieces together at low temperatures (Kuraishi 1997). There have been a number of studies on the use of mTGase in meat products including beef (Pietrasik 2003), pork (Pietrasik and Jarmoluk 2003), chicken (Kilic 2003) and lamb (Carballo 2006).

The most advanced area of dairy product processing using mTGase is in yoghurt manufacturing (Lorenzen 2002). Milk caseins have been shown to be excellent substrates for TGases due to their open structure (Traore and Meunier 1992) and it has been well demonstrated that incubating milk with mTGase results in a yoghurt with an improved gel strength (Ozer 2007). Other characteristics of mTGase cross-linked yoghurts include decreased post-acidification during storage, lower whey separation, milder taste as well as a smoother, dryer and slightly whiter surface than the untreated yoghurt (Ozer 2007).

Other uses within the food industry include improving the functional properties of gluten proteins in baking leading to an improved crumb texture. TGase can be used to improve dough properties and baking performance for flours with weak gluten and poor baking performance (Wang 2007). Microbial TGase has been shown to directly affect the properties of functional proteins under high hydrostatic pressure, effects that could not be induced under

atmospheric pressure (Lauber 2003). This presents new possibilities in food production and protein modification. Newer isoforms of the enzyme would undoubtedly open new avenues to pursue.

1.6.2.2. mTGase and Leather

In the leather industry, traditionally casein has been used to coat leather in a complex procedure involving the use of toxic hardening agents such as aldehydes, isocyanates or aziridine. A more environmentally friendly method of treatment involves treating hides with a combination of casein and mTGase, applied as a coating and subsequently dried. The finished product has similar properties to the sole casein coatings without the associated toxicity problems (Rasmussen 1996).

1.6.2.3. mTGase and Textiles

Although disulphide bonds are the most abundant cross-links in wool and are the main bonds stabilising wool fibres, several other types of cross-links may be present in small amounts (Cortez 2004). Such bonds include the isopeptide bond, formed as a function of TGase-mediated cross-linking, which serves to further stabilise the fibre matrix (Nakachi 2001). Additionally TGases may used be to alter other properties of the wool proteins with an increase in the amount of $\varepsilon(\gamma$-glutamyl) lysine cross-links being responsible for an increase in wool strength and reduced shrinkage (Cortez 2005). Traditional anti-felting treatments, involving the use of specific proteases (subtilisin) to degrade the scales on the wool fibre, thereby reducing the friction experienced during fibre root movement, leading to the finished product having limited strength and decreased weight (Shen 1999), properties which are unattractive to the consumer. Treatment with TGase has been shown to reverse the negative effects caused, in some instances completely reversing the loss caused by protease treatment (Fornelli 1993). In the future microbial TGase will no doubt find an increased number of applications in industry with newer forms of the enzyme being actively sought after.

1.6.3. Commercial Production of mTGase

Transglutaminase derived from *Streptomyces sp.* is produced by the culture of the micro organisms followed by subsequent purification. This conventional process for producing mTGase introduces problems with regards to supply amount and production costs. As such mTGase has been produced recombinantly in *E. coli*, *Corynebacterium glutamicum*, *Saccharomyces cerevisiae* and *Streptomyces lividans* (Takehana 1994; Washizu 1994; Yurimoto 2004) However, even the most successful of these approaches (Washizu 1994) resulted in a production yield lower than that of the wildtype strain despite extensive optimisation.

The *Streptomyces sp.* TGase gene has been chemically synthesized and expressed in the periplasm of *E. coli*, but the expression level was low (Takehana 1994). The protein was also produced as inclusion bodies and subsequently purified, however the yield and specific activity of the mTGase was low (Kawai 1997). Further attempts have been made to synthesize the mTGase gene taking into account the codon usage of *E. coli*. Despite an improvement in the expression level the TGase was expressed as inclusion bodies and the subsequent protein recovery was poor (Yokoyama 2000). TGase from *Streptoverticillium ladakanum* BCRC 12422 has been expressed in *E. coli* as inclusion bodies though enzymatic activity was not recovered after protein refolding (Tzeng 2005). Since mTGase acts independently of Ca^{2+}, it is likely that expression of active mTGase in the micro organism is fatal since the enzyme acts on the endoprotein. Expression as inclusion bodies is not without its limitations, since it is necessary to solubilise with a protein denaturant, the denaturant is subsequently removed and the mTGase reactivated to obtain the active mTGase. These processes have problems when practiced on an industrial scale. The amount of product is often small; this then requires an expensive purification procedure.

Currently, to help alleviate this problem, TGase has been purified from *Streptomyces fradiae* and cloned into both *E. coli* (Liu 2006a) and the parent strain (Liu 2006b) resulting in an enhancement of TGase production. In *E. coli* the TGase protein, expressed as inclusion bodies, is effectively refolded

utilising a cation exchange chromatography-type technique giving a yield of 105mg (active TGase)/400ml culture. In the parent strain TGase production was improved 1.3 fold. The improved recovery of active enzyme from both inclusion bodies and from the parent strain demonstrates the possibility of cheaper production costs provided the procedures may be further optimised.

Microbial TGase, due to its novel 3-D structure and active site orientation, exhibits different substrate specificities to mammalian-type TGases. Indeed the substrate reactivity of mTGase may be altered, as demonstrated by Kashiwagi (Kashiwagi 2002b). Mutation of residues proximally located to the active site resulted in improved substrate reactivity with primary amines. This illustrates the ability to potentially enhance TGase activity making it a more attractive prospect for industrial processes.

1.7. Project Aims

The aim of the project is to identify, isolate and characterise novel microbial cross-linking enzymes (TGases) that possess the characteristics required to be used as a biocatalyst in new commercial sectors. The ideal candidate would address the problems currently posed by both microbial and mammalian TGases, discussed previously. They should possess a wide pH and temperature range whilst also exhibiting a broad substrate range making it a valuable biocatalyst for a number of industrial and commercial processes.

Initial studies searching both the nucleotide and protein databases with characterised TG sequences indicated a number of micro organisms that may contain potential TG enzymes (Makarova 1999). The project will therefore be initially concerned with identifying a range of bacteria that may harbour a TGase, identification techniques to highlight such organisms include evolutional relationships between known species, comparable physiology, and results of protein and nucleotide database homology searches. Identified organisms will be screened for TGase activity using a variety of assay formats to identify the presence of the enzyme. Assays based on a multiwell format will include incorporation of labelled primary amines (e.g. biotinylated cadaverine) into a lysine blocked glutamine rich protein (Slaughter 1992) and a crosslinking assay involving the incorporation of a biotinylated glutamine rich peptide into casein (Trigwell 2004). The use of site-directed irreversible inhibitors synthesised at Nottingham Trent University (TNTU) provides a key tool to confirm that the activity observed is TG mediated.

If suitable levels of expression can be obtained in a candidate organism(s), then the objective is to purify the enzyme on a laboratory scale using conventional methods. For characterisation of the candidate gene and to facilitate future large scale expression of the enzyme, the gene will be identified, isolated and ultimately subcloned into a suitable expression system and its ability to modify proteinaceous substrates assessed.

Chapter 2

Materials & Methods

2.1. Materials

2.1.1. General Chemicals

Unless otherwise stated reagents were purchased from Sigma-Aldrich Chemical Company (Dorset, UK). Other chemicals were obtained from the following suppliers:

Bristol University Peptide Synthesis Facility; biotinylated peptide TVQQEL

Molecular Probes, USA: Biotin cadaverine

A list of commonly used buffers may be found in Appendix I.

2.1.2. Radiochemicals

Amersham Pharmacia Biotech, UK; [1, 4 ^{14}C]-Putrescine (50μCi/ml)

2.1.3. Immunochemicals

Covalab; France; Anti-Isopeptide MAb

Dako Ltd., UK; Anti-Rabbit IgG HRP conjugate

NeoMarkers, USA; CUB7402 antiTG2 MAb

N-Zyme Biotech, Germany; Anti-mtg PAb

2.1.4. Protein Purification Media

Amersham Pharmacia Biotech, UK; SP Sepharose, Mono S, Q Sepharose, Gelatin Sepharose 4B, Phenyl Sepharose, Sephacryl S-200, EAH Sepharose 4B, HiTrap Selection Kit

Novagen, UK; Ni^{2+} resin

2.1.5. Western Blotting Reagents

Amersham Pharmacia Biotech, UK; Enhanced chemiluminescence detection kit

Bio-Rad Laboratories Ltd., UK; Protean II wet blotting system

2.1.6. Protein Reagents

Amersham Pharmacia Biotech, UK; Rainbow markers

Bio-Rad Laboratories Ltd., UK; Lowry Protein Assay Kit

Marvel dried milk powder was purchased from domestic supply outlets.

2.1.7. Molecular Biology Kits and Reagents

Reagents and kits for the purpose of molecular biology were purchased from Promega, UK, unless stated otherwise.

Invitrogen, UK; Platinum *pfx* DNA polymerase

Novagen, UK; 1 kb DNA ladder

2.1.8. Plasmid Vectors (see appendix II)

Novagen, UK; pET22(b)+, pET21(d)+

Promega, UK; pGEM-T Easy Vector System I

MoBiTec, Germany; pWH1520

Amersham Bioscience, UK; pGEXtTG, pGEX2T

2.1.9. Synthetic TGase Inhibitors

N-Benzyloxycarbonyl-L-phenylalanyl-6-dimethylsulfonium-5-oxo-L-norleucine bromide salt (R281) and 1,3-dimethyl-2-(2-oxopropylsulfanyl)-3H-1,3-diazol-1-ium-chloride (R283) (Griffin et al. 2006) were synthesized at TNTU. Purity was determined by mass spectrometry and nuclear magnetic resonance.

2.1.10. Other Consumables

All water was deionised using the Elgastat system 2 water purifier (Elga Ltd. UK) or Milli-Q water purifier (Millipore, UK).General laboratory consumables were obtained from the following suppliers:

Canberra-Packard, UK; Scintillation vials

Corning/Bibby-Sterilin, UK; Petri dishes, 6-well plates

Millipore, UK; Immobilon PVDF membrane

Nalgenenunc International, USA; Cryovials

Nunc. Ltd., USA; 96-well plates

Osmonics. UK; nitrocellulose membrane

Oxoid, UK; bacteriological growth medium

Sarstedt Ltd., UK; 0.5, 1.5- and 2-ml microcentrifuge tubes, 5 ml scintillation vial inserts, 15 and 50 ml sterile centrifuge tubes, 10 ml sterile pipettes, 1 ml and 200 µl pipette tips, 0.22µm filters

Whatman Ltd, Maidstone, UK; Whatman 1- and Whatman 3- MM chromatography paper

2.2. Equipment:

General laboratory equipment was purchased from the following suppliers:

Amersham Pharmacia Biotech, UK; AKTAPrime protein purification system, P1 peristaltic pump, FPLC protein purification system.

B & L Systems, The Netherlands; Atto-minigel protein electrophoresis system

Beckman, UK; Spectrophotometer Model DU-7, centrifuges MSE Centaur 2, MSE Microcentaur.

Bio-Rad Inc., UK; Mini-protean II electrophoresis module, DNA-submarine gel electrophoresis module and power pack

Canberra-Packard, UK; Tri-Carb 300 Scintillation counter

Corning, UK; pH meter 130

Edwards High Vacuum, UK; Freeze dryer (Modulyo System)

Grant Instruments, UK; water baths

LH Engineering, UK; 5L and 20L fermentation vessels and module, French press

MSE, UK; Chilspin refrigerated centrifuge

New Brunswick Scientific, USA; 10L Bioreactor system, orbital incubator

Tecan UK Ltd., UK; Spectrafluor 96-well ELISA plate reader and XFlour4 software

Techne, UK; Flexigene thermocycler

2.3. Micro-organisms

As part of the European funded Project HIPERMAX a large number of related Actinomycetes were screened for the production of extracellular TGase (work carried out by European partners).

2.3.1. Growth Media

Unless specified otherwise growth medium was purchased from Oxoid, UK.

2.3.2. Escherichia coli Strains

Strains and their genotypes may be found in Appendix IIIA.

2.3.3. Other Organisms

Fungi and Yeast

A complete list of fungi and yeasts may be found in Appendix IIIB.

Bacteria

A thorough list of bacteria may be found in Appendix IIIC.

2.4. Growth of Micro-organisms

2.4.1. Growth Media

Media was prepared using dH_2O and sterilised by autoclaving. A comprehensive list of media recipes may be found in Appendix IV. Where applicable freeze-dried cultures were reconstituted in 250 µl sterile H_2O prior to spreading onto nutrient rich agar and incubation at the required temperature. Micro-organisms were re-streaked a minimum of three times prior to use in either growth investigations or in the preparation of extracts for TGase assays.

2.4.2. Scaled up Growth (Bioreactor)

Where required, *Bacillus subtilis* NCIMB 10106, *Pseudomonas putida* NCIMB 9872 or *Streptomyces* sp. (*S. baldaccii* NRRL B3500, *S. paucisporogenes* ATCC 12596, *S. platensis* ATCC 13865 and *S. viridis* ATCC 15386) were grown in a stirred batch bioreactor (L.H. Engineering Co. Ltd., UK) with either a 5, 10 or 20 L vessel in place. Medium was autoclaved in situ at the desired pH. In each case frothing was prevented by the addition of organic antifoam 204.

The organism *B. subtilis* was inoculated into 20 ml Schaeffer's sporulation medium and incubated overnight at 37 °C with shaking at 225 rpm and used as the seed medium. The seed medium was diluted 5 in 500 ml of fresh medium in a 1 L conical flask. The organism was cultured for 8 hours until the absorbance at 600 nm $\approx$ 0.6 after which it was inoculated into 4.5 L Schaeffer's sporulation medium contained within a 5 L vessel. The culture was subsequently grown (with gentle aeration and 150 rpm) to the relevant stage with the rate of growth monitored by OD at λ=600 nm. After the final incubation cells were collected by centrifugation at 15, 000 x g for 5 min before proceeding as described in Section 2.4.7.1.

Pseudomonas putida was initially grown from a single colony in 20 ml LB medium overnight at 30 °C with shaking at 200 rpm. The starter culture was diluted 1 in 100 into two separate 1 L conical flasks containing 500 ml LB medium. The culture was grown to OD=0.6 at 600 nm before being added to 18 L sterile LB medium in a 20 L bioreactor vessel. *Ps. putida* growth (grown at 30 °C with agitation at 200 rpm) was monitored, with cells being collected by centrifugation (at 10, 000 rpm for 5 min) when the OD=0.6-0.7 at 600 nm after which the cell pellet was processed as described in Section 2.4.7.1.

For the purpose of TGase purifications actinomycete strains were inoculated, from a sporulated plate, into 25 ml GYM medium in a 250 ml conical flask, and incubated for 7 days at 30 °C with agitation (225 rpm). After which 10 ml of the resulting culture broth was transferred into 2 x 490 ml fresh medium (in 1 L shake flasks) and grown at 30 °C for 6 days or until maximal TGase activity was observed, determined by extracting a small amount of culture broth and assaying its TGase activity using the hydroxamate assay. After the final incubation cells were removed by filtration through Whatman No. 1 before protein purification.

2.4.3. Actinomycete and Fungal Growth

For the purpose of TGase identification Actinomycete and fungal strains were grown on a small scale. Spore inoculum was harvested from the agar medium, using a wire loop, and used to inoculate shake flask media (250 ml

in a 1 L conical flask). Cultures were grown at 30 °C, 225 rpm until TGase could be detected from the clarified culture supernatant, assayed as described under Section 2.5. Thermophilic / thermotolerant strains were cultivated at 45 °C.

2.4.4. Storage of Micro-organisms

Micro-organisms were stored on nutrient agar slants at 4 °C. For long term storage colonies/spores from the agar surface were collected and mixed with 60% sterile glycerol prior to storage at -80 °C.

2.4.5. Determination of Bacterial Growth Stage

Overnight cultures were diluted 1 in 100 in the appropriate media, incubated at the required temperature (see Appendix III) and the optical density measured at λ=600nm every hour. A growth curve was then plotted.

Due to the more complex life cycle of the actinomycetes and fungi, and as such their response to growth in liquid medium, an alternative approach was adopted to determine the growth characteristics of such organisms. A spore suspension, harvested from an agar plate, was inoculated into the appropriate medium and incubated at 30 °C with shaking at 225 rpm. At regular intervals (daily) 5 ml of culture was extracted and the cells collected by centrifugation at 8,000 rpm for 10 min. The wet cell weight was determined using a balance and used to establish a growth curve.

2.4.6. Effect of Inhibitors on Actinomycete Growth

To determine the effect of the irreversible inhibitor R281 (Figure 2.2) on the growth rate of S. baldaccii, the organism was grown in 25 ml GYM medium in a 250 ml conical flask (seed medium) for 6 days at 30°C, 200 rpm. The seed culture was diluted 1:10 with fresh medium in 6-well plates. Inhibitor R281 was added at varying concentrations (250 µM -1 mM) and the plate was incubated at 30°C. Fresh inhibitor was added daily to maintain effect. Cell differentiation was monitored visually after 6 days incubation.

2.4.7. Preparation of Micro-organisms for TGase Assays

2.4.7.1. Bacterial Extracts

Bacteria were cultured, in 500 ml nutrient medium in a 1 L conical flask, to specific growth phases (early-log, mid-log, late-log and stationary phase). Cells were collected by centrifugation at 8,000 rpm for 10 min at 4 °C. A portion of the medium was retained for TGase testing. Cells were completely resuspended in 5 ml lysis buffer (50 mM Tris HCl pH7.4, 1.5 mg/ml lysozyme, 1 mM DTT, 2 mM EDTA, 1 mM PMSF) prior to incubation at 4 °C for 2-4 hours with shaking at 70 rpm. After incubation cellular debris was removed by centrifugation at 10,000 rpm for 15 min at 4 °C and the supernatant removed. The cellular debris was resuspended in 5 ml Tris. HCl pH7.4. Preparations of culture medium, soluble and insoluble lysate were deemed ready for TGase activity determination.

Alternatively cells were collected by centrifugation, resuspended in 1/100 volume of lysis buffer and lysed by passing through a French press at a pressure of 6,000-9,000 PSI. Lysis was monitored by plating out 20 µl diluted to 500 µl sterile H_2O on nutrient rich medium and incubating overnight.

2.4.7.2. Actinomycete and Fungal Extracts

Microbial TGase is known to be secreted into the surrounding medium (Ando 1989). As such, to generate actinomycete and fungal extracts, organisms were initially inoculated into 20 ml nutrient rich medium and grown at the required temperature (see Appendix III) for 6 days to generate the starter culture. The starter culture was diluted 10 in 100 in fresh medium (250 ml medium in 1 L conical flask) and incubated for the desired amount of time after which cells and debris were removed by centrifugation at 8,000 rpm for 5 min, alternatively cells were removed by filtration through Whatman No. 1 to generate the clarified culture supernatant. The clarified supernatant generated was assayed for TGase activity by the either the hydroxamate or biotin cadaverine assays in triplicate on at least three separate occasions.

2.5. Measurement of TGase Activity

2.5.1. Hydroxamate Assay

For the routine assay of TGase, the hydroxamate assay was used according to the method of Folk and Chung, 1985 (Folk and Chung 1985)(Figure 2.1). A 25 µl portion of a sample was mixed with 75 µl of Reagent A in a 96-well microtitre plate well. After incubation at 37 °C for 10 min, 75 µl of Reagent B was added to terminate the reaction and form an iron complex (Table 2.1). Thereafter, the absorbance of the complex was measured at 492nm in a microtitre plate compatible spectrophotometer (Tecan UK Ltd.). The colour development was read immediately after quenching.

Figure 2.1; Schematic Representation of the Hydroxamate TGase Assay
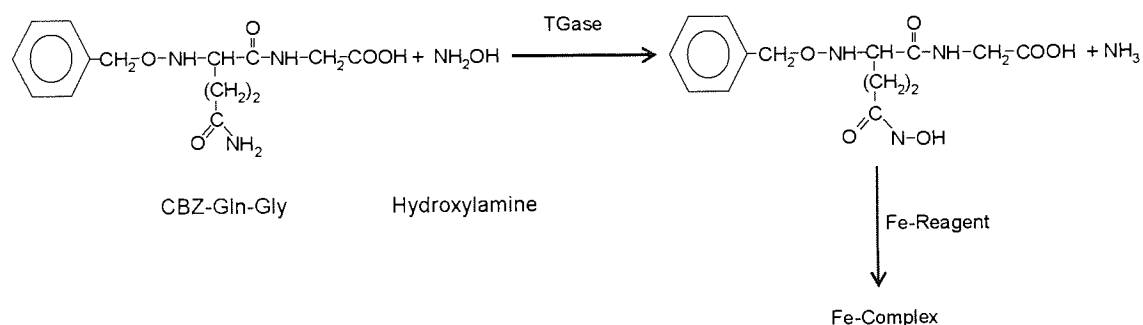
Enzyme solutions, inactivated by heating, were used as negative controls in the same reaction procedure, and their absorbance at 492 nm were subtracted from that of the active enzyme solution. A calibration curve was prepared using the same procedure except that γ-mono-hydroxamic acid L-glutamate was used instead of the enzyme solution.

Table 2.1; Hydroxamate Reagent Components

Reagent A	Reagent B
0.2 M Sodium acetate pH6.0	1 volume 3 N HCl
0.1 M Hydroxylamine	1 volume 12% TCA
0.03 M CBZ-Gln-Gly	1 volume 5% $FeCl_3.6H_2O$
	(Dissolved in 0.1 N HCl)

One unit of the enzyme activity was defined as the amount of enzyme catalyzing the formation of one micromole of hydroxamic acid per minute under the described reaction conditions. For the purpose of enzyme purifications activity is defined as the absorbance at 492 nm.

2.5.2. Biotin-Cadaverine Incorporation into N,N`-Dimethylcasein

Transglutaminase activity was determined by the incorporation of biotinylated cadaverine as described previously by Slaughter et al. 1992 (Slaughter et al. 1992) with the following modifications: 96 well plates were coated with 100 µl of 10 mg/ml N,N'-dimethylcasein in 10 mM Tris-HCl overnight at 4 °C, before being washed twice with TBS-Tween 20 and then once with TBS. 90 µl of a reaction mixture comprising of 0.132 mM BTC in 50 mM Tris-HCl (reaction buffer) was applied to each well before adding 10 µl of sample to initiate the reaction. Samples were assayed in triplicate. The reaction was allowed to proceed for 1 hour at 37 °C. Plates were washed as before and then blocked with 3% BSA in PBS (blocking buffer) for 1 hour at 37 °C. Incorporated BTC was resolved with a 1:5000 dilution of EXAP in blocking buffer for 1 hour at 37 °C. Plates were washed again, prior to development and preincubated for 5 min in 0.005 M phosphate citrate buffer containing 0.014% (v/v) H_2O_2. This solution was removed and replaced with 100 µl of the same buffer containing 75 µg/ml TMB (developing solution). The reaction was allowed to proceed for 15 min before being quenched by the addition of 50 µl of 2.5 M H_2SO_4. The resulting colour was read at 450 nm using a microtitre plate compatible spectrophotometer (Tecan UK Ltd.). Enzyme activity is defined as the absorbance at 450 nm.

2.5.3. Incorporation of Synthetic Peptide into Casein

The assay makes use of a glutamine rich synthetic peptide (TVQQEL) that is biotinylated at the N-terminus (Trigwell et al. 2004). Plates were pre-incubated overnight at 20 °C with 250 µl casein at a concentration of 1 mg/ml in 50 mM sodium carbonate buffer, pH 9.8. Plates were emptied by brisk inversion and washed three times with dH_2O before being blocked with 250 µl 0.1% BSA in sodium carbonate buffer, pH 9.8 (blocking buffer) at 37 °C for

1 hour. Wells were again washed and 150 µl 5 µg/ml biotin-TVQQEL in 100 mM Tris-HCl pH7.4 (reaction buffer) was applied to each well. 50 µl of sample was added to initiate the reaction. After incubation for 1 hour at 37 °C wells were emptied and washed as before. Incorporated biotinylated peptide was detected by the addition of 200 µl probing solution (EXAP diluted 1:10,000 in 1% (w/v) BSA/100 mM Tris-HCl) to each well; plates were incubated for 1 hour at 37 °C. Following washing, EXAP binding was detected by adding 200 µl developing solution (75 µg/ml TMB in 0.005 M phosphate citrate buffer containing 0.014% (v/v) H_2O_2) and allowing the reaction to proceed for 15 minutes before being stopped by the addition of 50 µl 5 M H_2SO_4. Absorbance was measured at a wavelength of 450 nm using a microtitre plate compatible spectrophotometer (Tecan, UK Ltd.). Enzyme activity is defined as the absorbance at 450 nm.

2.5.4. Incorporation of [^{14}C]-Putrescine into N,N'-Dimethylcasein

This assay relies on the ability of the enzyme to catalyse the incorporation of a radioactively labelled primary amine into a protein acceptor substrate (Lorand et al. 1972). At 30 second intervals, 45 µl of sample was added to 45 µl reaction mix containing 10 µl 50 mM Tris-HCl pH 7.4, 10 µl 12 mM [1,4-^{14}C]-putrescine (Amersham Pharmacia; specific activity 3.97 mCi/mmol), 20 µl 25mg/ml N,N'-dimethylcasein in 50 mM Tris and either 5 µl 50 mM Tris-HCl pH 7.4 or 5 µl 200 mM EDTA pre-equilibrated at 37 °C in a water bath. 15 min after initiation of the reaction 10 µl aliquots from each sample were removed, in triplicate, and spotted onto 10 mm^2 of 3MM filter paper (Whatman, UK) at 30 second intervals. Squares were then placed in 10% (w/v) ice cold TCA for 10 min, washed three times for 5 min in 5% (v/v) TCA, once in acetone: ethanol (1:1, v/v) for 5 min and once in acetone for 5 min. A triplicate set of filter papers containing no sample was also washed as a reaction blank. Also a triplicate set of filter papers were spotted with the reaction mix and allowed to air dry, so as to measure counting efficiency. Filter papers were then air dried for 15 min. Once dry, filter papers were placed in scintillation tubes, to which 2 ml Optiphase High Safe liquid scintillation fluid was added, and ^{14}C counts were obtained in a Packard

Liquid Scintillator (Packard Biosciences, UK). Enzyme activity is defined in counts per minute.

2.5.5. Influence of Inhibitors on TGase Activity

Synthetic chemicals created on site in the Department of Chemistry, Nottingham Trent University, R281 and R283 (Figure 2.2), were designed to irreversibly bind to the active site of transglutaminase. Both are site-directed irreversible inhibitors of transglutaminase that act by covalent modification of the cysteine residue present within the active site. Concentrations used ranged from 100 μM - 1 mM and were pre-incubated with the sample to be tested for 1 hour at 37 °C.

Figure 2.2; Chemical Structure of Synthetic Inhibitors R281 and R283

R281 (N-Benzyloxycarbonyl-L-phenylalanyl-6-dimethylsulfonium-5-oxo-L- norleucine bromide salt);
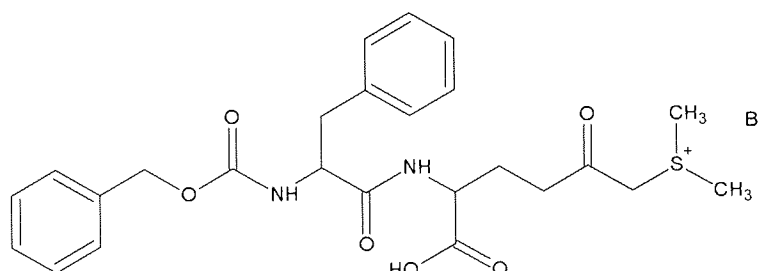
R283 (1, 3-dimethyl-2-(2-oxopropylsulfanyl)-3H-1,3-diazol- 1-ium-chloride);
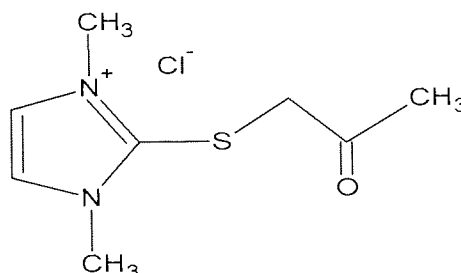
2.6. Production of TGase Affinity Purification Resin

The TGase active site directed suicide inhibitor R281 (Figure 2.2) was coupled to EAH Sepharose 4B (Amersham Biosciences, UK) by reaction of the free carboxyl on R281 with EDC and subsequent conjugation to the free amine of the EAH Sepharose 4B.

EAH Sepharose 4B was washed with three cycles of alternating low pH (0.1 M sodium acetate, 0.5 M NaCl pH 4) and high pH (0.1 M Tris HCl, 0.5 M NaCl pH 8.3) buffers, with a final wash of distilled water. R281 was dissolved in 20ml distilled water pH 6 at a concentration of 10 mM and 10 ml of EAH Sepharose 4B (settled resin volume) was added. To this mixture, solid EDC was added to a final concentration of 0.1 M and the mixture was rotated for 24 h at 4 °C. The pH of the reaction was monitored during the first 2 hours and adjusted to pH 6 if necessary with 0.1 M NaOH. The resin was then washed as above with three cycles of alternating high and low pH buffer and a final water wash. Resin was stored at 4 °C until required.

2.6.1. Affinity Isolation of TGase using R281 Coupled Resin

The cell free lysate of the micro-organism to be investigated was prepared as described in Section 2.4.7.1. The lysate was mixed with 1 ml (settled volume) of R281-Sepharose resin. $CaCl_2$ and DTT were added to a final concentration of 5 mM each and the mixture incubated overnight at 4 °C with shaking at 50 rpm. The resin was packed into a column and installed into the FPLC machine (Pharmacia FPLC). Protein was eluted with a 20 ml linear gradient of 0-5.5 M NaCl in PBS pH 7.4, collecting 1 ml fractions. The resin was then washed with 5 ml of PBS pH 7.4. Resin was recovered from the column, 30 ng/ml trypsin was added and the resulting suspension was incubated at 37 °C overnight so as to digest covalently attached protein off the resin. Fractions generated were analysed by SDS-PAGE to determine the protein elution profile.

2.7. Preparation of Guinea Pig Liver and Microbial TGase

2.7.1. GPL-TGase Purification

Tissue transglutaminase was purified using the protocol described previously (LeBlanc et al. 1999) with a few modifications;

2.7.2. Liver Homogenisation

Five guinea pigs were sacrificed by cervical dislocation and extracted livers (200-250g) were homogenised to a 50% (w/v) homogenate first using a bench-top blender in 500 ml 5 mM Tris-HCl, 2 mM EDTA, pH 7.5, 0.25 M sucrose, 5 mM benzamidine, 1 mM PMSF (homogenizing buffer), and then using an homogeniser. Nuclei and large debris were removed by centrifugation at 30,000 x g for 1 hour at °C. The supernatant was subjected to further centrifugation at 100,000 x g for 1 hour at 4 °C to remove cell membranes and other cell debris. The clarified supernatant was filtered through Whatman No.1 filter paper (Whatman, UK).

2.7.3. Anion Exchange Chromatography

Clarified homogenate was loaded onto 300 ml of Q-sepharose Fast Flow (Amersham Pharmacia, UK) at a rate of 10 ml/min and the column washed with Buffer 1 until an eluant A_{280} ≤ 0.5 Au was reached. Protein was eluted with a linear gradient of NaCl from 0 to 0.720 M over 190 min at a flow rate of 10 ml/min collecting 1 min fractions. TGase activity and protein content was determined by the hydroxamate activity assay and the Lowry protein assay, respectively, as described in Sections 2.5.1 and 2.9.1. TGase rich fractions were pooled.

2.7.4. Size Exclusion Chromatography

TGase fractions collected from anion exchange chromatography were precipitated by gradual addition of $(NH_4)_2SO_4$ to 80% (w/v) on ice with gentle stirring for 45 min. Precipitate was removed by centrifugation at 30,000g for 1 hour at 4°C. Pellets were reconstituted in 10ml of 50mM Tris-Acetate, 1mM

EDTA, 0.16M KCl, pH 6.0 before being loaded onto a 500 ml Sephacryl S-200 column (Bio-Gel A-0.5m, fine mesh), equilibrated in the same buffer, at a flow rate of 2ml/min. Dextran blue (10mg) was added to the resuspended precipitate so as to track protein progression. Protein was recovered by continued flow of 50 mM Tris-acetate buffer, pH 6.0, containing 1 mM EDTA and 0.16M KCl. Fractions (5ml) were collected from when the dextran blue tracking dye reached the bottom of the column until TGase activity could be detected. Fractions were assayed for TGase activity, by hydroxamate, and protein content (Lowry assay) with highly active fractions being pooled.

2.7.5. GTP-Agarose Gel Filtration

Size exclusion chromatography fractions were applied onto a 5 ml bed volume GTP agarose column (Sigma, UK) equilibrated in 50 mM Tris-Acetate, 1 mM EDTA, 1 mM DTT, pH 7.5 and run through at 2 ml/min using a peristaltic pump. The flow-through was then reloaded to ensure saturation of the binding sites. The column was then washed with 50 ml of the same buffer before elution with 5 mM GTP in 50 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 7.5 Resulting 1 ml fractions were assayed once again for protein content and TG2 activity using UV absorbance and biotin cadaverine activity assay respectively. Purity was verified by SDS-PAGE and compared to that of commercially available gpl TGase (Sigma Aldrich, UK).

2.7.6. Dialysis

Following gel filtration GTP contamination was removed by overnight dialysis with three buffer changes at 4 °C in PBS/2 mM EDTA pH 7.4 using dialysis tubing with a 5kDa MW cut-off (Perbio Science, UK).

2.8. mTGase Purification (Commercial Source)

Microbial TGase was purchased from Ajinomoto Co. (Ajinomoto European Sales, Germany) as a 1 % preparation combined with maltodextrin. Single step purification was undertaken to remove extraneous maltodextrin.

50 g of TGase preparation (500 mg TGase protein) was dissolved in 20 mM sodium phosphate and its pH adjusted to 6, by the addition of NaOH, prior to filtration through Whatman No. 1 filter paper. The resulting solution was loaded onto an SP Sepharose column (100 ml resin), equilibrated in sodium phosphate pH 6, at a rate of 5 ml/min. After washing with 5 column volumes of 20 mM sodium phosphate pH 6, protein was eluted with a linear gradient of 0-0.5 M NaCl over 50 min, at a rate of 5 ml/min, collecting 5 ml fractions. Fractions were assayed for protein content (Lowry assay), TGase activity (hydroxamate assay) and assessed by SDS-PAGE. Fractions with high activity were pooled lyophilized and stored at -20°C until required.

2.8.1. Purification of Streptomyces TGase

TGases were purified according to the method described by Duran (Duran et al. 1998) with the following modifications:

2.8.1.1. Cation Exchange Chromatography

Cells were removed by centrifugation at 8,000 rpm for 15 min and the supernatant filtered through Whatman No. 1 filter paper. The pH of the filtrate was adjusted to pH 6 by the addition of NaOH and loaded onto an SP Sepharose column (100 ml resin) at a rate of 5 ml/min. After washing with 5 column volumes of 20 mM sodium phosphate pH 6, protein was eluted, at a rate of 5 ml/min, with a linear gradient of 0-0.5 M NaCl over 92 min collecting 5 ml fractions at 4°C. Fractions were assayed for activity (by hydroxamate), protein content (Lowry assay) and assessed by SDS-PAGE. Fractions with high specific activity were pooled.

2.8.1.2. Hydrophobic Interaction Chromatography

NaCl was added to the active fractions to a final concentration of 3M prior to loading onto a phenyl Sepharose column (25 ml resin) equilibrated with (running buffer (20 mM Sodium phosphate, 3 M NaCl pH6) at room temperature. Protein was eluted with a linear salt gradient from 3-0 M NaCl in 20 mM sodium phosphate pH 6 at a rate of 2 ml/min. Fractions, collected at 1

min intervals (2ml), were assayed for activity (by hydroxamate or biotin cadaverine assays), protein content and assessed by SDS-PAGE. Fractions with high specific activity were pooled and exhaustively dialysed against 20 mM sodium phosphate pH 6 at 4 °C.

2.8.1.3. Gelatin Sepharose Chromatography

Dialysed protein, possessing TGase activity, was loaded onto Gelatin Sepharose 4B column (25ml resin) (Amersham Pharmacia, UK) equilibrated in 20 mM sodium phosphate pH6 at a rate of 2 ml/min. After washing, with 5 column volumes of equilibration buffer, protein was eluted with a linear gradient of 0-0.4 M NaCl in 20 mM sodium phosphate buffer pH6 at a rate of 2 ml/min collecting 2 ml fractions. Active fractions were assayed for TGase activity (by hydroxamate), protein content (UV absorbance) and enzyme purity analysed by SDS-PAGE as described in Section 2.9.2.

2.8.1.4. Cation Exchange Chromatography (MonoS)

Protein was applied to a 1 ml MonoS column equilibrated in 20mM sodium phosphate pH 6, at a rate of 0.5 ml/min at 4 °C. After washing with 5 column volumes of 20 mM sodium phosphate pH 6, protein was eluted with a linear gradient of 0-0.4 M NaCl in 20 mM sodium phosphate pH 6 collecting 1 ml fractions. Fractions were assayed for activity (hydroxamate assay), protein content (UV absorbance) and assessed by SDS-PAGE. Fractions with high specific activity and high purity were pooled.

2.9. Analysis of Protein
2.9.1. Protein Estimation

The Lowry protein assay (Lowry et al. 1952) was carried out using the Bio-Rad kit (Bio-Rad, UK). 25 µl of reagent A and 200 µl of reagent B were added sequentially to 5 µl of sample or BSA standards ranging from 0.15-10 mg/ml in a 96-well plate format. The plate was incubated for 15 minutes at room temperature and the absorbance was read at 750 nm using a spectrophotometer (Tecan, UK Ltd.).

Protein concentration was evaluated by comparing sample absorbance at 280 nm to known standards using UV-spectrophotometry. This technique was used, as required, to evaluate protein samples from gpl-TGase fractions containing high amounts of DTT and used for microbial TGase to detect low amounts of enzyme.

2.9.2. SDS-PAGE

The method employed was a modification of that described by Laemmli for use with vertical slab gel apparatus (Laemmli 1970). Gels were cast using the Mini Protean 3 Electrophoresis System (Bio-Rad, UK) and consisted of a 3% (w/v) polyacrylamide stacking gel with a 10-12% (w/v) resolving gel enabling effective separation of proteins of differing molecular weights. Resolving gel recipes for different concentrations of acrylamide are listed in Table 2.2. Resolving gels were cast according to the manufacturer's protocol.

Table 2.2; Polyacrylamide Gel Recipe

Stock Solutions	Final Acrylamide Concentration (%)	
	10	12
30% Acrylamide/0.8% bisacrylamide	5.00 ml	6.00 ml
4x Tris. HCl/SDS pH8.8	3.75 ml	3.75 ml
H_2O	6.25 ml	5.25 ml
10% Ammonium Persulphate	0.05 ml	0.05 ml
TEMED	0.01 ml	0.01 ml

The resolving gel was overlaid with isopropanol and the gel was allowed to polymerise for 45 min at room temperature. The upper surface of the polymerised resolving gel was washed with dH_2O to remove residual isopropanol, and dried using filter paper. The stacking gel, comprising; 0.65 ml acrylamide solution, 1.25 ml Tris/SDS pH 6.8, 3.05 ml of dH_2O 25 µl 10% (w/v) ammonium persulphate and 5 µl of TEMED was applied between the glass plates and a comb was inserted to form the sample wells. Polymerisation was allowed to proceed for 45 min at room temperature.

The sample comb was removed and wells were washed and filled with Tris-glycine running buffer pH 8.5. Samples to be loaded were normalised against protein concentration, with 15-20 μg protein being combined with the appropriate volume of 5x Laemmli loading buffer and boiled for 5 min prior to being loaded into the sample wells. Electrophoresis was carried out for approximately 1.5 hours at 150 V until the bromophenol blue marker dye had reached the bottom of the gel. For visualisation gels were stained with Coomassie brilliant blue.

2.9.3. Western Blotting of SDS-PAGE Separated Proteins

Electrophoretic transfer of SDS-PAGE separated proteins to nitrocellulose was performed as described by Towbin (Towbin et al. 1979) using the Protean II Cell System (Biorad, UK). First gels were equilibrated in transfer Buffer for 15 min, before being assembled as follows; a scanting pad was layered onto the cathode plate of the assembly cassette followed by presoaked filter paper and the washed polyacrylamide gel. Nitrocellulose membrane (Osmonics, UK) was layered on top of the gel followed by additional filter paper and the final scanting pad. Before the cassette was closed air bubbles were removed. Western blotting was carried out for 1.5 hours at 160 mA in pre-chilled transfer buffer.

Upon completion, the membrane was recovered and transfer of protein was verified by staining the membrane with Ponceau Red solution and the gel with Coomassie Blue. Ponceau Red was removed by washing in dH$_2$O. Membranes were initially blocked by incubation with 5% (w/v) fat-free dried milk powder in PBS, pH 7.4, and 0.05% (v/v) Tween20 for 2 hours at room temperature, before being probed with the appropriate primary antibody, in the same buffer, overnight at 4 °C with gentle agitation.

2.9.3.1. Immunoprobing of Western blots

Following primary antibody binding the membrane was washed three times in PBS, 0.5% (v/v) Tween20, for a total of 15 min and incubated with a species-specific secondary HRP-conjugated antibody, in blocking buffer, for 2 hours

at room temperature with agitation, after which a final three washes were carried out to remove the antibody background.

The HRP component of the secondary antibody was exposed for 60 seconds to Enhanced ChemiLuminescence substrate (ECL, Amersham Bioscience, UK). Excess substrate was removed and the membrane covered with cling film. The resulting light emission was detected by exposure to Kodak X-Omat chemiluminescence detection film (Sigma, UK) for 1-20 min, depending on antibody dilution and signal intensity. The film was developed using 20% (v/v) LX-24 developer (Sigma, UK), fixed in 20% (v/v) FX-40 fixer (Sigma, UK), and rinsed in tap water before being air dried.

2.9.4. Analysis of Amino Acid Content

Enzymes were concentrated to 10 mg/ml and hydrolysed in 6 M HCl at 100 °C prior to derivatisation using the EZ: Faast amino acid analysis kit (Phenomenex, UK), according to the manufacturer's instructions. Derivatised samples were analysed using Gas Chromatography and their amino acid content evaluated.

2.9.4.1. Cyanogen Bromide Digestion

Highly purified TGases (1mg) were dissolved in 500 µl of 70% TFA and treated with 1-2 crystals (50 M excess) of cyanogen bromide at room temperature for 24 hours. Digested peptides were dried at 80 °C and dissolved in 200 µl of 0.1% TFA.

2.9.4.2. Reversed-Phase HPLC

Cyanogen bromide digested proteins were first resuspended in water before being resolved using a HiPore RP-318 C18 25cm x 0.46cm HPLC column (BioRad, UK) equilibrated in water/0.1% trifluoroacetic acid at a flow rate of 1 ml/min. Peptides were eluted with acetonitrile/0.1% trifluoroacetic acid at room temperature using the following gradient:

Time	Water/ 0.1% TFA	Acetonitrile/ 0.1% TFA
0	99.5%	0.5%
30	70%	30%
45	0%	100%
46	0%	100%
47	99.5%	0.5%
55	99.5%	0.5%

Absorbance of the eluant was monitored at 230 nm and 280 nm with 1 ml fractions being collected. Fractions containing peaks of interest were reduced in volume to near-dryness and sent for N-terminal sequencing.

2.9.5. Peptide Mass Mapping

Proteins for analysis were purified as described and desalted before being sent to the University of Helsinki (Finland) for peptide mass mapping using a BiflexTM matrix assisted laser desorption ionization/time-of-flight (MALDITOF) mass spectrometer (Bruker-Franzen Analytik, Bremen, Germany).

2.10. TGase Cross-Linking of Model Substrates

To determine whether the different Streptomyces TGases demonstrated any differences in substrate specificity with respect to cross-linking, they were used to cross-link the model substrates gelatin, casein and BSA. To 1ml of either sodium caseinate, porcine gelatin (75 Bloom) or BSA solution (15 mg/ml in PBS pH7.4) was added a known amount of TGase. The solutions were incubated overnight at 37 °C with shaking. Aliquots were removed at various time points for analysis of polymer generation by SDS-PAGE and isopeptide bond detection and quantitation by Western blotting with the remainder being subjected to proteolytic digestion for direct $\varepsilon(\gamma$-glutamyl)lysine determination.

2.10.1. Proteolytic Digestion of Cross-Linked Substrates

In order to release the acid and alkali-labile ε(γ-glutamyl)lysine isopeptide cross-link from TGase-treated protein, an exhaustive program of proteolytic digestion utilising a number peptidases was required. The following method, similar to that described by Griffin and Wilson, 1984(Griffin and Wilson 1984), was undertaken with modifications as detailed below.

Briefly, samples of protein (approximately 15mg) were precipitated by the addition of 100% TCA to a final concentration of 10% and incubated on ice for 10 min. Protein was collected by centrifugation at 13,000 rpm for 5 min at room temperature. The supernatant was discarded and the pellet thoroughly resuspended in 250 µl of 10% TCA. Protein was collected by centrifugation as before and washed three times with 250 µl of diethyl ether/ethanol (1:1), ensuring complete dispersion of the pellet. The pellet was then washed in the same way three times with diethyl ether, and dried at room temperature for 30 min. The pellet was rehydrated in 50 µl of water and sonicated briefly to aid dispersion. To this was added 950 µl of 0.1 M NH_4HCO_3 pH8.0 and 1 crystal of thymol to prevent bacterial growth during subsequent incubations. Proteolytic treatments were then performed in the following way: 10 µl of 0.5 M $CaCl_2$ and 10 µl of collagenase (Clostridiopeptidase A) (10 mg/ml in 0.1 M NH_4HCO_3 were added and incubated at 32 °C for 16 hours with shaking; 10 µl of subtilisin carlsberg (10 mg/ml in 0.1 M NH_4HCO_3) was added and incubated at 32 °C for 16 hours with shaking. The subtilisin digestion was repeated twice more. A protein assay was performed at this point to give an indication of protein concentration for subsequent calculations. 10 µl of pronase (15 mg/ml in 0.1 M NH_4HCO_3) was added and incubated at 32 °C for 16 hours with shaking. This digestion was repeated once. Proteases were inactivated by heating to 100 °C for 15 min. Magnesium chloride was added to a final concentration of 5mM. Leucine aminopeptidase was activated by mixing the following components and incubating for 3 hours at 37 °C: 10 µl $MnCl_2$ (50 mM), 90 µl Tris HCl pH8.0 (10 mM) and 100 µl leucine aminopeptidase (22.75U). Prolidase was activated by mixing the following and incubating at 37 °C for 3 hours: 20 µl of $MnCl_2$ (50 mM), 80 µl of Tris HCl pH8.0 (10 mM), 80 µl of distilled water, 20 µl of prolidase (38.6U). 90 µl of the

activated leucine aminopeptidase and 75 µl of activated prolidase was added and incubated at 37 °C for 16 hours with shaking. The leucine aminopeptidase and prolidase digestions were repeated once more. The pH of the sample was adjusted to between 6.75 and 7.0 with HCl and 10 µl of carboxypeptidase Y (20 mg/ml) was added and incubated at 30 °C for 16 hours with shaking. Samples were freeze dried and stored at -20 °C prior to analysis.

2.10.2. Quantitation of ε (γ-glutamyl) lysine by Cation Exchange Chromatography

Lyophilised samples were resuspended in 0.1 M HCl and sonicated for 2 min to aid dispersion. An aliquot (90 µl) was mixed with 110 µl of loading buffer (0.2 M lithium citrate, 0.1% phenol, 0.2% thiodiglycol pH 2.2) and loaded onto an UltroPac 8 resin column 0.46cm x 20cm using a Biochrom 30 amino acid analyser running a lithium citrate buffer system. The program used was a modification of the standard separation and was optimised for ε (γ-glutamyl)lysine separation using a series of lithium citrate buffers of different molarity and pH:

Buffer 1

0.2M lithium citrate, 1.5% propanol, 0.1% phenol, 0.2% thiodiglycol, pH 2.80

Buffer 2

0.3M lithium citrate, 1.5% propanol, 0.1% phenol, 0.2% thiodiglycol, pH 3.00

Buffer 3

0.5M lithium citrate, 0.1% phenol, 0.2% thiodiglycol, pH 3.30

Buffer 4

0.9M lithium citrate, 0.1% phenol, 0.2% thiodiglycol, pH 3.50

Buffer 5

1.65M lithium citrate, 1.5% propanol, 0.1% phenol, 0.2% thiodiglycol, pH 3.55

Buffer 6

0.3M lithium hydroxide

Derivatisation was performed post column using o-phthaldialdehyde (0.8 M boric acid, 0.78 M KOH, 600 mg/ml o-phthaldialdehyde, 0.5% (v/v) methanol, 0.75% (v/v) 2-mercaptoethanol, 0.35% (v/v) Brij 30) and the fluorescence was measured at 340 nm excitation and 450 nm emission using a fluorescence detector (Cecil CE-4500). Dipeptide was determined by standard addition of known amounts of ε(γ-glutamyl)lysine to the sample and comparing peak areas. An example of the resulting chromatogram may be seen in Figure 2.3.

Figure 2.3; Chromatogram Depicting the Separation of ε,γ(glutamyl)lysine cross-link (1 nmol) from a Mixture of Standard Amino Acids (1nmol each) using Cation Exchange Chromatography.
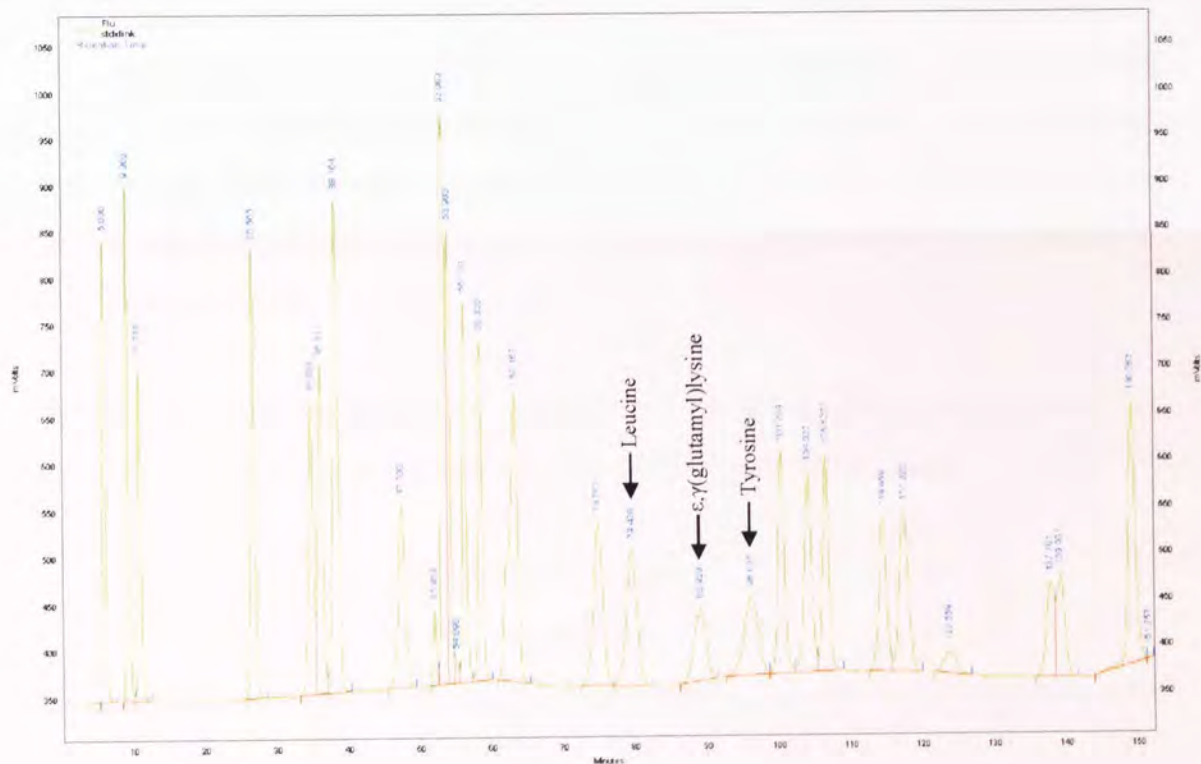
2.11. Molecular Biology Techniques

2.11.1. Bioinformatics

BLASTp, BLASTn and PSI-BLAST were accessed by the NCBI website (www.ncbi.nig.gov/BLAST). PSI-BLAST program identifies conserved domains using the Conserved Domains (CD) database compiled by NCBI.

DNASTAR (Lasergene) programs Primerselect and Megalign were used to calculate oligonucleotide characteristics and complete multiple alignments, respectively. The Megalign program uses the CLUSTAL method for alignments.

2.11.2. Genomic DNA Isolation

Genomic DNA was isolated from micro-organisms following overnight growth in nutrient rich medium, utilising the GenElute™ Bacterial Genomic DNA Miniprep Kit (Sigma-Aldrich, UK) according to the manufacturer's instructions. DNA isolation from *Streptomyces sp.* and related genera was carried out according to the CTAB/NaCl protocol as described by Kieser et al, 2000 (Kieser 2000).

Prior to downstream applications extracted DNA was analysed by both UV spectrophotometry and agarose gel electrophoresis as detailed.

2.11.3. Determination of DNA Concentration

Samples were diluted 1/100 in dH_2O, loaded into quartz micro-cuvettes, and readings were recorded at 260 and 280 nm against a dH_2O blank. DNA solutions with an A260nm /A280nm ratio ≥ 1.6 were deemed free of contaminating substances.

2.11.4. Agarose Gel Electrophoresis

Analysis of DNA was achieved by agarose gel electrophoresis using 0.75% to 2% (w/v) agarose gels. Gels were prepared by heating the desired amount of agarose in 1xTAE buffer until it had dissolved. Upon cooling ethidium bromide was added to final concentration of 0.5 mg/ml and the gel was cast using an electrophoresis tray (Bio-Rad, UK) and allowed to set.

DNA solutions to be analysed were supplemented with 10x DNA loading buffer, mixed and applied to the sample wells. Electrophoresis was performed at 90 V for 90 minutes in 1x TAE running buffer or until the desired separation was achieved. DNA samples were visualised using a transilluminator and compared to known molecular weight standards.

2.11.5. Oligonucleotides

All oligonucleotides were synthesized (desalted) by Sigma Genosys (UK). Oligonucleotides used in this work are detailed in Table 2.3.

2.11.6. Polymerase Chain Reaction

PCR reactions were set up in thin walled tubes. A typical reaction mix comprised of 1x Reaction Buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTP's, 100 ng of each primer, 0.1-10 ng of template DNA, 2.5 U thermophilic DNA polymerase and sdH_2O to give a final volume of 50 µl. Following cycling, amplified products were analysed by gel electrophoresis.

Cycling conditions proceeded as follows:

$$94°C \text{ for 4min}$$
$$\left.\begin{array}{l} 94°C \text{ for 1min} \\ 57°C \text{ for 1min} \\ 75°C \text{ for 1min} \end{array}\right\} \text{ 35 cycles}$$
$$75°C \text{ for 5min}$$

For degenerate PCR $MgCl_2$ concentration was increased to 3.5 mM to aid primer annealing, the annealing temperature was reduced to 50 °C for 45s and the extension time was reduced to 45s. Following cycling, amplified products were analysed by gel electrophoresis.

Table 2.3; PCR Oligonucleotides

Where indicated A= adenine, T=thymine, G=guanine, C=cytosine, N=A/C/G/T, Y=C/T, M= A/C, H= A/T/C, R= A/G, W= A/T

Primer Name	Sequence (5´-3´)	Function
Rlp_1	GCTGGCCAAAATTATTGTATCAGGACAA	*B. subtilis* forward primer with *Msc*I site for pET22b(+)
Rlp_2	TTGCCTCGAGGCGGACGATGCGGAAAA	*B. subtilis* reverse primer with *Xho*I site for pET22b(+) and pET21d(+)
Rlp_3	GGATCCATTATTGTATCAGGAC	*B. subtilis* forward primer with *Bam*HI site for pET21d(+)
Rlp_4	GAATTCTTAGCGGACGATGCGG	*B. subtilis* reverse primer with *Eco*RI site for pGEX1λT
Rlp_5	TTTTACTAGTATTATTGTATCAGGACAA	*B. subtilis* forward primer with *Spe*I site for pWH1520
Rlp_6	TTTTGGATCCTTAGCGGACGATGCGG	*B. subtilis* reverse primer with *Bam*HI site for pWH1520
Ps_1	TTGCTGGCCAAACTGTCCATTCGCCA	*Ps. putida* TGase-like primer with MscI site
Ps_2	TTGCTCGATCACTGGCGATGCACG	*Ps. putida* TGase-like primer wth Xhol site
Degen_F1	GTNAAYAAYTAYATHMGN	Degenerate primer to: aa 123-135
Degen_F2	CARCARATGACNGARGAR	aa 142-150
Degen_R1	NCKNARNGCNSWRTARAA	aa 262-273
Degen_R2	GTNAAYTTYGAYTAYGGN	aa 344-354
Bal_F	GGATCCGGCTCTTACCGACGACAGG	*S. baldaccii* forward primer with BamHI site
Bal_R	GAATTCTTACGGCCAGCCCTG	*S. baldaccii* reverse primer with EcoRI site

2.11.7. Extraction of DNA from Agarose Gels

DNA excised from agarose gels was purified using Wizard SV PCR clean up kit (Promega, UK) according to the manufacturer's protocol. 5 µl of the resulting DNA solution was checked by electrophoresis as detailed in Section 2.11.4.

2.11.8. Restriction Enzyme Digestion of DNA

Digestion of DNA with restriction endonucleases was carried out according to the manufacturer's guidelines. A typical reaction would contain 2-4 µg DNA, 1 x reaction buffer and 5-10 U enzyme, usually to a final volume of 30 µl. This was incubated at 37 °C for 2-3 hours prior to analysis by electrophoresis and purification of digested DNA as detailed in Section 2.11.4.

2.11.9. DNA Ligation

Following extraction and purification of digested DNA, ligations were set up. Each ligation contained 1x ligase buffer, 200 ng insert, 600 ng vector and 1µl T4 ligase in a final volume of 10 µl. The resulting vector: insert ratio was 3:1, which maximized the number of recoverable recombinants. Reactions were incubated at room temperature for 4 hours or overnight at 4 °C before being transformed into competent *E. coli* DH5α.

2.11.10. Subcloning of DNA Fragments

PCR products amplified using *Taq* polymerase were subcloned using pGemT Easy Vector System I (Promega, UK) according to the manufacturers instructions prior to transforming competent *E.coli* DH5α. Recombinants were plated on LB agar containing 75 µg/ml ampicillin along with 0.1 mM IPTG and 50 µg/ml X-Gal enabling blue/white colour selection. Insertion of the DNA of interest interrupts the β-galactosidase gene preventing the metabolism of the chromogenic substrate X-Gal. Non-recombinants appear blue whilst recombinants appear white.

2.11.11. Isolation of Plasmid DNA

Plasmid DNA was recovered by using the Wizard SV Miniprep Kit (Promega, UK) according to the manufacturer's guidelines. Restriction enzyme digestion was carried out to ensure presence of insert.

2.11.12. Preparation of Competent Cells

E. coli from an overnight culture was diluted 1 in 100 and incubated at 37°C, 200 rpm until an OD=0.5 at λ=600nm. Cells were harvested by centrifugation at 8,000 x g for 5 min. Medium was removed and cells resuspended in 1/10 original volume of ice cold 0.1 M $CaCl_2$ and incubated on ice for 45 min. Cells were again collected by centrifugation, resuspended in 1/100 original volume of ice cold 0.1 M $CaCl_2$. Cells were then deemed ready for use.

2.11.13. Bacterial Transformation

Between 2-5 µl ligation solution was added to 50 µl competent cells and mixed by gentle aspiration before being incubated on ice for 15 min. Cells were heat shocked at 42 °C for 90 s before being returned to ice for 5 min, after the addition of 200 µl LB medium. 100 µl transformed cells were spread onto LB plates containing the appropriate antibiotic and incubated overnight at 37 °C.

Confirmation of transformants was achieved by conducting plasmids preparations and restriction digests.

2.11.14. DNA Sequencing

Samples were sent to DNA Sequencing Facility, University of Oxford, for confirmatory analysis.

2.11.15. Expression and Purification of Recombinant Protein

Cells from an overnight culture harbouring the desired plasmid were diluted 1 in 20 in fresh LB, supplemented with 75 µg/ml ampicillin, and grown at 37 °C with agitation to an OD= 0.6-0.7 (λ_{600}). A 500 µl aliquot was retained as the un-induced control. IPTG was added to a final concentration of 1 mM and cells were incubated for a further 3 hours at 28 °C with agitation. Cells from 500 µl of induced culture were collected as the induced sample. Both the un-induced and induced samples were then analysed by SDS-PAGE as outlined in Section 2.9.2.

For the purpose of protein purification the culture was scaled up. After induction cells were collected by centrifugation at 15,000 x g for 10 minutes and lysed by the addition of Bugbuster Reagent (Novagen, UK) according to the manufacturer's protocol. The resulting cell lysate was used as the starting material for batch purification of recombinant hexa-histidine tagged protein following the manufacturer's protocol.

Chapter 3

Method Development

3.1. Introduction

Previous studies in search of novel TGase enzymes have been undertaken with a limited amount of success (Motoki et al. 1991; Bech et al. 1996; Kobayashi et al. 1998c). In earlier studies the hydroxamate assay, described by Folk and Chung (Folk and Chung 1985), in which hydroxylamine is converted to hydroxamic acid through reaction with CBZ-glutaminyl glycine, was used for the detection of TGase activity. The use of this assay diminished in favour of the [^{14}C] putrescine incorporation into N,N` dimethyl casein assay, described by Lorand (Lorand et al. 1972), which was used to identify TGase producing micro-organisms from a wide variety of sources (Bech et al. 1996; Kobayashi et al. 1998c). It should be noted however, that despite an apparently wide distribution of TGase enzymes, put forward as a result of [^{14}C] putrescine assay screening (Bech et al. 1996; Kobayashi et al. 1998c), to date no additional TGase enzymes have been purified from alternative sources other than from *Streptomyces* and *Bacillus sp.*. This may suggest that the assay is not entirely TGase specific since putrescine may act as a substrate for a wide range of enzymes including polyamine/diamine acetyltranferases, which exist extensively in many micro-organisms (Haywood and Large 1985) and polyamine oxidases, flavin adenine dinucleotide-dependent enzymes involved in polyamine catabolism (Tavladoraki et al. 2006). Despite this, the [^{14}C] putrescine and hydroxamate assays have predominantly been used for the assessment of microbial TGase activity. However a plethora of equally suitable alternatives exist these include; fluorescent based assays whereby, for example, transglutaminase substrates or synthetic peptides may be labelled with a flourochrome which are subsequently immobilised by TGase mediated cross-linking and detected thereafter (Kusch et al. 2006). Alternatively upon transamidation, a fluorescent by-product may be released and subsequently detected (Gillet et al. 2005). Other assays include either calorimetric or radioactive ELISA based assays (Slaughter et al. 1992; Trigwell et al. 2004; Madi et al. 2005) or those that detect the production of reaction by-products. Since NH$_3$ is produced as a by-product of TGase protein cross-linking, it may be detected using o-phtaldialdehyde and used as an indication of extent of ε(γ-glutamyl) lysine bond formation (Flanagan and Fitzgerald 2003). All the previously mentioned assays have been used in the assessment of

mammalian TGase activity, however little is known about both their suitability and sensitivity for the assessment of microbial TGase activity.

It is the main aim of this chapter to assess and develop suitable assays capable of detecting physiological levels of TGase such that the assays may be used to detect TGase from micro-organisms *in vivo*. Additionally, the responses of both microbial and gpl TGase to a number of inhibitors shall be monitored. These shall include the synthetic molecules R281 and R283 (both potent inhibitors of eukaryotic TGase isoforms) (Griffin et al. 2006). The compounds comprise a group of structurally related analogues that have been designed to inhibit TGase mediated protein modifications where such modifications may contribute to human disease states such as fibrosis, scarring (Verderio et al. 2004) and cancer (Folk et al. 2006).

3.2. Results

3.2.1. Purification of Mammalian and Microbial Transglutaminases

Prior to assay validation purified TGase enzymes were required. TGase from *S. mobaraensis* was purchased from Ajinomoto Co. Inc. (Japan), as a 1% preparation containing 99% maltodextrin (as both a filler and preservative), and partially purified by cation exchange chromatography (Section 2.8) to remove extraneous maltodextrin (Figure 3.1). Upon SDS-PAGE analysis multiple bands surrounding the corresponding molecular weight were observed. Upon Western blotting, a large number of cross-reacting bands were highlighted. These bands are most probably due to degradation of the enzyme either during the laboratory purification or during the original industrial preparation and lyophilisation process.

Purified guinea pig liver (gpl) TGase (TG2) was used as a mammalian representative and purified accordingly (LeBlanc et al. 1999) (Figure 3.2). Due to purity and contamination issues surrounding commercially available gpl-TGase, mammalian TGase used in this study was purified 'in house' according to the method of LeBlanc (LeBlanc et al. 1999). Protein was purified from homogenised guinea pig livers by successive steps of ion exchange, gel filtration and affinity chromatography (Figure 3.2 A-C) (Section 2.7.1) pooling fractions of high specific activity. The affinity step requires comparatively large amounts of GTP in order to elute bound gpl-TGase. These amounts were found to interfere with both the hydroxamate activity assay and the Lowry protein assay. As such TGase activity was determined by the biotin cadaverine incorporation assay (represented as A450 nm) with protein content assessed by the Lowry assay post dialysis. Both protein electrophoresis and Western blotting were used to assess both purity of the purified protein confirming the presence of protein corresponding to the published molecular weight of gpl-TGase (78 kDa) (Figure 3.2 D). Western blotting confirmed a single cross-reacting band, corresponding to that observed by SDS-PAGE, whilst the commercial preparation highlighted the presence of a large amount of cross-reacting degradation products of the enzyme (Figure 3.2 E).

Figure 3.1; Purification of Microbial TGase

Microbial TGase was purified by ion exchange chromatography as described in Section 2.8. After elution with a linear gradient of NaCl (0-1M) fractions were assayed for both the presence of protein (by the Lowry assay) and their TGase activity (by the hydroxamate assay) (A). Active fractions were pooled. To determine overall TGase activity a sample was heated at 80°C for 5 min, activity was determined by hydroxamate assay (B). Protein (15μg) was visualised by SDS-PAGE (12% acrylamide gels) with coomassie blue staining. The protein was also analysed by Western blotting as described in 2.9.3, probing with 1:10,000 anti-mTG antibody and chemiluminescent detection (Ci and Cii respectively).

(A) Ion exchange Chromatography
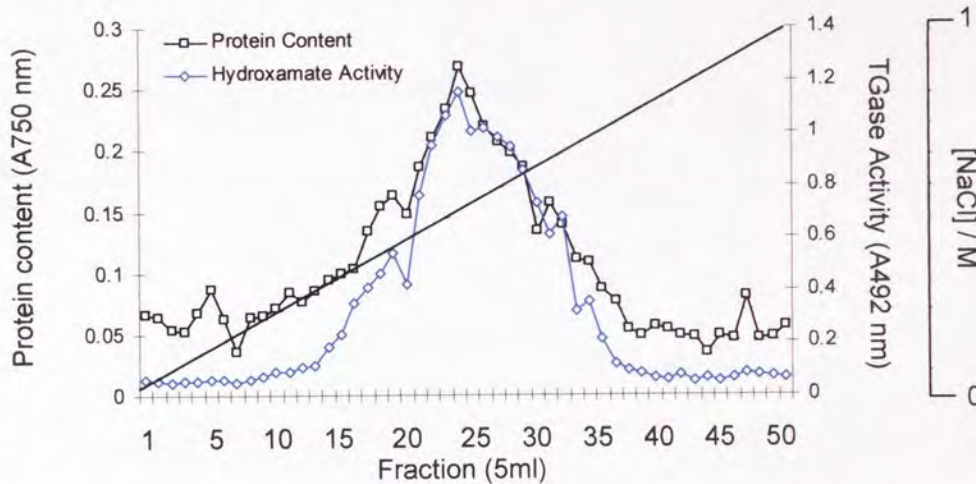
(B) Purified mTGase Activity Assessment
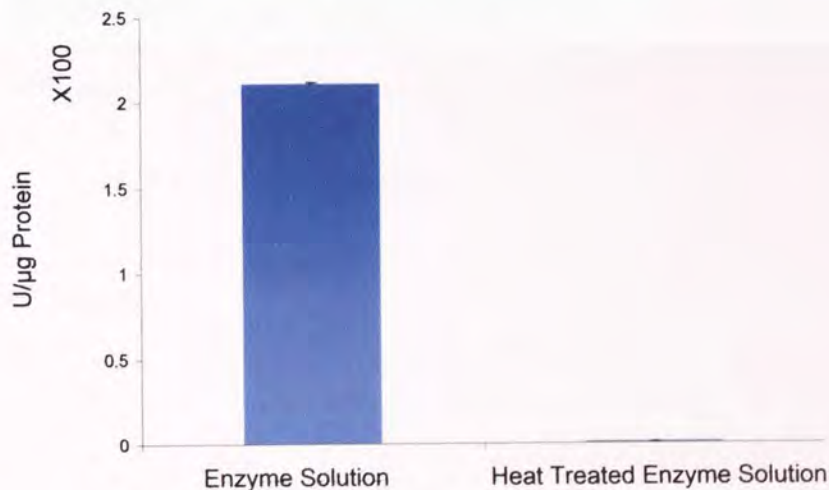
(Ci)   SDS-PAGE
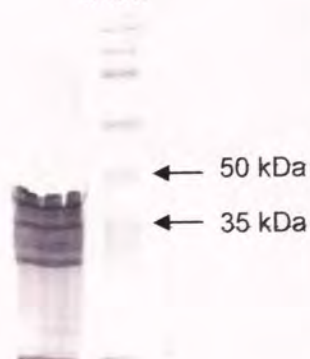
(Cii) Western Blotting
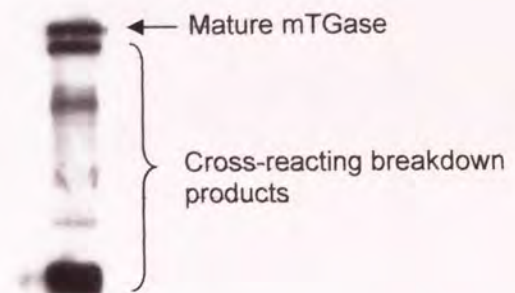
Figure 3.2; Purification of Guinea Pig Liver TGase

Guinea pig liver TGase was purified according to the methods described in Section 2.7.1. At each purification stage both the protein content (by Lowry or A280 nm) and specific absorbance values were determined (A-C). The hydroxamate assay was used for anion exchange and gel filtration with the biotin cadaverine assay used for affinity purified gpl TGase. At each stage samples of pooled active fractions were retained for SDS-PAGE analysis, using 10% acrylamide gels, stained with coomassie blue (D). In house purified gpl-TGase was compared to the commercial preparation (Sigma, UK) using SDS-PAGE (10% resolving gel) with subsequent analysis by Western blotting, as described in Section 2.9.3, using 1:10,000 CUB7402 antibody (NeoMarkers, USA). Incorporated antibody was detected by chemiluminescence.

(A) Anion Exchange Chromatography
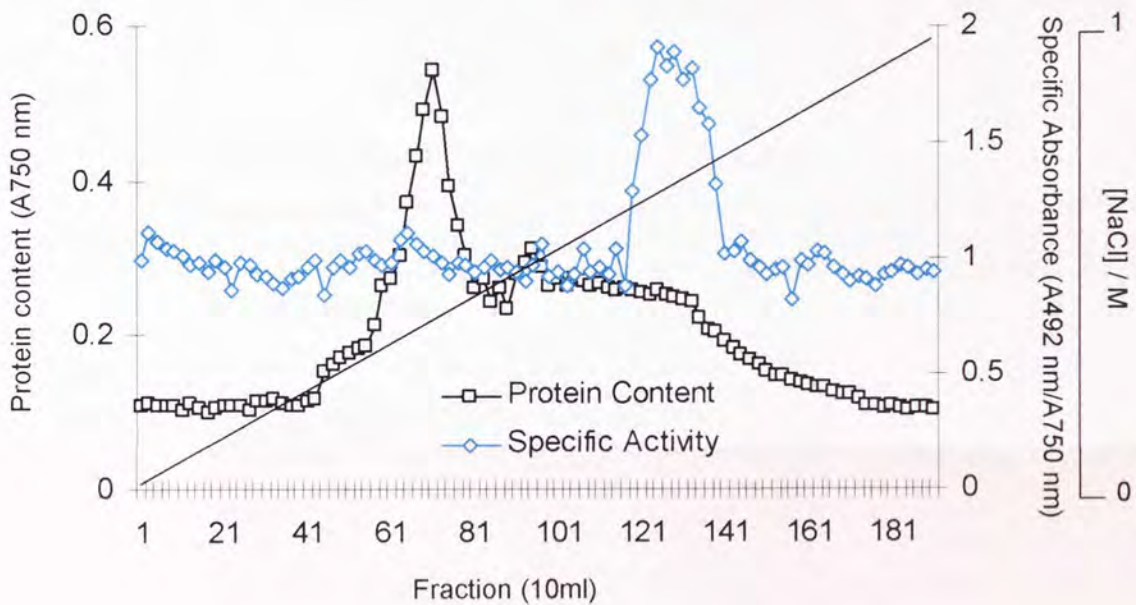
(B) Gel Filtration
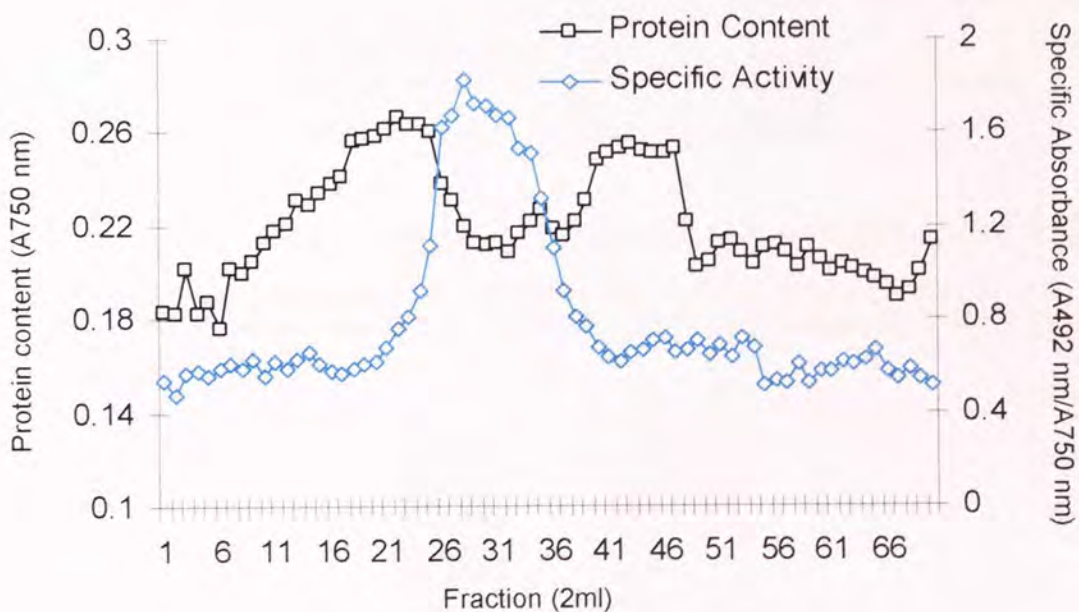
(C) *Affinity Chromatography of GPL-TGase*
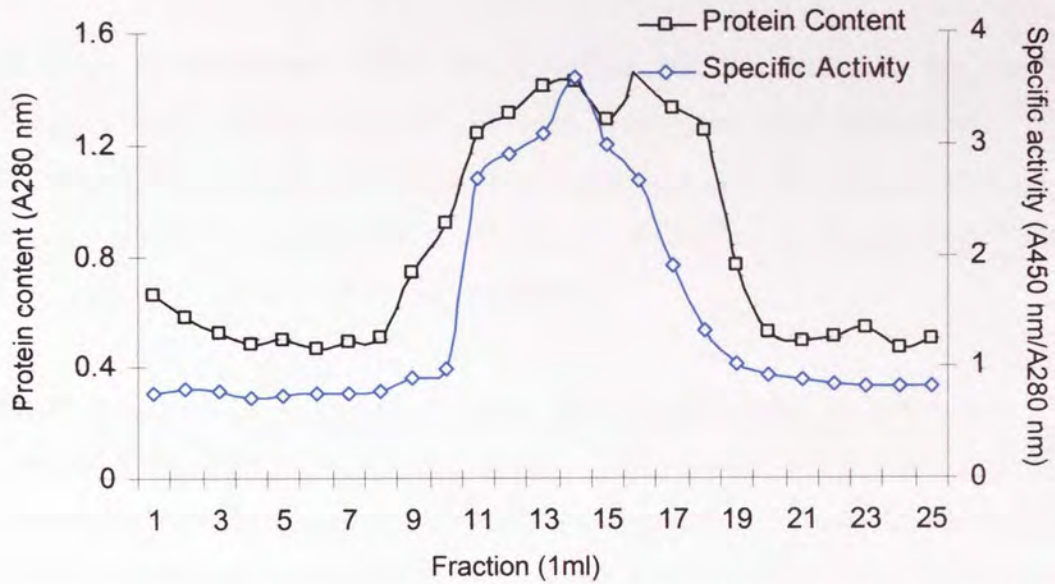
(D) *SDS-PAGE Analysis of GPL-TGase Purification*

Guinea pig liver homogenate (1), post ion exchange, gel filtration and affinity purified gpl TGase fractions (2,3,4 respectively) were separated by SDS-PAGE (10% resolving gel) and stained by coomassie brilliant blue. The boxed region represents TGase protein. MwM denote molecular weight markers.
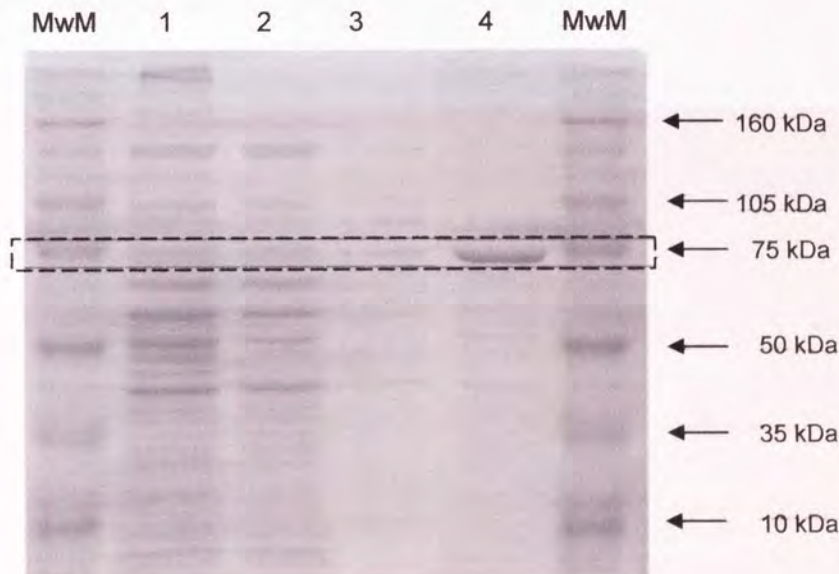
(E) *Western Blotting of Commercial and 'in house' Purified GPL TGase*

Commercial gpl-TGase (1) and affinity purified gpl-TGase (2) were separated by SDS-PAGE (10% resolving gel) and probed with anti-tTG CUB-7402 (Neomarkers, USA, 1:10,000) as described in Section 2.9.3.

3.2.2. Validation of TGase Assays

In order to ascertain which assay would be of most use for identifying TGases each detection method was evaluated and compared. It was envisaged that such investigations would gauge the merits of individual assays whilst comparing the activities of mammalian and microbial TGase in response to a number of TGase inhibitors.

Each assay was performed using known amounts of microbial TGase ranging from 200 µg/ml to 20 ng/ml. The results were then analysed to determine sensitivity and reproducibility (Figure 3.3). The hydroxamate assay was the least sensitive of the three assays used (reliable to 2µg/ml), whilst the biotin cadaverine and peptide cross-linking assays were both more sensitive and more reproducible and capable of detecting TGase activity of 20 ng/ml and 5 ng/ml respectively. However, the peptide cross-linking assay resulted in a reduced background count and was also capable of distinguishing lower levels of TGase (1-10 ng/ml).

Figure 3.3; Validation of TGase Assays

To validate TGase assays initially a serial dilution (200µg/ml-20ng/ml) of mTGase was performed. Samples of mTGase were assessed for activity in the hydroxamate (A), biotin cadaverine (B) or peptide-cross-linking (C) assays as described in Section 2.5. After completion absorbance values were read enabling the assessment of assay sensitivity. Samples were assayed in triplicate with bars representing the mean values from three independent experiments with error bars representing standard error.

(A)   Hydroxamate Assay
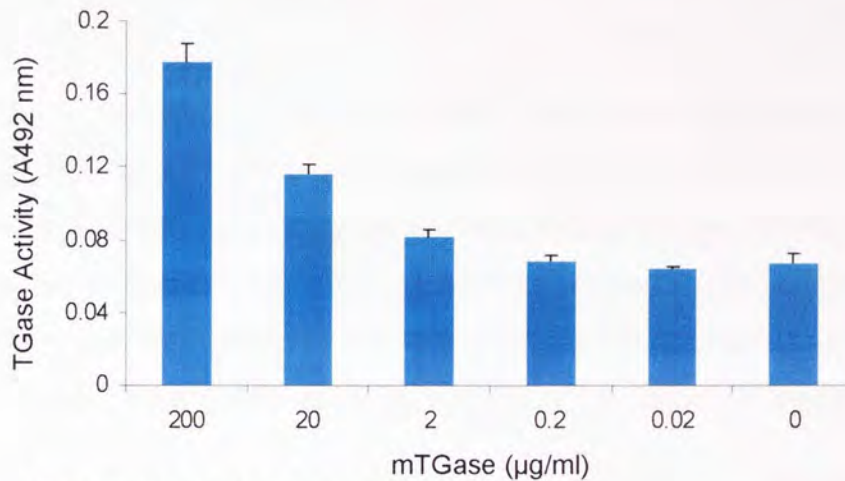
(B)   Biotin Cadaverine Assay
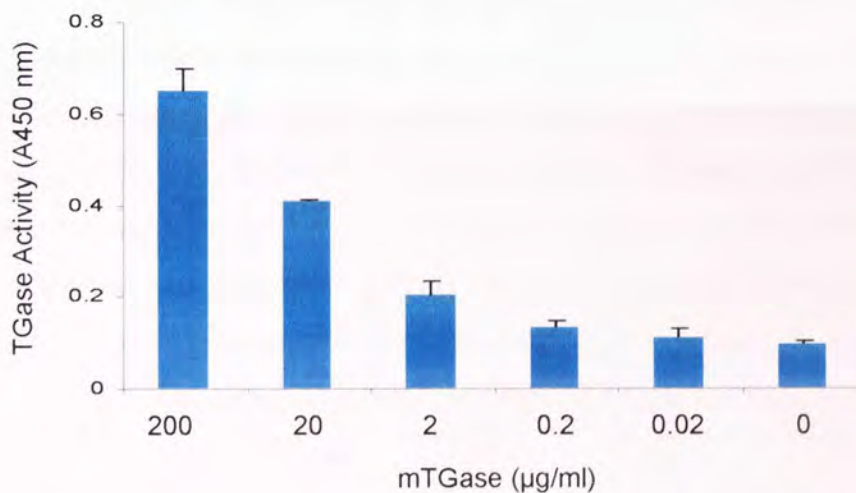
(C)   Peptide Cross-Linking Assay
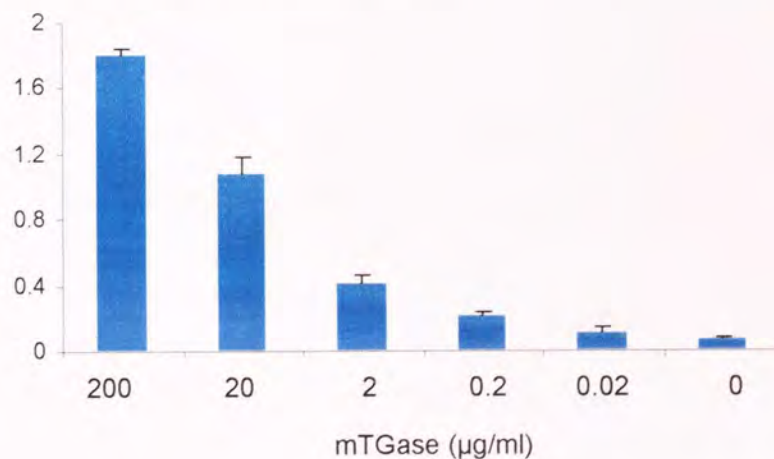
3.2.3. Validation of TGase Assays using TGase Derived from B. subtilis

It was not known whether the TGase assays described would be successful in detecting physiological levels of microbial TGase. As such, an investigation was undertaken to address whether such amounts of enzyme may be detected and whether the presence of additional proteins/proteases would interfere with the TGase assays used.

Microbial TGase activity has been described previously in *B. subtilis* (Kobayashi et al. 1998b) and was selected for the validation studies, as an alternative to the *Streptomyces* enzyme. Since TGase activity in *B. subtilis* is attributed to spore coat formation during sporulation (Kobayashi et al. 1998b) the organism was grown to the relevant stage of sporulation prior to assaying for activity. The spore suspension was then used to validate the TGase assays.

As can be seen from Figure 3.4 mTGase activity was detected in the resuspended spore suspension using either the biotin cadaverine or peptide cross-linking assays. Cells collected prior to lysis were also spiked with a known amount of *S. mobaraensis* TGase to assess both the amount of protein degradation under physiological conditions, and the affect of the lysis procedure on the stability of the enzyme (Figure 3.4). After treatment a large proportion of enzyme activity remained (92%) indicating that the lysis process and endogenous proteases had a minimal effect on the viability of the *S. mobaraensis* TGase enzyme.

Figure 3.4; Effect of the Lysis Procedure on Viability of S. mobaraensis TGase

B. subtilis was grown, in 500ml Schaeffer's sporulation medium, for 6 hours after sporulation was initiated. Cells were collected by centrifugation at 13,000 rpm for 15 min and divided in two. One half was treated with lysis solution whilst to the second half was added 10μg/ml (final concentration) mTGase in lysis solution as described in Section 2.4.7.1.. Resuspended cells were incubated for 2 hours at 4°C with end over end rotation. The lysates generated were assessed for mTGase activity using the biotin cadaverine assay as described in Section 2.5.2. In addition, the lysis solution and 10 μg/ml mTGase (in 50 mM Tris.HCl) were assessed also. Presented are the mean values + standard error of three independent experiments, each completed in triplicate.
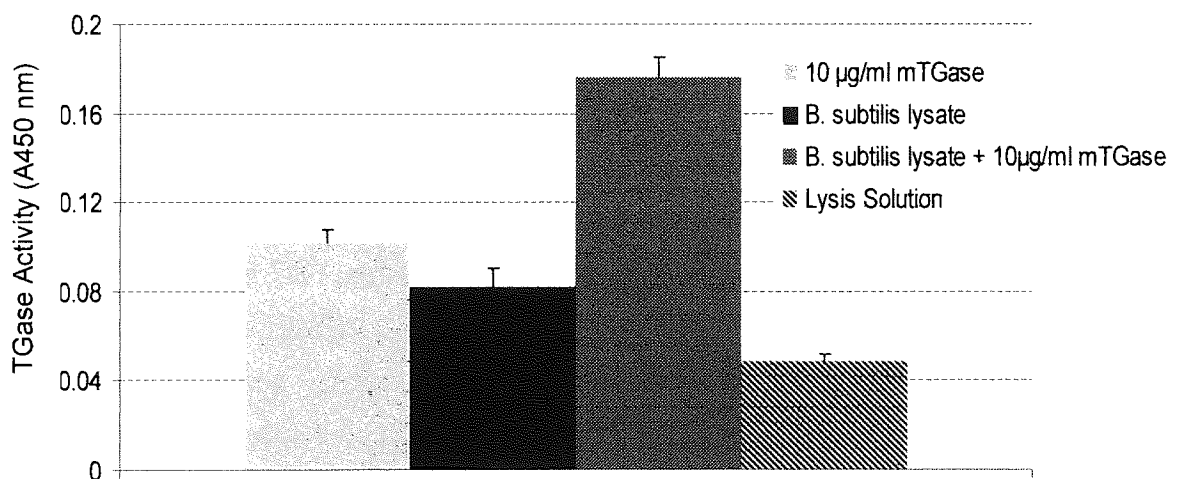
3.2.4. Effect of Inhibitors on TGase Activity

To assess the inhibition profile of likely TGase enzymes the affect of a number of inhibitors on the action of both microbial and mammalian type TGase (derived from *S. mobaraensis*, and Guinea pig liver, respectively) was monitored (Figure 3.5).

In most cases the enzymes behaved in a similar manner to each other. However pronounced differences were uncovered following challenge with the synthetic irreversible inhibitors R281 and R283. Challenge with primary amines produced a dose dependent reduction in the amounts of biotinylated substrate incorporated into casein. In both cases inhibition with N-ethylmaleimide confirmed that enzymatic activity was dependent on a cysteine residue. It is envisaged that novel TGase enzymes should exhibit similar enzymatic characteristics upon challenge with the same inhibitors as do the enzymes already known.

Both R281 and R283 were found to have a profound effect on the activity of gpl TGase reducing activity by up to 90% (molar ratios of $1:9.6 \times 10^5$ R281 and $1:3.8 \times 10^7$ R283 for 200 µg/ml gpl TGase). With *S. mobaraensis* TGase, R281 reduced enzyme activity by approximately 90% though R283 had little effect on its activity (molar ratios of $1:4.72 \times 10^5$ R281 and $1:1.9 \times 10^7$ R283 for 200 µg/ml mTGase). Prior incubation (1 hour at 37 °C) of the effective synthetic irreversible inhibitor with the TGase solution was found to exacerbate the amounts of inhibition, though the extent of increase appeared dependent on the amount of enzyme present (Figure 3.6).

Figure 3.5; Effect of Inhibitors on the Activity of Microbial and GPL TGase

The response of purified microbial (A) and gpl-TGase (B) to various substances was assessed. In each case 20µg/ml TGase (final concentration) was applied to the wells of a 96-well plate (coated with10 mg/ml N'N' dimethylcasein in 50 mM Tris.HCl pH 7.4) in triplicate. Inhibitors were added to the biotin cadaverine reaction solution to the concentrations shown. The biotin cadaverine assay was completed as described in Section 2.5.2 with the absorbance of samples read at 450 nm and compared to uninhibited enzyme samples. In each case data represents the mean of three independent experiments + standard error.

(A)
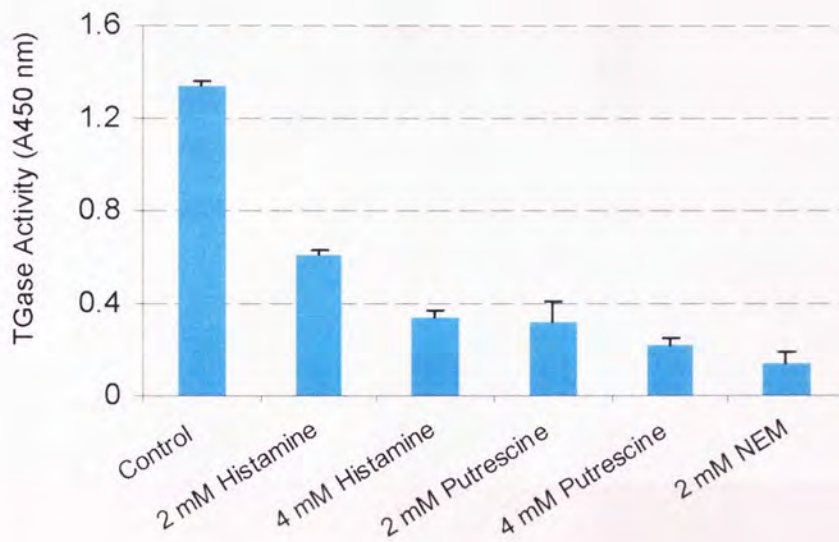
(B)
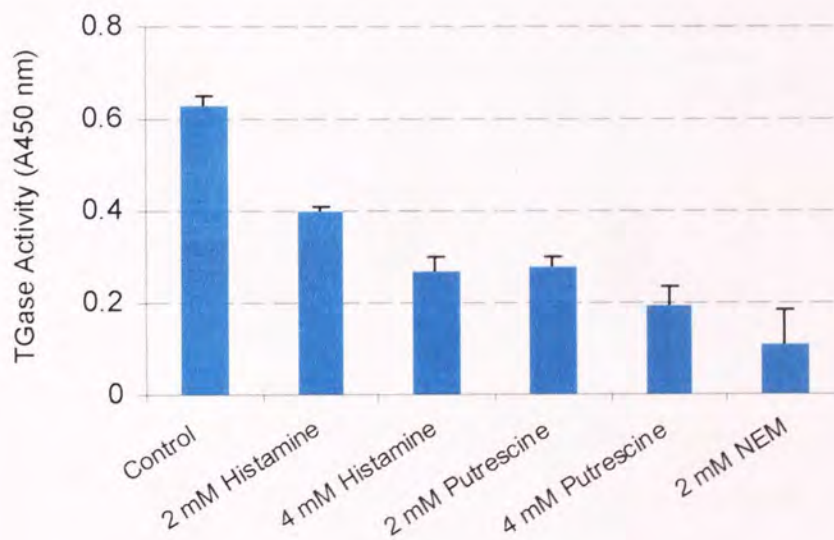
Figure 3.6; Effect of R281 and R283 on the Action of Microbial and GPL TGase

Varying concentrations of both microbial (A) and gpl (B) TGase (200µg/ml – 2 µg/ml) were assayed in the presence of the synthetic inhibitors R281 and R283 (250 µM and 1 mM, respectively) using the biotin cadaverine assay as described in Section 2.5.2. Prior to assessment enzyme solutions were also incubated (Inc.) with inhibitors for 1 hour at 37°C prior to assay. After quenching, by the addition of H_2SO_4, absorbance values at A450 nm was determined and values compared. Data presented are from three separate experiments.

(A)
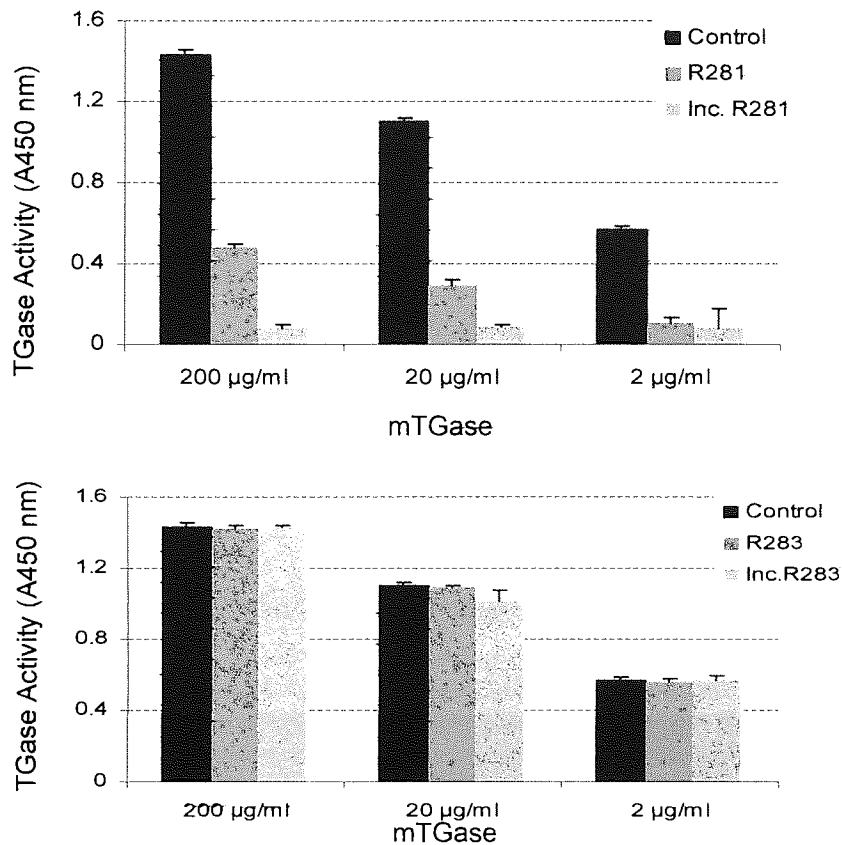
(B)
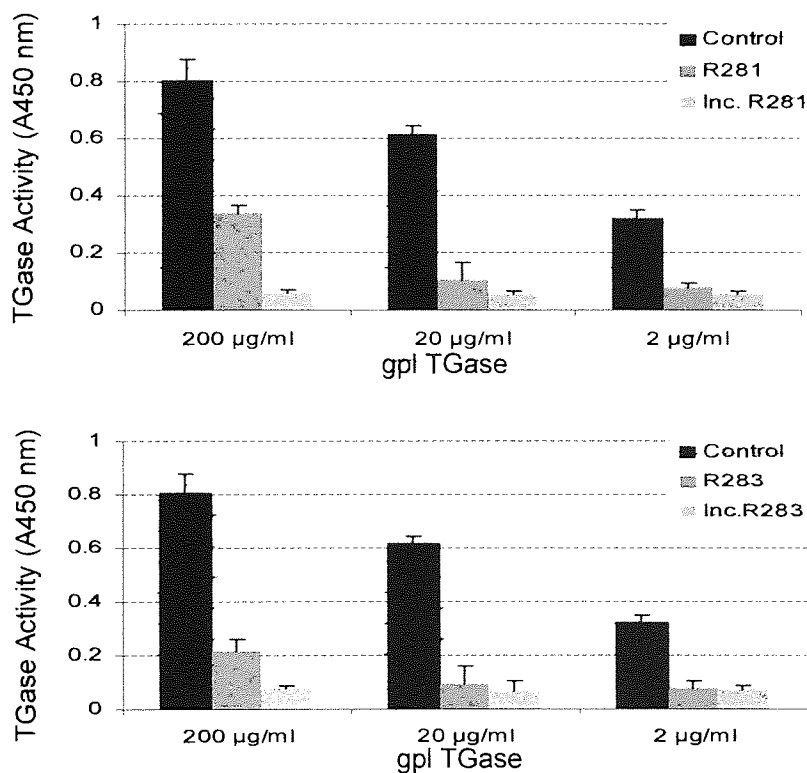
3.2.5. Purification of Microbial TGase using R281-Coupled Sepharose

The synthetic irreversible TGase inhibitor R281 has been shown to be effective at suppressing TGase enzyme activity (Section 3.2.4) (Griffin et al. 2006). It would prove useful if the inhibitor could be used not only to confirm enzyme activity but also to covalently bind novel TGase proteins. The chemical structure of R281 allowed the creation of R281 coupled Sepharose by reaction of the free carboxyl on R281 with EDC (N-ethyl-N'-(3-dimethylaminopropyl)carbodiimide) and subsequent conjugation to the free amine of the EAH Sepharose 4B (Section 2.6).

After washing, bound protein was removed from the resin by treatment with trypsin. Visualisation by SDS-PAGE revealed that the majority of protein appeared to bind to the resin after 4 hours incubation (Figure 3.7). Potential TGase positive strains identified by alternative assays should be incubated for 4 hours with in this method.

Figure 3.7; Protease Treatment of mTGase Coupled R281-Sepharose

200 µg mTGase in 50 mM tris-HCl pH7.4 was incubated with 1 ml (settled volume) of R281-coupled sepharose, previously washed three times with 50 mM Tris-HCl. The resin and mTGase mix was incubated with end over end rotation for 2, 4 and 6 hours at 4 °C to enable ligand binding. After incubation, the resin was washed with a linear gradient of NaCl (0-5M) to remove unbound material with covalently attached protein removed by trypsinisation, as described in Section 2.6.1. Protein was then visualised by SDS-PAGE (12% resolving gel) and coomassie blue staining.

MwM 6 Hours 4 Hours 2 Hours

50 kDa

35 kDa

30 kDa

Full length
mTGase released
from R281 resin
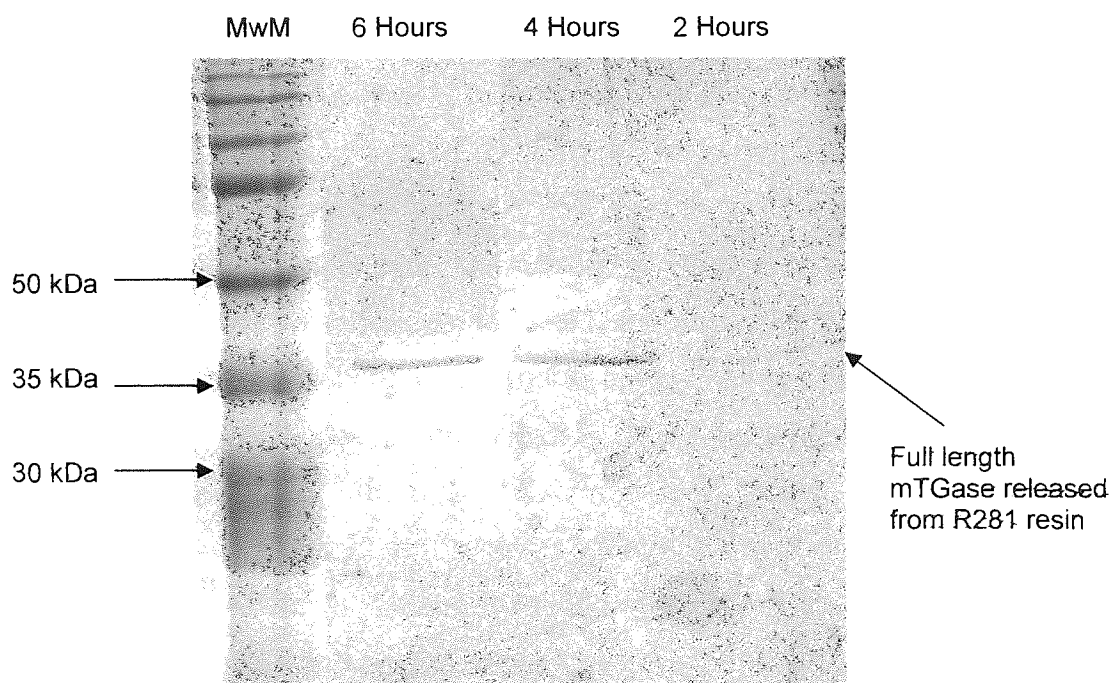
3.3. Discussion

Prior to assay validation, high quality enzymes were required to ensure that the activity observed was that catalysed by TGase. Microbial TGase was purified from the commercially available preparation and was found to be greater than 90% pure despite a large amount of protein degradation, a likely product of the preservation process. Mammalian TGase was purified to homogeneity (assessed by SDS-PAGE) in house from Guinea pig liver ensuring superior activity and purity to the commercially available enzyme (Sigma Aldrich, UK).

A number of assays, previously developed to detect mammalian type TGases (Folk and Chung 1985; Slaughter et al. 1992; Trigwell et al. 2004), were assessed for their ability to reliably detect microbial TGase. The hydroxamate assay, whilst quick to carry out, was least sensitive and perhaps explains why only a limited number of organisms were identified as TGase positive originally (Motoki et al. 1992). However, the speed of the assay proved advantageous for assessing TGase activity during enzyme purifications, limiting protein degradation and purification time.

The ELISA based assays described by Slaughter (Slaughter et al. 1992) and Trigwell (Trigwell et al. 2004) demonstrated improved reliability, reproducibility and sensitivity compared to the hydroxamate assay. At the beginning of the study only the biotin cadaverine assay was available and was assessed to ensure it was capable of detecting physiological levels of TGase from the TGase producer *B. subtilis*. This provided, to some extent, a degree of confidence in the ability of the assays to detect the low levels of TGase that one might expect to encounter. Upon availability, the peptide cross-linking assay was deemed an excellent assay for confirming potential TGase activity since it measures peptide cross-linking as opposed to amine incorporation. Additionally the assay also demonstrated a reduced background giving it improved discriminatory power over the biotin cadaverine assay Since both the biotin-cadaverine and peptide cross-linking assays assess two of the different reactions catalysed by TGase enzymes, (cross-linking ability and amine incorporation) thereby increasing both the specificity and sensitivity to potential microbial TGase enzymes.

Since the hydroxamate assay was least sensitive it was omitted from subsequent TGase analyses involving inhibitors. The competitive primary amine substrates putrescine and histamine evoked a dose dependent reduction of TGase activity with both microbial and gpl TGase. The presence of N-ethylmaleimide almost completely abolished TGase activity confirming the requirement of a cysteine residue for enzymatic activity. Interestingly, stark differences were observed between microbial and gpl TGase upon challenge with the synthetic irreversible inhibitors R281 and R283. In the case of gpl TGase, both inhibitors were effective in reducing enzyme activity. However, only R281 was effective against mTGase with little or no effect caused by R283. Since both microbial and gpl TGases share a high degree of similarity on their active site domains one possible explanation for this is that R281 is more specific for the active site cysteine. The inhibitor R283, instead of covalently binding to the active site cysteine, may bind to any other cysteine residue resulting in a disrupting protein conformation and ultimately, at high concentrations, sequestering enzyme activity. Since gpl TGase comprises eighteen cysteine residues and mTGase contains a single cysteine (cys64 Figure 1.5) one might predict that R283 would have a greater effect on mammalian-type TGases, this may present an interesting avenue for further investigation.

As an alternative strategy for identifying novel microbial TGases, the site directed inhibitor R281 was bound to a solid phase and used for its affinity for TGase proteins. This assay was effective at binding *Streptomyces* TGase though large scale studies were not completed due to time constraints. It is thought that the procedure would aid the identification and purification of novel TGase proteins by identifying proteins possessing the TGase active site. Such proteins could be released from the resin by protease treatment, separated and sequenced.

The assays described provide a sound basis for a biochemical screen to search for novel TGase enzymes. The evidence provided in this study shows that the biotin cadaverine assay can be used for routine analysis of micro-organisms and that confirmation of cross-linking activity should be completed by the peptide cross-linking assay. In addition, lysates appearing positive

should be subjected to inhibition analysis since one might expect novel TGases to behave in a similar manner in their response to the inhibitors. One would expect enzymes that carry out similar functions to possess similar structural and enzymatic characteristics as is the case with mammalian type TGases. Such characterisation would therefore aid the identification of novel microbial protein cross-linking enzymes.

Chapter 4

Screening for Novel Microbial Protein Cross-Linking Enzymes

4.1. Introduction

To date a multitude of TGase isoforms have been identified from an increasing number of organisms (Section 1.2). From an evolutionary perspective, eukaryotic TGases share similarities in both their active site and reaction mechanism with the papain-like thiol proteases (Pedersen, et al. 1994) and as such have been grouped together within the same superfamily in the Structural Classification of Proteins (SCOP) database. Since TGases from bacteria do not resemble their eukaryotic counterparts they have not been included in this group, and occupy a place within the cysteine proteinase superfamily (West, et al. 2003).

The identification of enzymes originating from bacteria, based on structural relationships between either prokaryotic or eukaryotic TGases, remains a slow process. However, using both computer analyses of profiles generated from the PSI-BLAST program (Altschul, et al. 1997) and using sequence similarity searches coupled with fold-recognition programs (ORFeus and meta-BASIC) (Ginalski, et al. 2004) have identified protein superfamilies homologous to eukaryotic TGase, as well as a novel family of bacterial TGase like cysteine proteinases originating from microbes have been identified (Makarova, et al. 1999). A number of proteins homologous to animal TGases were identified in the genomes of bacterial and archaeal representatives (Makarova, et al. 1999). These proteins represent a family of TGase-like proteins that share a number of structural features with mammalian TGases including: the location of small amino acid residues directly upstream of the catalytic cysteine; the aromatic residues found two amino acids downstream of the catalytic histidine and the aromatic residues found flanking the catalytic aspartate N-terminal region. Although, the genomes of the organisms in question have been sequenced, most proteins are awaiting functional characterisation (Makarova, et al. 1999).

A family of Bacterial TGase-Like Cysteine Proteinases (BTLCP) have been identified based on fold-recognition programs (Ginalski, et al. 2004). This family has been predicted to comprise an invariant Cys-His-Asp catalytic triad and shown to possess an N-terminal signal sequence suggesting an

extracellular target for BTLCP function, either by modifying cellular components or assisting bacterial virulence. However, the exact substrates and reaction kinetics remain to be determined. The BTLCPs described exclusively belong to the proteobacteria, showing a scattered distribution, with one exception, *Anopheles gambiae,* thought to be as a result of a contaminating bacterial sequence in the genome assembly (Mongin, et al. 2004). Many of these proteins belong to an uncharacterised cluster of orthologs COG3672 (Tatusov, et al. 1999). Intriguingly a transitive PSI-Blast search (using Gi: 1612442) reveals an archaeal sequence, originating from *Archaeoglobus fulgidus (*Gi: 11498787), that has been previously described as homologous to animal TGases (Makarova, et al. 1999). This implies, at least to some extent, a relationship between the BTLCP family and mammalian TGases.

TGases from *Bacillus* and *Streptomyces sp.,* Cytotoxic necrotising factors 1 and 2 from *E. coli,* and dermonecrotising toxin from *Bordetella pertussis* are not included in the profiles generated (Makarova, et al. 1999; Ginalski, et al. 2004). Despite possessing the catalytic cysteine and histidine residues, they share no significant similarity to mammalian TGases. This would suggest that there exists two separate lineages within microbes one homologous to animal TGases that encode proteases whilst the other encodes TGases carrying out protein cross-linking. The latter of which would appear to be restricted to a few genera.

Microbial TGase was initially identified by biochemically screening a large number of organisms for a protein cross-linking enzyme (Ando et al., 1989). Using fold recognition programs and detailed structural similarity searches to identify microbial homologues to mTGase would present an incredibly difficult task, despite a degree of homology in the active site region; mTGase from *S. mobaraensis* has a unique 3-D structure (Kashiwagi, et al. 2002a). Despite advances in identifying potential sources of bacterial TGases none have been definitely confirmed as exhibiting TGase activity experimentally (Makarova, et al. 1999; Ginalski, et al. 2004). These searches, however, may provide an insight as to the likely origins of microbial derived isoforms and in addition to gene homology searches, based on pre-existing bacterial

TGases, shall be used to identify candidate organisms that may comprise a TGase enzyme. Upon confirmation of TGase activity, using the assays described in Chapter 3, the protein responsible shall be purified and characterised using standard protocols.

4.2. Results

4.2.1. Database Searching for TGase Homologues

Prior to screening, a comprehensive search of the non-redundant databases was completed in an attempt to identify organisms harbouring potential TGase enzymes. Searches were completed using the protein sequence of TGases from both *Streptomyces mobaraensis* (Figure 4.1) and *Bacillus subtilis* (Figure 4.2). These sequences were used to probe the translated nucleotide database (tBLASTn) (Altschul, et al. 1997) for potential homologues.

In both cases, TGases were identified in closely related species. Within the *Streptomyces* genus, publicly available TGase sequences were found to share a high degree of identity (>79%), with greatest diversity appearing in a strain of *S. cinnamoneum*. Likewise a number of genes encoding TGase proteins were highlighted when a search was initiated using the protein sequence of TGase from *B. subtilis*. Of those identified, a large proportion resided in closely related species, however, their extent of identity varied with the most closely related proteins sharing 60% identity (*B. licheniformis*) and the most diverse members showing only 36% identity (*B. cereus* and *B. anthracis*). A notable addition to the B*acillus*-type TGases is that from the thermophile *Geobacillus kaustophilus.* This bacterium resides within its own distinct phylogenetic cluster (Ash, et al. 1991) and as such reiterates the ubiquitous distribution of *Bacillus*-type TGases throughout the *Bacillus* family.

Organism (Accession Number)	E value	Identity (%)
S. mobaraensis IFO13819 (AF531437)	0	100
Eukaryotic synthetic construct (AB181959)	0	100
Stv. mobaraensis (Y18315)	0	100
S. mobaraensis WSH-Z2 (DQ132977)	0	100
S. mobaraensis (AY502065)	0	99
S. mobaraensis (AY762265)	0	93
S. mobaraensis (AY129279)	0	93
S. mobaraensis (AY 241675)	0	93
S. mobaraensis (AY 502064)	0	92
S. mobaraensis (AY615816)	1e-173	93
S. mobaraensis (AY422087)	3e-163	82
S. cinnamoneus (AB085698)	2e-162	81
S. platensis (AY555726)	1e-159	79
S. cinnamoneum (Y08820)	2e-155	79

Figure 4.1; tBLASTn Search of mature TGase from S. mobaraensis

The results of the search are aligned with the 331 amino acids of mature TGase from S. mobaraensis.
e-values of 0 represent that highlighted sequences are identical.
The tBLASTn Program was accessed at www.ncbi.nih.gov/blast and the search was completed May 2006.

Organism (Accession Number)	E Value	Identity (%)
B. subtilis (Z99119)	9e-140	100
B. subtilis (Z93935)	9e-140	100
B. licheniformis ATCC 14580 (Cp000002)	3e-82	60
B. licheniformis DSM 13 (AE017333)	3e-82	60
B. clausii KSM-K16 (AP006627)	9e-47	42
Geobacillus kaustophilus HTA426 (BA000043)	1e-144	44
B. halodurans C-125 (BA000004)	8e-42	40
B. cereus E33L (CP000001)	1e-39	36
B. thuringiensis (AE017255)	4e-39	36
B. anthracis (AE017334)	4e-39	36
B. anthracis (AE017225)	4e-39	36
B. cereus ATCC 14579 (AE016877)	4e-39	36
B. anthracis (AE016879)	4e-39	36
B. cereus ATCC 10987 (AE017194)	3e-27	42
Polaromonas sp. JS666 (CP000316)	0.017	26

Figure 4.2; tBLASTn Search of TGase from B. subtilis

The results of the search are aligned with the 245 amino acids of TGase from B. subtilis. e-valuates of 0 represent that highlighted sequences are identical. The tBLASTn Program was accessed at www.ncbi.nih.gov/blast and the search was completed May 2006.

4.2.2. *Which Micro-organisms to Screen for TGase Activity?*

The results of the database searches undertaken failed to highlight novel organisms that may harbour a potential TGase enzyme (Section 4.2.1). It was therefore necessary to select organisms to screen for TGase activity using the methods described in Chapter 3. Obvious candidates included those within the same genus, and closely related organisms. Organisms were also selected based on their unique characteristics to include, but not limited to: psychrophiles; thermopiles; osmotolerant organisms, organisms exhibiting diverse metabolic characteristics. Large scale bacterial screening in search of novel TGase has been undertaken previously (Motoki, et al. 1992; Andou, et al. 1993; Bech, et al. 1996). From those organisms screened a number have been shown to possibly possess TGase activity, though to date the protein has not been purified. Such organisms therefore make excellent candidates to screen.

4.2.2.1. Screening Methodology

TGase from *Streptomyces sp.* and *Bacillus sp.* act on extracellular and intracellular environments, respectively (Kanaji, et al. 1993; Kobayashi, et al. 1998b). Therefore it is reasonable to assume that any potentially new TGase homologues could be active in either of these environments. For this reason the culture medium of organisms as well as cell lysates would be assayed for TGase activity. Activity of each bacterial TGase isoform would appear to be focussed around the late exponential and stationary phases of growth; these phases involve the transition from vegetative cell to spore/spore bearing structure. Therefore analysis of sporulating micro-organisms (*Bacillus sp.*, higher bacteria, and fungi) was completed around this transitive stage in the life cycle. Samples of other bacteria screened were to be taken at various time points and their TGase activity assessed by one of the aforementioned enzyme assays (Section 2.5).

4.2.2.2. European Project HIPERMAX (Project No. NMP-3-CT-2003-505790)

HIPERMAX is focused on the production of high performance protein matrices for industrial use in, for example, the textiles arena. High performance protein matrices would be produced via TGase catalyzed reactions. Such modification of protein matrices would require a TGase enzyme that could be widely used and patented for commercial applications. This is important because the existing patent for *S. mobaraensis* TGase is close to expiry. The focus of the project is therefore to identify a novel source of microbial origin TGase that could be tested and eventually applied in a number of industrial processes.

To identify novel TGase enzymes, large scale screening of micro-organisms has been undertaken by a number of European partners, concentrating on the higher bacteria, and fungi. Supplementary testing has been carried out as part of the current study, to confirm or deny TGase positive strains.

4.2.3. Screening Micro-Organisms for TGase Activity
4.2.3.1. Detection of TGase from Fungal Isolates

In liquid media, fungi adopt a filamentous growth pattern similar to that of the Actinomycetes (Kalakoutskii and Agre 1976). As such, culture aliquots were extracted at various time-points and analysed by the hydroxamate assay (Section 2.5.1) for TGase activity (Figure 4.3). Five *Aspergillus sp.,* including three belonging to the genus *Aspergillus terreus* (G. Szackacs, Technical University, Budapest), were identified as hydroxamate positive and selected for further study using alternate TGase assays.

For all *Aspergillus sp.* maximal potential TGase activity was observed between 24 and 30 hours after inoculation into medium TG-7 (Appendix IV)(Figure 4.3). However, despite apparent hydroxamate activity no enzyme activity could be detected using either the biotin-cadaverine or peptide cross-linking assays using the same culture supernatants, possibly as a result of protease activity. In an attempt to clarify the presence or absence of a TGase enzyme, 500 ml culture supernatant of *A. terreus* IFO 7079 (found to exhibit maximum potential TGase activity by hydroxamate) was collected after 24 hours (grown in medium TG-7) filtered and partially purified by anion exchange chromatography (Figure 4.4); Enzyme activity (by hydroxamate) was found to trail the protein profile (Figure 4.4). Further activity assays (biotin cadaverine and peptide cross-linking assays) verified that there was no TGase activity. Furthermore, no cross-reactivity was detected when partially purified fractions were probed with anti-transglutaminase antibodies raised against tTGase (CUB7402, NeoMarkers, USA) and mTGase (N-zyme Biotech., Germany). As such, purification from fungi was halted with in favour of other classes of micro-organisms.

Figure 4.3; Detection of TGase from Fungal Cultures

Isolates were inoculated into 250 ml of the relevant nutrient rich medium (see Appendix IIIb) and incubated at 30°C with shaking (200 rpm). Media (1 ml) was extracted at the time points outlined below and cellular debris was removed by centrifugation (5 min at 13,000 rpm) to generate cell free extracts. Extracts were analysed for TGase activity by the hydroxamate assay.
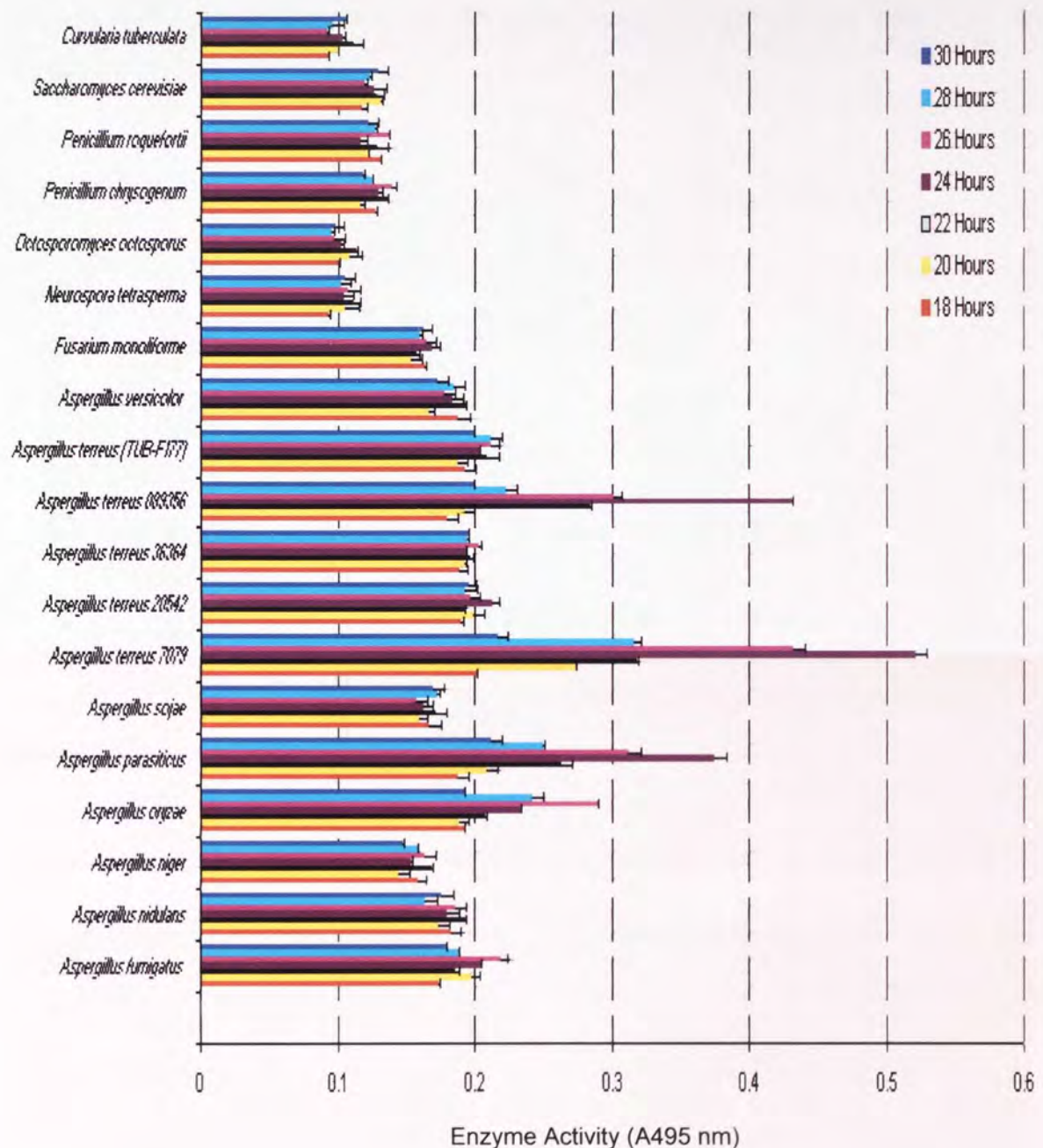
Figure 4.4; Partial Purification of Potential TGase from A. terreus IFO 7079.

A. terreus IFO 7079 was grown in 250 ml medium TG-7 for 26 hours at 30°C with shaking (225 rpm). Cells were removed by centrifugation at 8,000 x g for 15 min at 4 °C. The resulting supernatant was loaded onto Q Sepharose and eluted with a linear gradient of NaCl (0-1M) over 60 min at a flow rate of 5 ml/min collecting 5 ml fractions. Fractions collected were assessed for protein content (Lowry assay) and hydroxamate activity and have been shown as specific activity (A495/A750) (A). Fractions 13 to 29 were analysed by SDS-PAGE (B) using 10% acrylamide gels as described in Section 2.9.2.

(A)
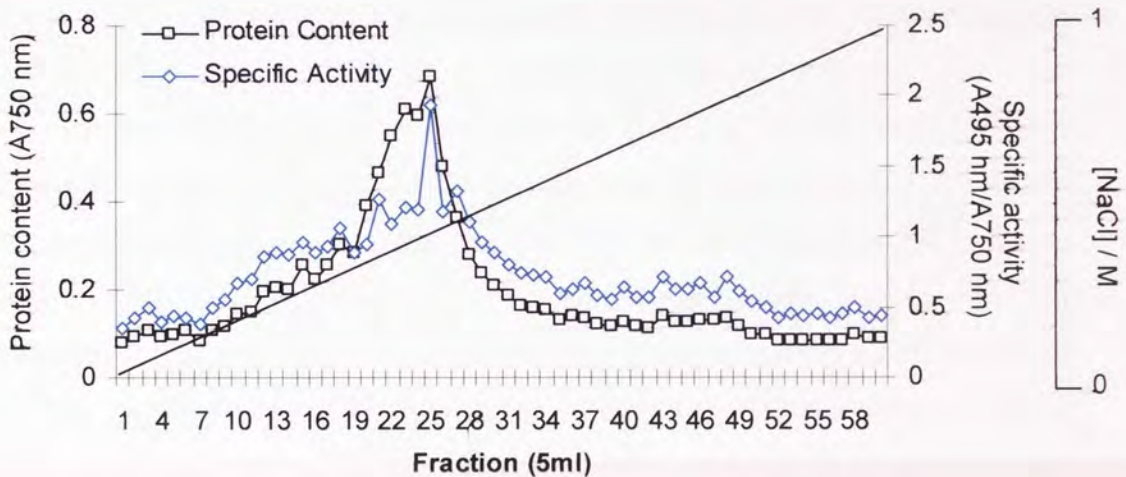
(B)
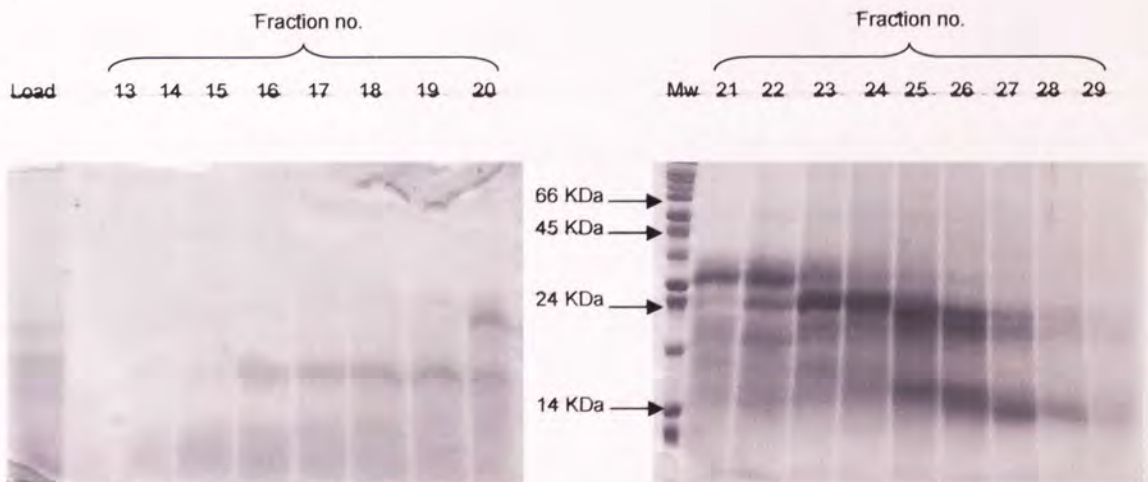
4.2.3.2. Identification of TGase from the Actinomycetes

As part of the European Project HIPERMAX a number of organisms were identified as exhibiting TGase activity by the hydroxamate assay (identified by G. Szakacs, Technical University of Budapest) and were sent by partners to Aston University for confirmatory testing. Extracellular TGase activity was confirmed in six species (Figure 4.5), using both the biotin cadaverine and peptide cross-linking assays (Section 2.5.2 and 2.5.3). These assays are more sensitive and therefore more reliable than the hydroxamate assay.

In each case, Actinomycete species were inoculated onto GYM agar and incubated for 10 days at 30 °C to provide the spore innoculum for liquid growth studies. Maximum TGase activity was achieved after 6 days growth in medium GYM. Growth rate was measured daily by extracting 5 ml of culture broth and determining the wet cell weight as outlined in Section 2.4.5. A time course of activity was determined, however for simplicity, activity between days 4 and 8 (encompassing maximum activity) have been presented (Figure 4.5). The organism *Saccharomonospora viridis* exhibited maximum TGase activity after 4 days, with actinomycetes strains exhibiting maximal TGase activity after 6 days growth.

TGase activity was only detected in a few of the actinomycetes screened belonging to the *Streptomyces* and *Saccharomonospora* genera. The purification, characterisation and physiological studies carried out on these enzymes are presented in Chapter 5.

Figure 4.5; Actinomycetes Screened for TGase Activity

Organisms were plated on nutrient rich medium (GYM) to generate the spore innoculum. The spore innoculum was grown in 5 ml GYM medium, at 30°C with shaking at 225 rpm, as a starter culture to inoculate 25 ml GYM medium under identical conditions. Medium was sampled daily and cellular debris was removed by centrifugation at 15,000 x g for 5 mins. The resulting cell-free medium was assessed for TGase activity by the biotin cadaverine assay as described in Section 2.5.2.
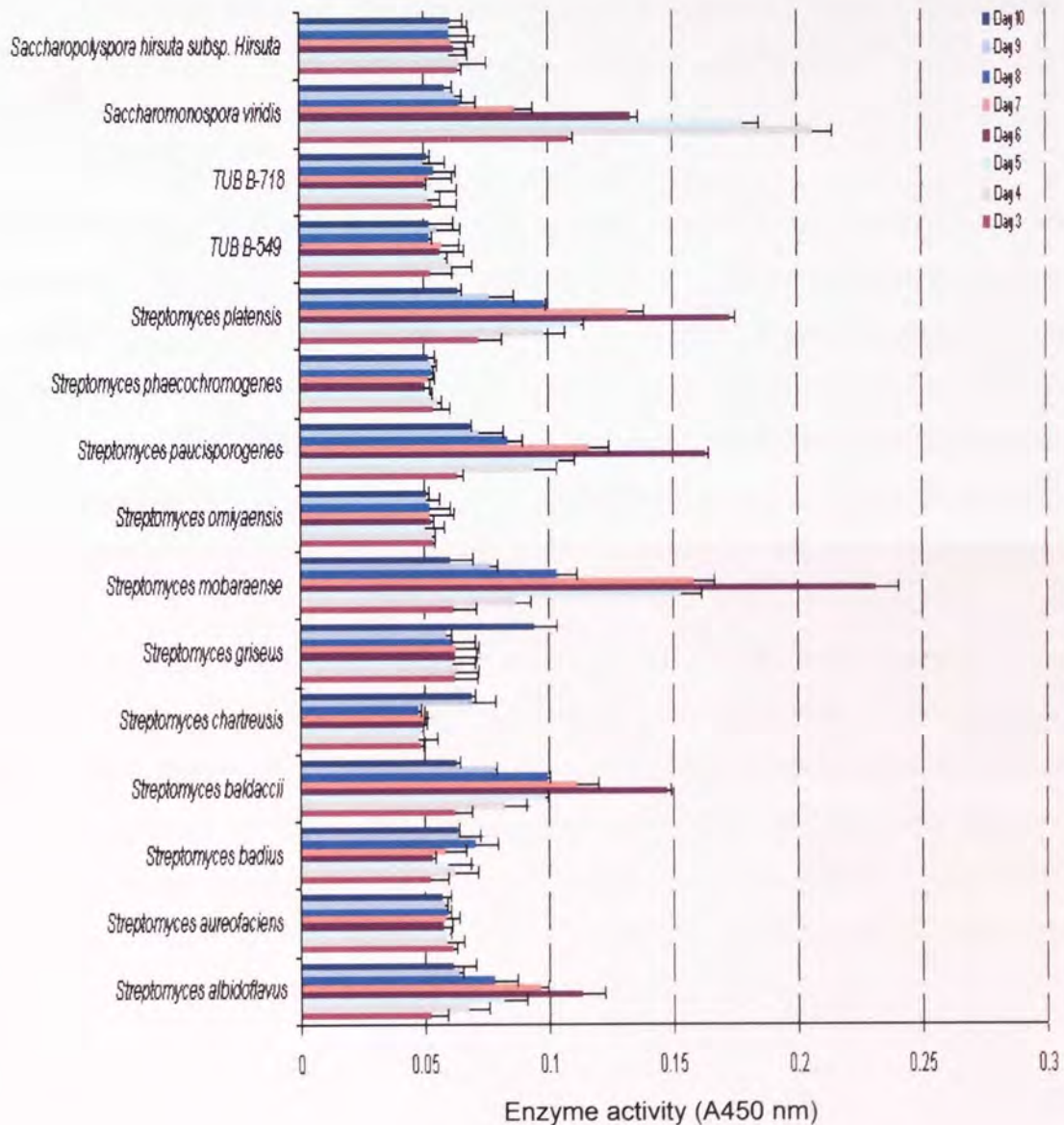
4.2.3.3. Identification of TGase from Bacillus sp. and Related Organisms.

The results of BLAST searches highlighted the possibility of identifying a TGase enzyme similar to those described in other *Bacillus sp.* (Figure 4.6). Organisms selected for study were revived from long term storage and sub-cultured three times on nutrient rich medium (LB) prior to growth, using Schaeffer's sporulation medium, for the purpose of enzymatic assays. In addition to members of the *Bacillus* family a number of other organisms, closely related to the *Bacillus* family (by 16SrRNA relationship) (Xu and Cote 2003), were assayed for activity also.

No activity was detected in organisms closely related to the *Bacillus* genera. However, noticeable activity was detected in a number of related *Bacillus* species (Figure 4.6). Further analysis confirmed TGase activity to be attributed to the later stages of sporulation, as described by Kobayashi (Kobayashi, et al. 1998b). In each case enzyme activity (from crude lysates) was found to be susceptible to the competitive primary amine substrates, histamine and putrescine (Figure 4.7).

Initial purification attempts from the organism *B. subtilis* were carried out, to gauge available yield and recovery, according to the method described by Kobayashi (Kobayashi, et al. 1998b). Insufficient enzyme was recovered despite growth of 20 L *B. subtilis*. As such, the method was deemed impractical for purification of sufficient amounts of enzyme. An alternative strategy, involving the cloning and expression of Bacillus TGase, was adopted in an attempt to assess its enzymatic characteristics.

Figure 4.6; Detection of TGase in Bacillus sp. and Attribution to Sporulation

Bacillus sp. were grown (in 500ml) in Schaeffer's sporulation medium to the relevant growth stage. Culture broth was sampled at periodic intervals and assessed by spectrophotometry as described in Section 2.4.5, enabling a growth curve to be constructed (A). The growth curve of each organism was used as indication of growth stage for activity assays. Organisms were grown in 250 ml Schaeffer's medium per timepoint for the relevant time period at 37 °C and 225 rpm after which, cells were harvested by centrifugation at 15,000 x g for 15 mins at 4°C before being lysed and assayed for TGase activity (biotin cadaverine assay) as described in Sections 2.4.7.1 and 2.5.2 respectively (B).

Data are reported as the mean + S.D. of samples assayed in triplicate. Shown below is a representative of 3 independent experiments.
* denotes asporogenic strain of *B. megaterium*.

(A)
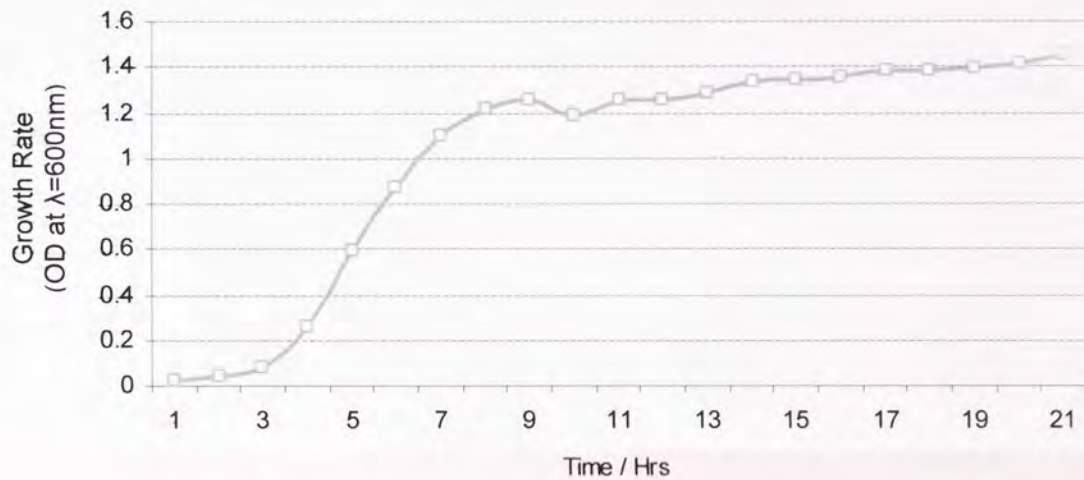
(B)
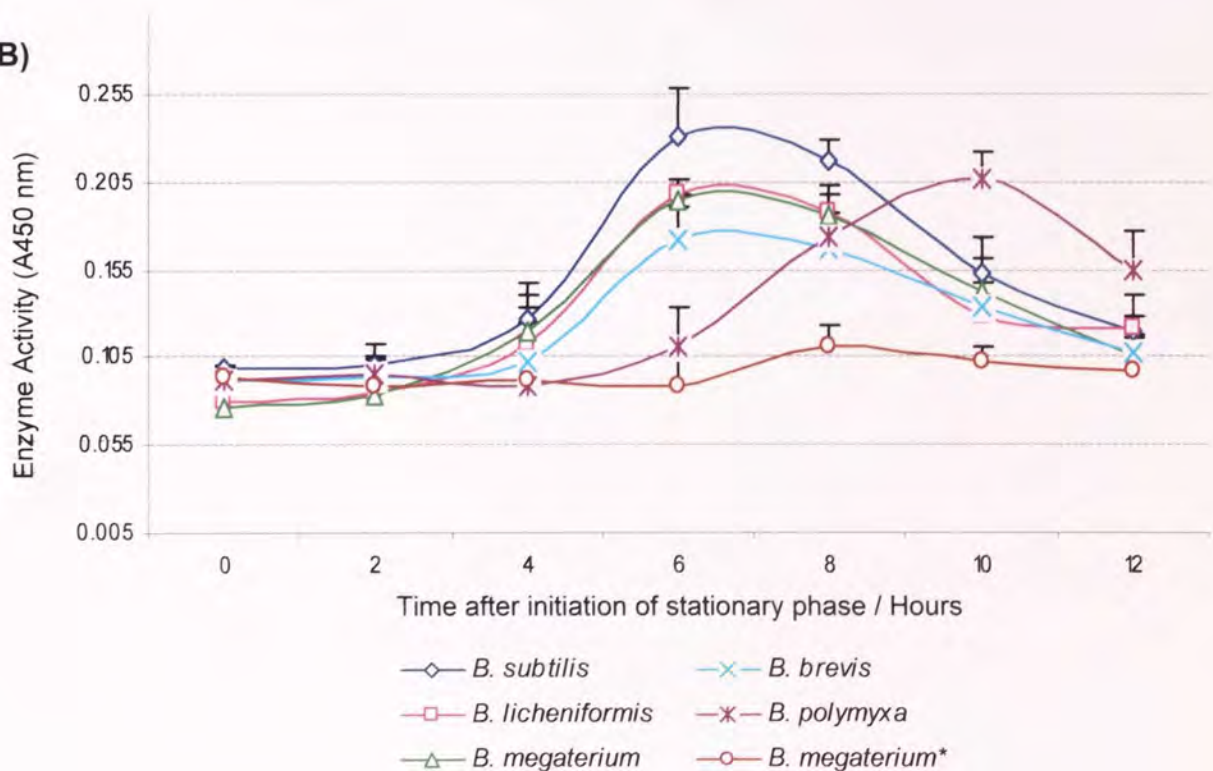
Figure 4.7; Effect of Inhibitors on TGase from B. subtilis

B. subtilis was grown in 500 ml Schaeffer's medium to 6 hours post sporulation initiation whereby cells were collected by centrifugation at 15,000 x g for 15 min prior to lysis as described in Section 2.4.7.1. The lysates (20 µl) generated were assayed for TGase activity in the presence of the competitive primary amines histamine (2mM), putrescine (2mM) and the thiol inhibitor N-ethylmaleimide (2mM) (NEM) using the biotin cadaverine assay as described in Section 2.5.2. Graph depicts mean values + S.D from samples assayed in triplicate from three independent experiments.
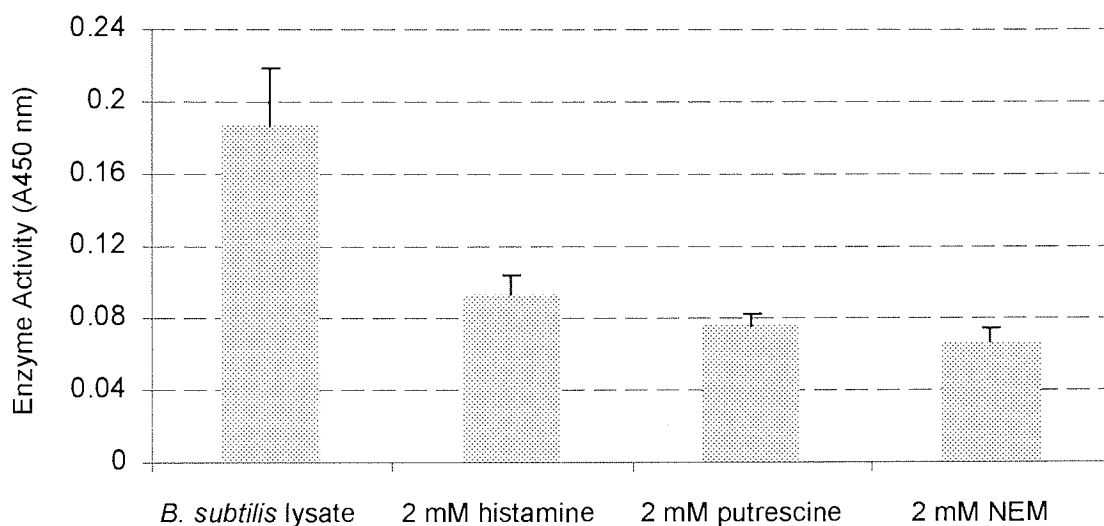
4.2.3.4. Molecular Cloning of TGase from B. subtilis

Due to the low amounts of TGase present in a number of *Bacillus* strains, (Section 4.2.3.3), a molecular approach was adopted with the aim of producing recombinant TGase originating from *Bacillus* sp.. Primers were designed, based on the published sequence of TGase from *B. subtilis* (Kobayashi, et al. 1998a) to enable cloning into a number of vectors. Vectors used allow both heterologous and homologous expression of proteins that were targeted to different cellular compartments. The primers were used to amplify the 735 bp open reading frame by PCR (Section 2.11.6) (Figure 4.8) which was ligated into pGEM-T (Promega, UK). The gene encoding TGase was then subcloned into the plasmid of choice (Figure 4.9).

Attempts at transforming the *B. megaterium* protoplasts (MoBiTec, Germany) were unsuccessful despite multiple attempts. *Bacillus* TGase was cloned into the vectors pET21d, pET22b and pGEX1λT for expression in *E. coli*. Expression, in *E.coli,* was only observed using the expression plasmid pET22BTG2, highlighting an expressed protein corresponding to the published molecular weight of *Bacillus subtilis* TGase (28.3 kDa) (Figure 4.10). Since the primers were designed to incorporate the poly histidine tag encoded by the vector recombinant protein was purified by Ni^{2+} affinity (Section 2.11.15). Protein corresponding to the correct molecular weight was purified (Figure 4.11); however, no enzyme activity could be detected

Due to the inherent problems of expressing *Bacillus* TGase subsequent screening of the genera was halted in favour of screening other micro-organisms for the presence of TGase activity.

Figure 4.8; PCR Amplification and Cloning of B. subtilis TGase

B. subtilis TGase was amplified by PCR as described in Section 2.11.6 using primers Rlp_1 and Rlp_2 and the following reaction conditions: 94°C for 4 min followed by 30 cycles of 94°C for 1min, 58°C for 1 min and 72°C for 1 min (A). The amplified DNA was analysed by agarose gel electrophoresis with ethidium bromide staining as described in Section 2.11.4, with the band corresponding to the correct size (≈750 bp) being extracted and ligated into pET22b(+)as outlined in Sections 2.11.7 and 2.11.9. respectively. The resulting plasmid, pET22BTG2, was restricted with *MscI* and *XhoI* to release *B. subtilis* TGase (B).

(A)
(B)
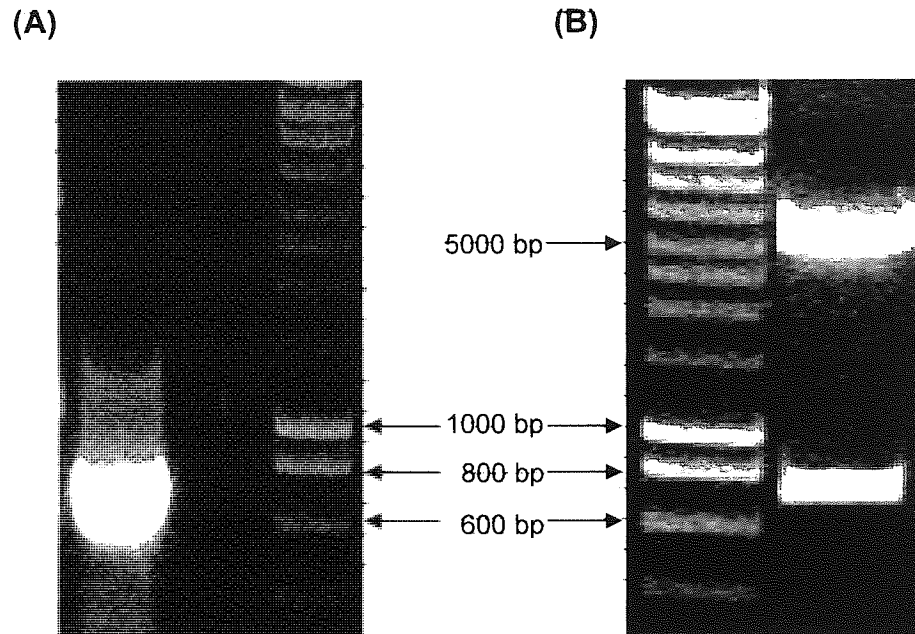
Figure 4.9; Molecular Cloning of B. subtilis TGase into Vectors pET21d, pGEX1λT and pWH1520.

B. subtilis TGase was amplified by PCR as described in Section 2.11.6 using the vector specific primers detailed in Table 2.3. with annealing temperatures ranging from 56- 58 °C. The amplified DNA was subcloned into pGEM-T allowing enzymatic digestion using the designed restriction sites. The gene of interest was ligated into the corresponding vector as described in Section 2.11.9. After transformation of the ligated vector plasmids were recovered and the gene of interest excised by enzymatic digestion using the relevant enzymes. The fragments were analysed by agarose electrophoresis, with ethidium bromide staining, to confirm the presence of the gene of interest in vector pET21d(+) (A), pGEX1λT (B) and pWH1520 (C).

(A)

6,000 bp —
5,000 bp —
                  ◄—Linearised pET21d

1,000 bp —
750 bp —   ◄—B. subtilis TGase
500 bp —

(B)

6,000 bp —
5,000 bp —   ◄—Linearised pGEX1λT

1,000 bp —
750 bp —
500 bp —   ◄— B. subtilis TGase

(C)

1,000 bp —►
750 bp —►   ◄— B. subtilis TGase
500 bp —►

118
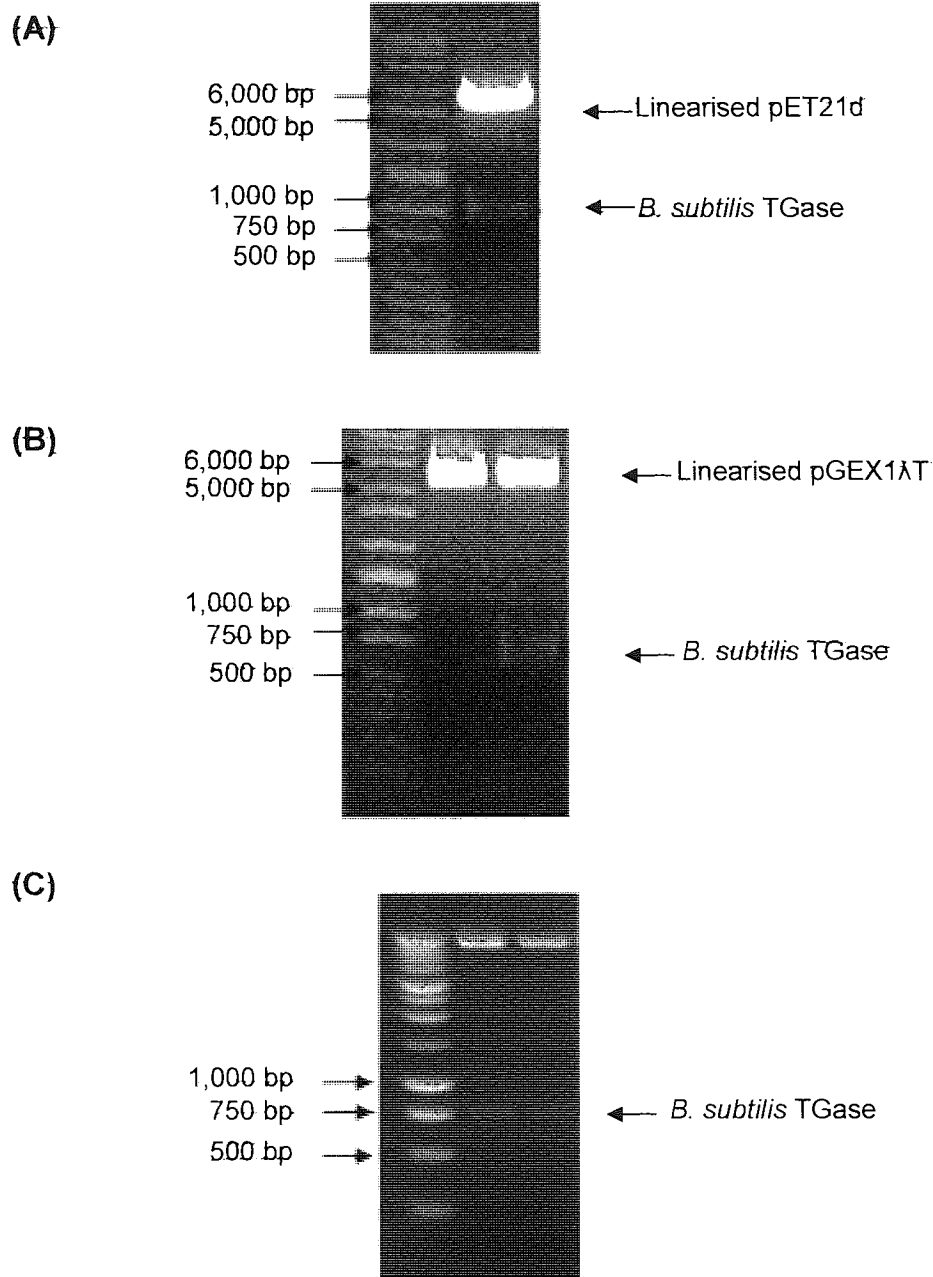
Figure 4.10; SDS-PAGE Analysis of Expressed Protein (B. subtilis TGase)

E.coli tuner cells harboring either pET22b(+) or pET22BTG2 were grown to an absorbance 0.6 (A600 nm). Prior to induction with IPTG an aliquot was taken and used as the uninduced control (U). Samples of induced cultures were taken at 1, 2 and 4 hours post induction and analysed by SDS-PAGE.
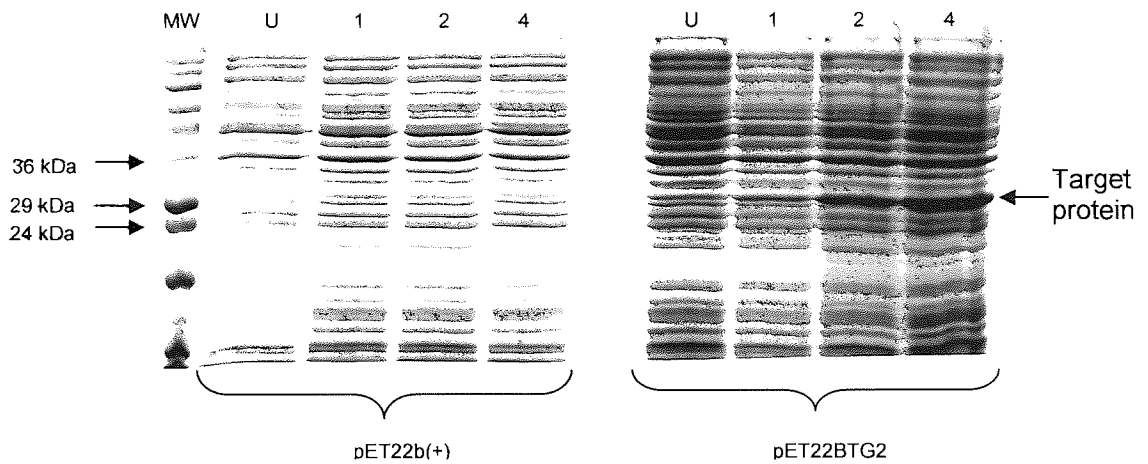
Figure 4.11; Purification of Histidine Tagged Bacillus TGase

E. coli tuner (Novagen, UK) transformed with pET22BTG2 was grown to an A_{600nm} of 0.75, in the presence of 75 µg/ml ampicillin, prior to induction by the addition of IPTG to a final concentration of 1 mM. Cells were incubated for a further 4 hours at 28°C at 200 rpm. Cells were harvested by centrifugation at 13,000 rpm for 10 min and treated with Bugbuster protein extraction reagent (Novagen) according to the manufacturers' instructions. Recombinant histidine tagged protein was purified according to the manufacturers' protocol.

L designates column load, FT; column flow through; W; Wash and E; Elution
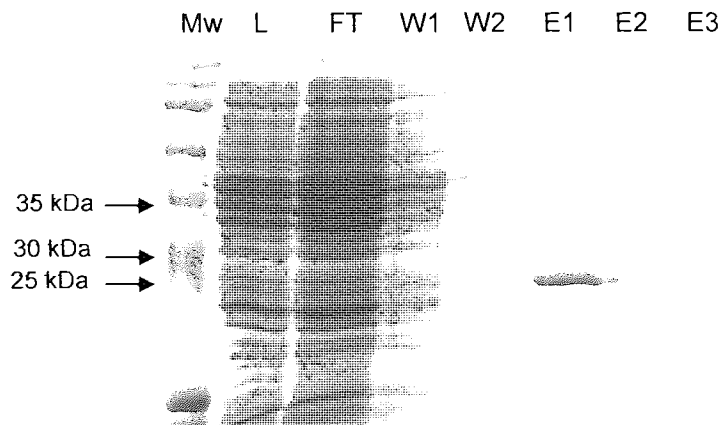
4.2.3.5. Screening of Other Micro-Organisms for TGase Activity

Other bacteria were screened for TGase activity. Culture supernatants and cell lysates from different time points were generated and analysed (Figure 4.12). Reproducible activity was observed from the cell lysates of the organisms *Pseudomonas putida* NCIB 9872 and *Ps. putida* NCIB 9876. No other organisms screened displayed significant activity using the activity assays described. Analyses of other *Pseudomonas sp.* highlighted that although activity appeared to be restricted to species of *Ps. putida*, activity was not detected in all *Ps. putida* isolates analysed (Figure 4.12). Apparent TGase activity appeared, in each case, to be related to cellular growth since maximal activity was detected in the mid-late logarithmic stage lysate.

Figure 4.12; Screening other Micro-Organisms for TGase Activity

Initially growth investigations were performed, as described in Section 2.4.5 to determine individual growth characteristics. Micro-organisms were inoculated into 500 ml medium in a 1 L conical flask, per growth stage analysed (see Appendix IIIc for relevant medium and growth conditions), and grown at the required temperature, indicated in Appendix III, with shaking at 200 rpm. Cultures were terminated at the early logarithmic, mid-late logarithmic and stationary phases of growth. Cells were collected by centrifugation at 13,000 rpm for 15 min at 4°C with an aliquot of cell free culture broth retained for assessment. Cells collected were lysed as described in Section 2.4.7.1 to generate the crude lysates. The crude lysates (20 µl) were analysed in triplicate for TGase activity using the biotin cadaverine assay as described in Section 2.5.2.
Pseudomonas strains boxed.
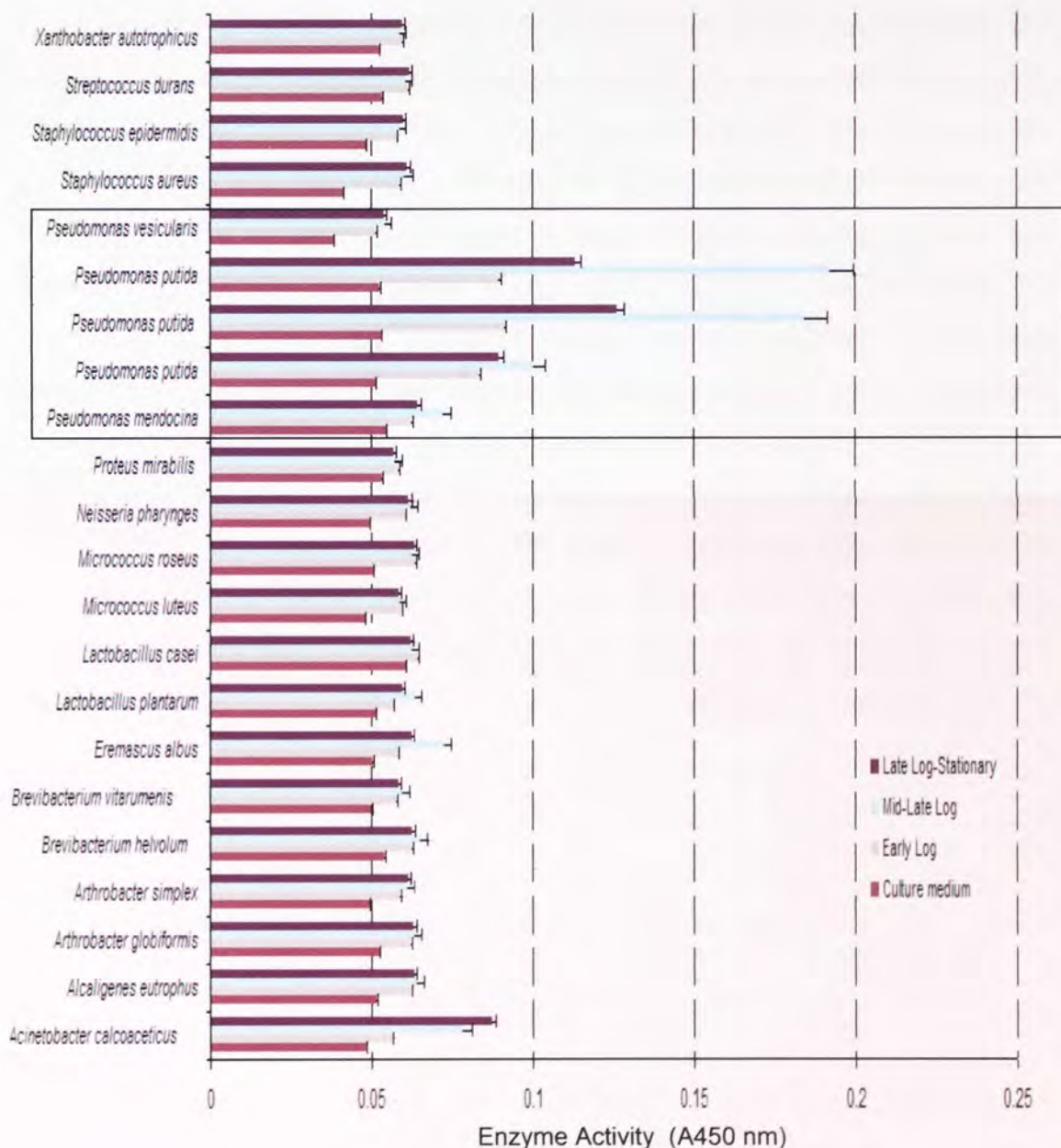
4.2.3.6. Identification of Potential TGase from Pseudomonas putida NCIB 9872

Noticeable TGase activity was detected in two out of three species of *Pseudomonas putida* screened, with most activity observed in the soluble fraction of cell lysates generated from strain NCIB 9872, during mid-late logarithmic growth. This organism was chosen for further analysis with a view to purification of the responsible entity.

In an attempt to confirm the presence of a TGase in *Ps. putida* 9872, the lysate, from which activity was easily detected, was subjected to inhibition with varying concentrations of the competitive primary amines: histamine and putrescine. As can be seen in Figure 4.13, no appreciable activity was observed in the resuspended insoluble lysate fractions confirming that the protein responsible was soluble in the lysis reagent used. Interestingly, instead of down-regulating apparent TGase activity, enzyme activity was increased in a dose-dependent manner by the presence of both competitive TGase substrates (Figure 4.13). Upon comparison putrescine was found to induce a more profound increase in response as opposed to histamine. Apparent activity was increased 1.66 and 2.08 fold in the presence of histamine (2 mM and 4 mM respectively). In the presence of putrescine apparent activity was increased 2.03 and 2.32 fold (2 mM and 4 mM respectively).

Figure 4.13; Effect of Primary Amine Substrates Apparent TGase Activity Derived from Cell Lysates of Ps. putida *NCIB 9872*

Ps. putida was grown in 1 L LB medium at 30 °C with shaking at 200 rpm for 6 hours (mid/late logarithmic growth stage). Cells were collected by centrifugation at 13,000 rpm for 10 min prior to lysis as described in Section 2.4.7.1. Lysates generated were separated by centrifugation, at 13,000rpm for 20 min, into soluble and insoluble fractions. The insoluble fraction was resuspended in 10 mM Tris-HCl pH 7.4 to an equivalent concentration to the soluble fraction and the TGase activity (by biotin cadaverine incorporation) of each was assessed in the presence of 2 mM and 4 mM histamine or putrescine using the peptide incorporation assay as described in Section 2.5.3. Data presented reflects mean values + S.D. from samples assayed in triplicate. Below is representative of 3 independent experiments.
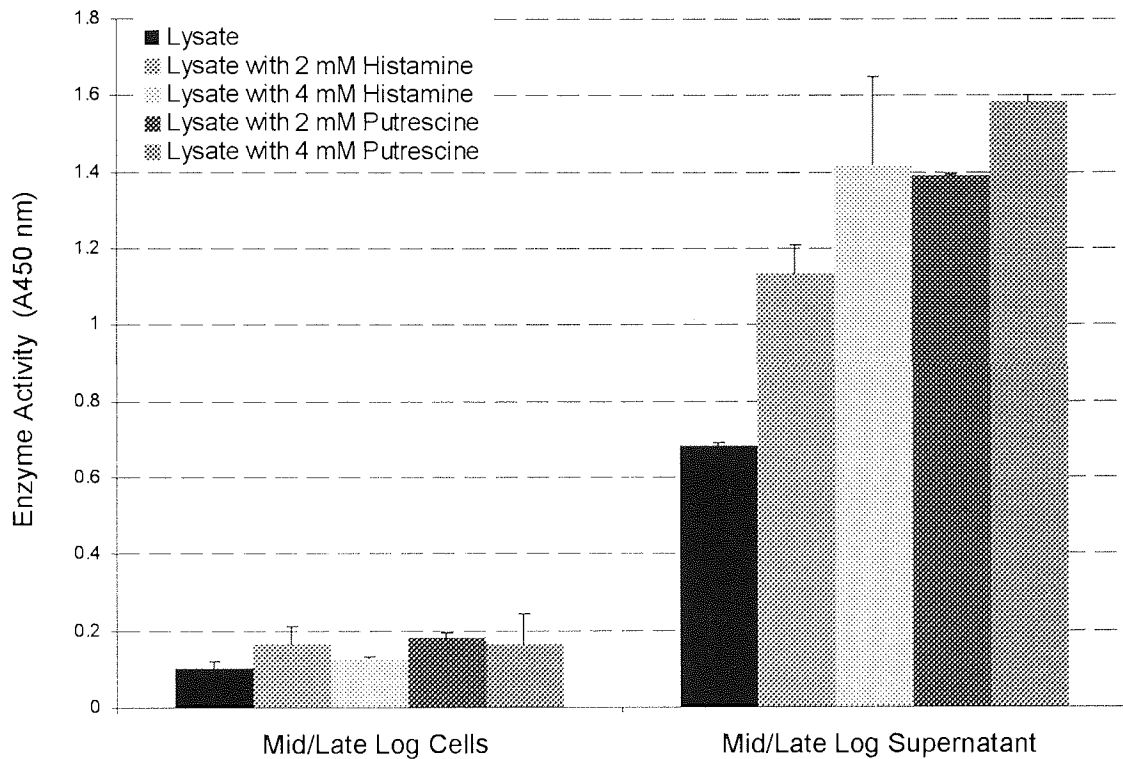
4.2.3.7. Partial Purification of Potential TGase from Ps. putida *NCIB 9872*

The reactivity of an enzyme present in the soluble lysate in both the biotin cadaverine and peptide cross-linking assays would appear to have highlighted the potential presence of a TGase-like enzyme in *Ps. putida* NCIB 9872, although in light of previous findings this waits to be substantiated. With a view to confirmation, it was envisaged that the protein responsible would be partially purified by a combination of ammonium sulphate precipitation and gel filtration (Figure 4.14). Notable enzyme activity was detected from the 75% (w/v) ammonium sulphate precipitation, activity that was reduced by the presence of the synthetic inhibitor R281. Additionally, activity was also reduced by the presence of 2 mM histamine.

The lyophilized lysate, generated from the partial purification procedure, was analyzed for the presence of $\varepsilon(\gamma$-glutamyl) cross links (as described in Section 2.10.1 and 2.10.11 respectively) as an indication of inherent TGase activity. The amount of cross-link formed was 3.2 nmol crosslink/mg protein indicating the potential presence of a TGase.

Figure 4.14; Ammonium Sulphate Precipitation of Enzyme Possessing Potential TGase Activity from Ps. putida *NCIB 9872*

Ps. putida NCIB 9872 grown in 1L LB medium at 30 °C with shaking at 200 rpm to the mid/late logarithmic growth stage. Cells were collected by centrifugation and lysed as described in Section 2.4.7.1. Insoluble material was removed by centrifugation at 13,000rpm for 20 min with the soluble supernatant used as the starting material for the ensuing purification. Protein was precipitated by the addition of solid ammonium sulphate to a final concentration of 25, 50 75% (w/v)and saturation (indicated as 100%), Protein was precipitated on ice for 2 hours with gentle agitation before being collected by centrifugation (13,000 rpm for 20 min). Precipitated protein was dialysed exhaustively against PBS and recovered to an equivalent concentration prior to assessment of TGase activity using the biotin cadaverine assay in the presence of 250 µM R281 and 2 mM histamine (Section 2.5.2).

Lysate was assayed in triplicate on three separate occasions. Error bars represent standard deviation.
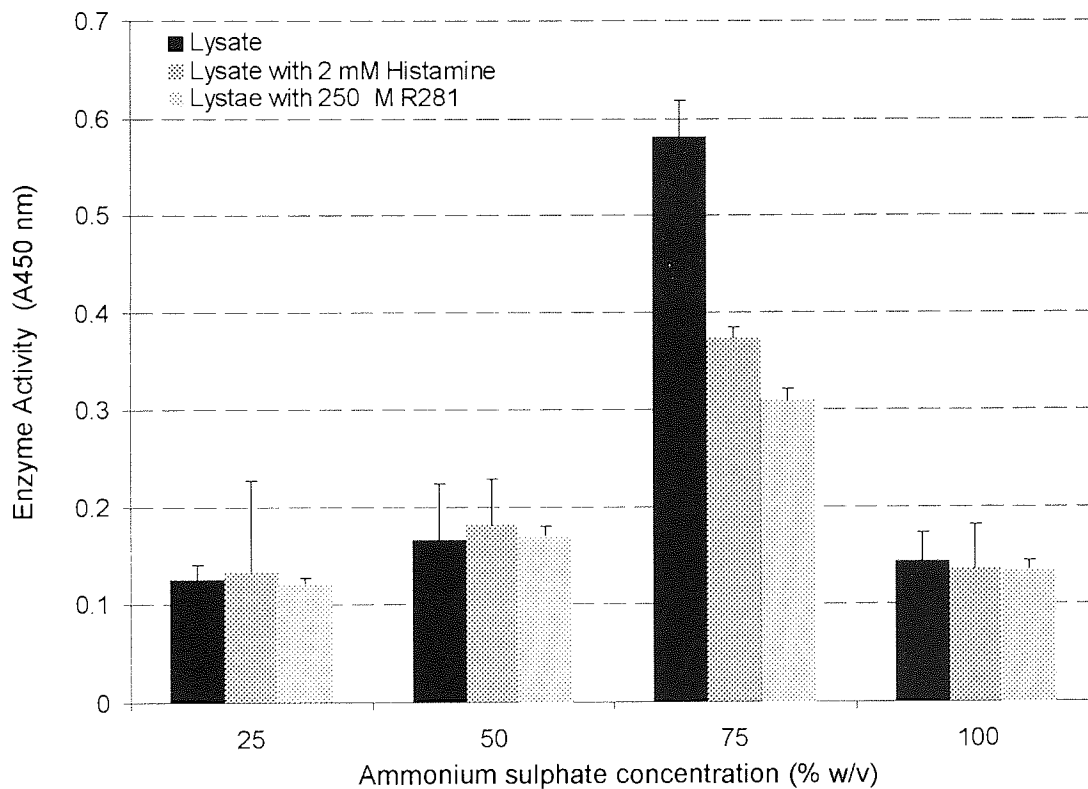
To elucidate an apparent size for this potential TGase enzyme the protein obtained from ammonium sulphate precipitation was separated by gel filtration with the resulting fractions assayed for both protein content and TGase enzyme activity (Figure 4.15). Although a crude approach it was thought that this would help elucidate the apparent molecular weight of the apparent TGase enzyme. Proteins ranging from 25-75 kDa in size were visualised, though none cross-reacted with antibodies to microbial and eukaryotic TGase as determined by Western blotting. Enzyme activity was determined by the incorporation of radio-labelled putrescine into N,N'-dimethylcasein in concert with the peptide cross-linking assay, reactions were allowed to proceed for 30 min and 1 hr. An increase in activity was detected using the peptide cross-linking assay; however no increase in activity was observed using the radio-labelled putrescine assay despite the fact that inhibition of potential activity occurred using R281 (Figure 4.15). The lack of sensitivity to putrescine could perhaps be a function of an altered substrate specificity.

Figure 4.15; Separation and Analysis of Ps. putida *NCIB 9872 Lysate by*

Gel Filtration

Cell lysate of *Ps. putida* was generated as described in Section 2.4.7.1 from 2.5 L culture harvested at mid/late logarithmic growth phase. Protein was initially precipitated by the addition of ammonium sulphate to 50% (w/v) which after collection by centrifugation was discarded. Remaining protein was precipitated by the addition of ammonium sulphate to saturation. Protein was collected by centrifugation (15,000 x g for 25 min). The protein pellet was resuspended in gel filtration running buffer and separated by gel filtration using Sephacryl S-200 resin (500ml resin at a flow rate of 5ml/min). Protein, collected at 1 min intervals, was analysed for protein content and TGase activity by the biotin cadaverine assay as described in Section 2.5.2 (A). Active fractions were pooled, lyophilized and visualised by SDS-PAGE using 10% acrylamide gels (B). Activity of the lyophilised lysate was assessed by both the peptide crosslinking (Ci) and [^{14}C] putrescine assays (Cii) as described in Sections 2.5.3 and 2.5.4 respectively. Assays were repeated in triplicate + standard error. Graphs are representative of at least 2 independent experiments. MwM = molecular weight markers.

(A) Gel Filtration
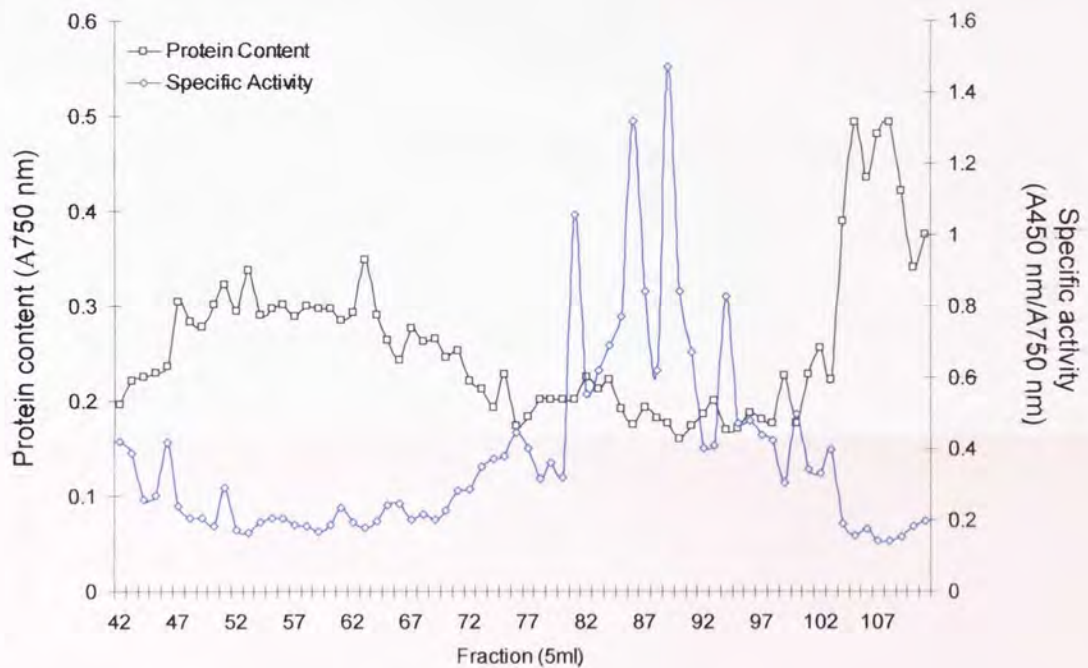
(B) SDS-PAGE
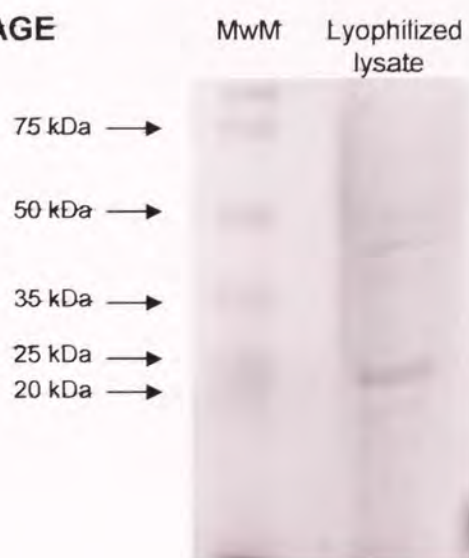
(Ci) Peptide Cross-Linking Assay
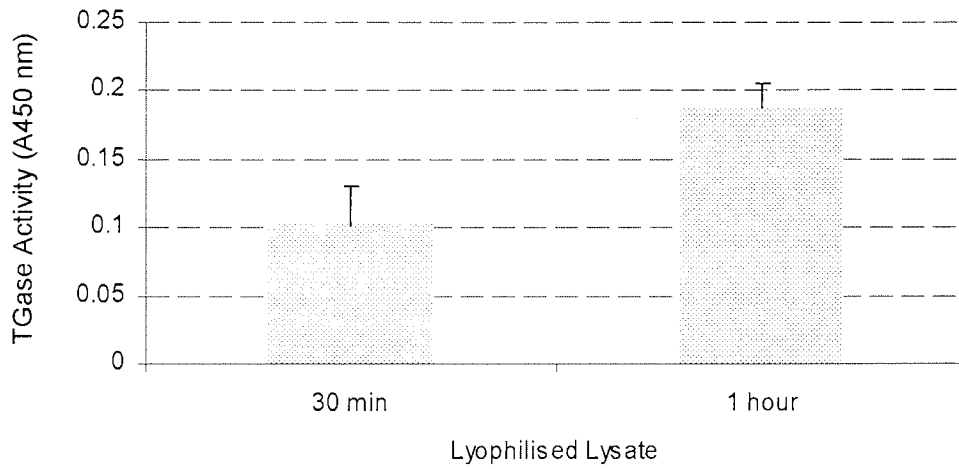
(Cii)[C^{14}] Putrescine Incorporation into N,N'-dimethylcasein
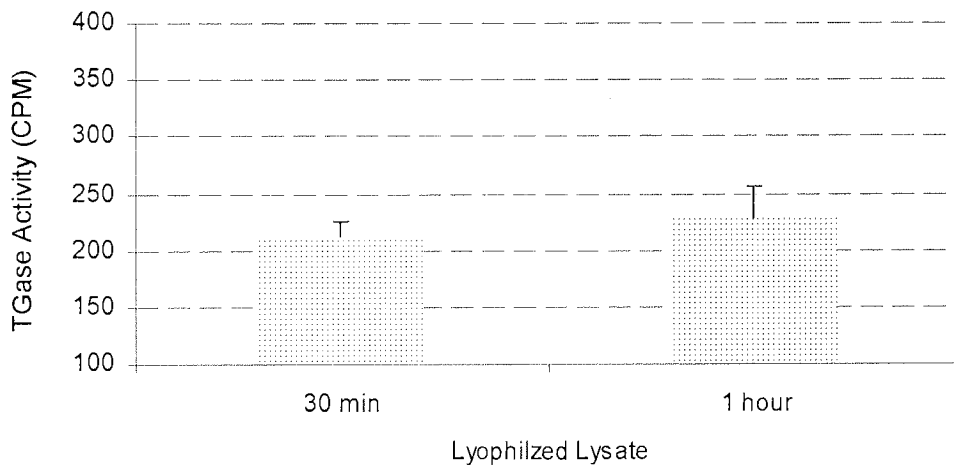
4.2.3.8. *Further Purification of Potential TGase from* Ps. putida *NCIB 9872*

Building on the work carried out previously, additional enzyme purification was undertaken.

The starting material was precipitated, initially by the addition ammonium sulphate to 50% (w/v) with the resulting pellet being discarded. Remaining protein was precipitated in the presence of 100% ammonium sulphate, and dialysed against PBS. An activity assay, post exhaustive dialysis, ensured the enzymes presence in the saturated ammonium sulphate fraction (Figure 4.16A).

Dialysed protein, derived from ammonium sulphate precipitation, was then separated by anion exchange chromatography using a Q sepharose column (Amersham Bioscience, UK). Resulting fractions were assessed for both protein content and enzyme specific activity (Figure 4.16B). Fractions of high specific activity were pooled and separated by gel filtration using Sephacryl S-200 resin (Figure 4.16C). Fractions encompassing the active peak were subsequently analysed by SDS-PAGE (Figure 4.16D) and found to comprise proteins from 20-70 kDa in accordance with work carried out previously whereby the protein exhibiting TGase activity was deemed to be of similar molecular weight. Further purification steps were abandoned due to loss of enzyme activity.

Figure 4.16; Purification of Potential TGase from Ps. putida NCIB 9872

The cell free lysate of mid-late exponentially grown *Ps. putida* (2 L) was generated by three passages through a French press; an aliquot was cultured on LB agar to ensure complete cell lysis. The cell free lysate generated was successively precipitated by the addition of 50 and 100% (w/v) (saturated) ammonium sulphate. Precipitates were collected by centrifugation (15,000 x g for 30 min) and resuspended in anion exchange running buffer. The TGase activity of each precipitation step was confirmed using the biotin cadaverine assay as described in Section 2.5.2 (A). The lysate containing TGase activity was separated by anion exchange chromatography using Q-sepharose (B) with the active fractions being pooled and separated by gel filtration using Sephacryl S-200 (C). After each chromatography step fractions were assayed for both protein content and TGase activity by biotin cadaverine incorporation (Sections 2.9.1 and 2.5.2). Fractions encompassing the TGase active peak (20 µg protein loaded) were analysed by SDS-PAGE using 10% acrylamide gels and stained with Coomassie brilliant blue (D). Bars represent mean values + standard error.

(A) Ammonium Sulphate Precipitation

Legend:
- No Ca2+, DTT, R281
- with 5 mM Ca2+ and DTT
- with 5 mM Ca2+, DTT and 500 mM R281
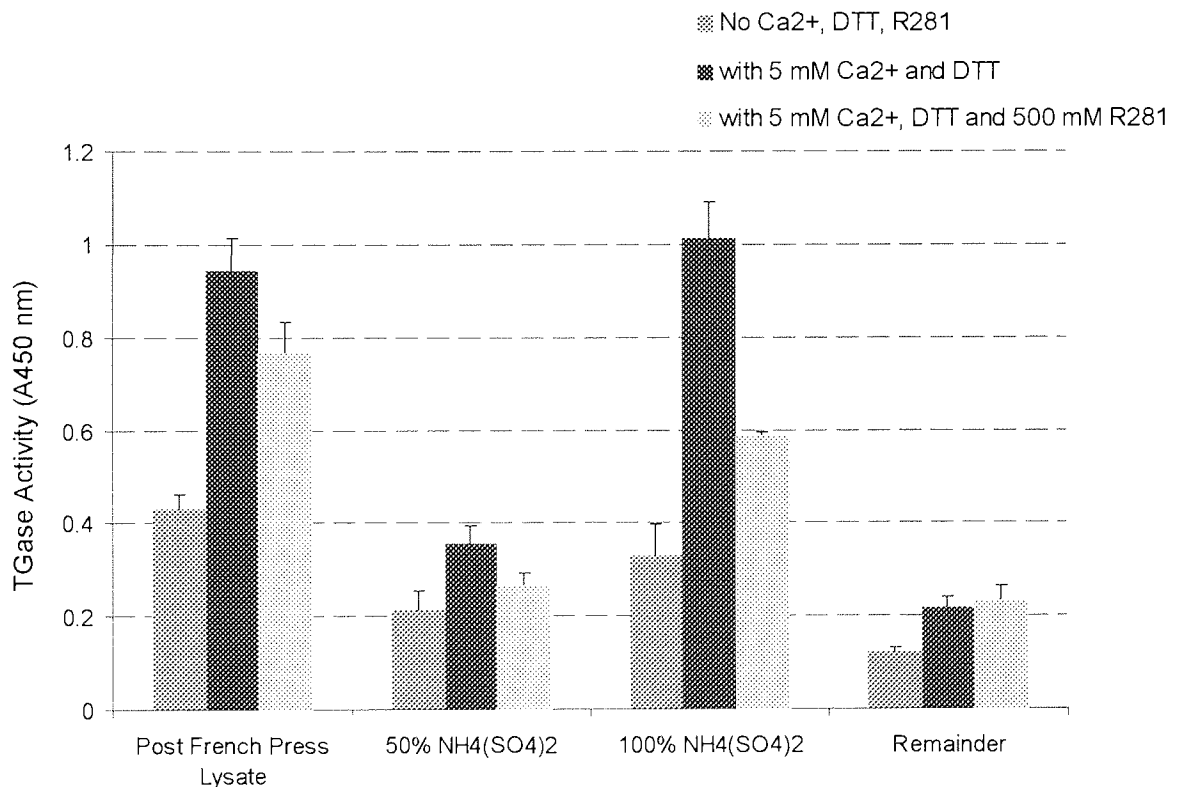
(B) Anion Exchange Chromatography (Q Sepharose)
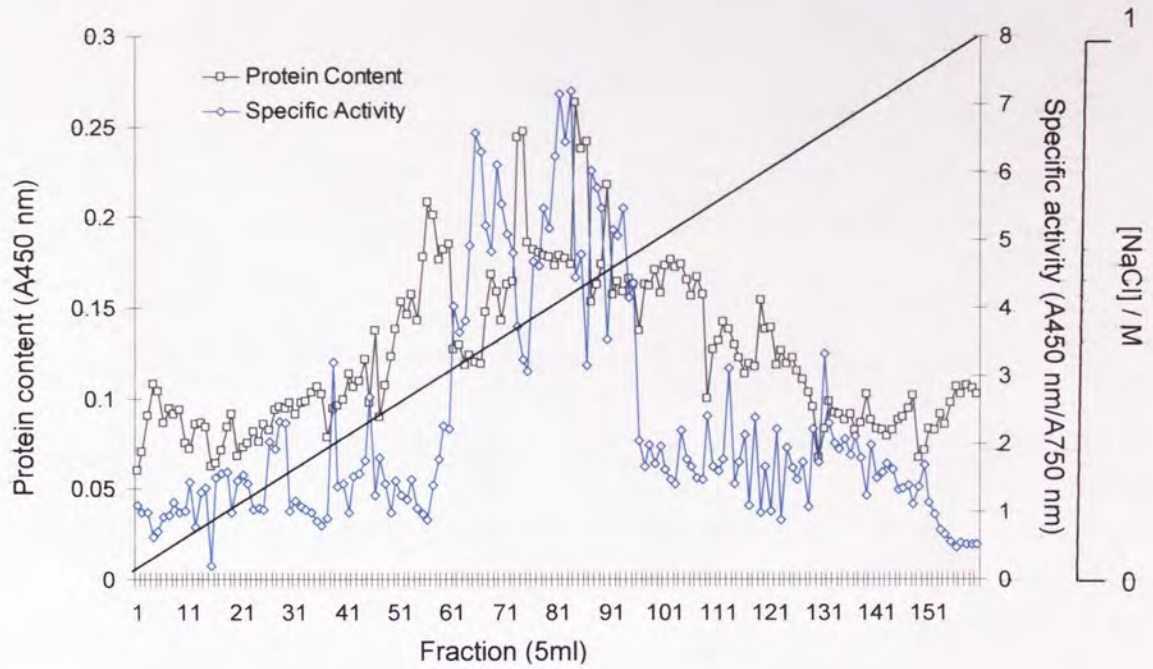
(C) Gel Filtration (Sephacryl S-200)

(D) SDS-PAGE Analysis

Fraction Number

MwM 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90

75 kDa →
50 kDa →
35 kDa →
25 kDa →
20 kDa →
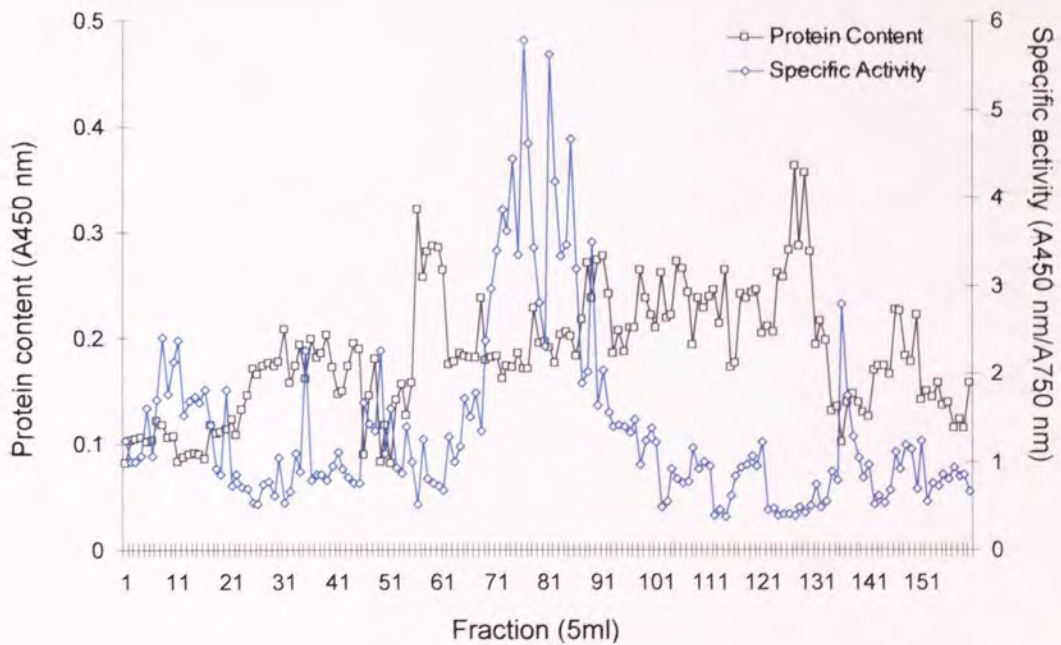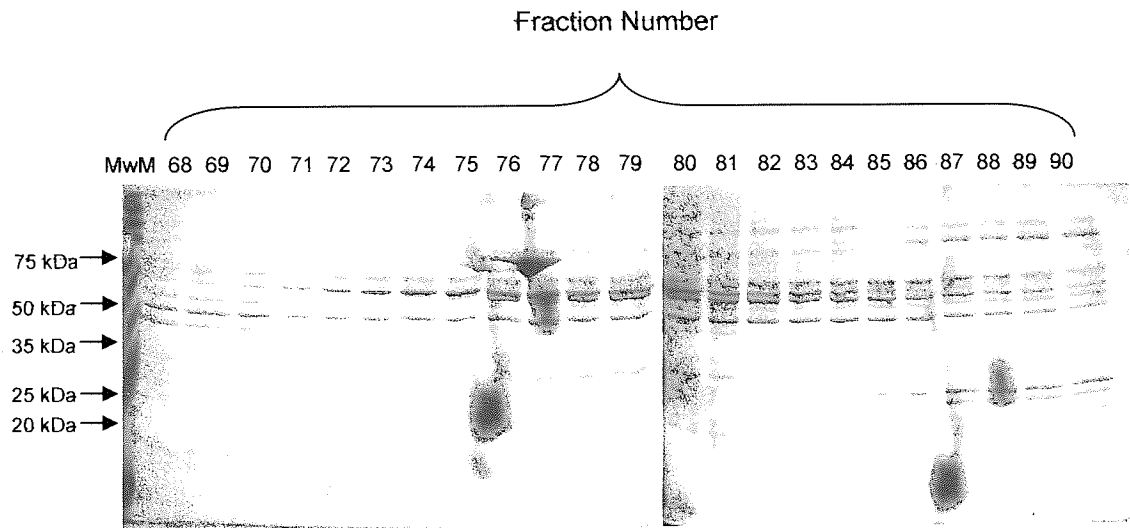
4.2.3.9. *Affinity Purification of TGase from* Ps. putida *NCIB 9872 Using R281-Sepharose*

The site directed TGase irreversible inhibitor R281 has been coupled to EAH Sepharose 4B (Amersham Bioscience, UK) to create an affinity resin capable of covalently binding TGase protein (Figure 4.17). This method may provide a useful step in the identification and purification of novel TGase proteins.

Figure 4.17; Production of R281 Affinity Resin

R281 coupled Sepharose was generated utilizing the free amino groups at the end of an 11-atom spacer arm present on EAH Sepharose. R281, which comprises a free carboxyl group, was linked to EAH Sepharose 4B following a two step carbodiimide coupling outlined below and described in Section 2.6. The N,N'-disubstituted carbodiimide promotes condensation between a free amino group (present on the resin) and a free carboxyl group (present on R281) to form a peptide link by acid catalyzed removal of water.

Step 1:

R281   N-(3-Dimethylaminopropyl)-   Active Ester
              N'-ethylcarbodiimide

Step 2:

Active Ester   EAH-Sepharose

R281 Coupled Resin

Urea Derivative
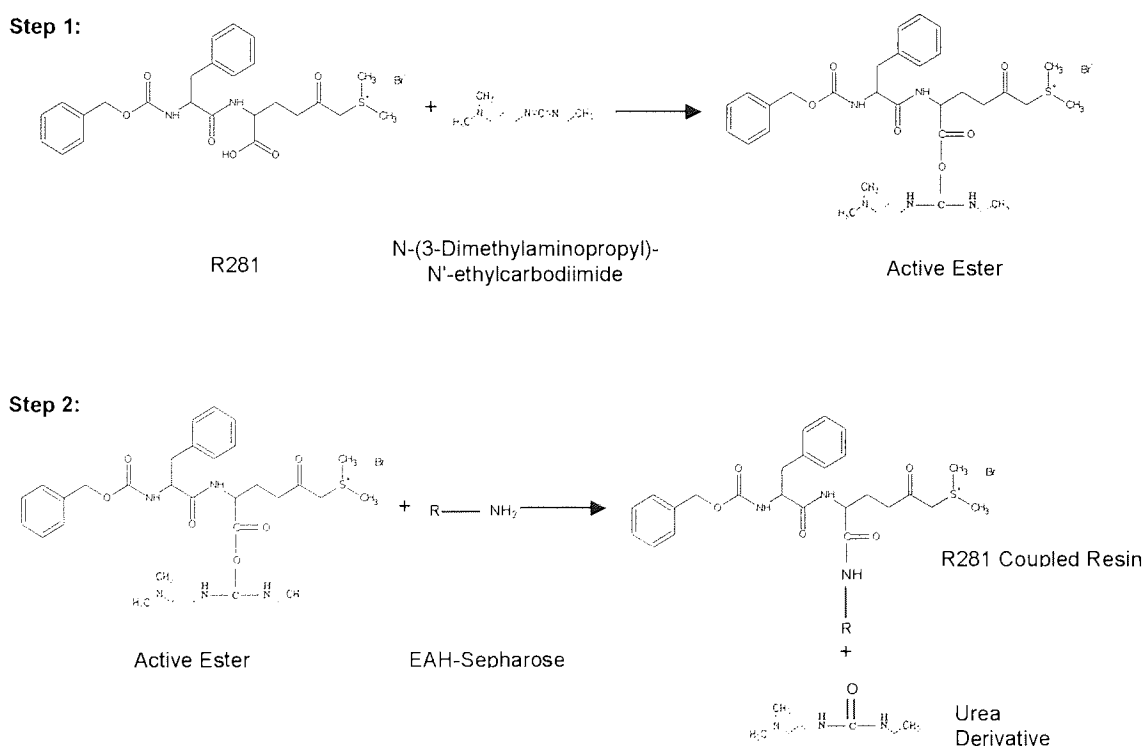
Results gathered from previous studies highlighted that 4 hours incubation to be adequate for effective mTGase binding to the affinity resin (Section 3.2.5). Using the method described (Section 2.6.1) and after SDS-PAGE analysis, a large proportion of protein failed to bind to the coupled resin, with loosely bound material being eluted during the salt wash (0-5M NaCl gradient)

(Figure 4.18). In the case of *Ps. putida* NCIB 9872, a large number of proteins bound to the resin (Figure 4.18). It would appear likely that instead of targeting solely the active site cysteine residue of TGase, that R281 may also target surface exposed cysteine residues. Additionally, it may be possible for the affinity of the inhibitor R281 to be altered by either coupling with the EAH sepharose or by reaction with N-(3-Dimethylaminopropyl)-N'-ethylcarbodiimide hydrochloride (EDC). Such non-specific binding limits the usefulness of the proposed application.

Figure 4.18; Analysis of Covalently Attached Ps. putida Proteins to R281- Sepharose.

The lysate, generated by French Press, from *Ps. putida*, thought to contain a potential TGase was generated from a 500 ml culture grown to mid-late exponential phase. The resulting cell free lysate, separated by centrifugation at 15,000 x g for 10 min, was incubated with R281 coupled EAH Sepharose for 4 hours (rotated end over end at 4°C) and washed with a 0-5 M sodium chloride gradient as described in Section 2.6.1. Irreversibly bound protein was removed by digestion with 20 ng trypsin incubated for 16 hours at 37 °C and visualised by SDS-PAGE using 10% acrylamide gels with subsequent coomassie blue staining.

Lane 1: cell lysate; lane 2: unbound protein; lanes 3-16 are representative fractions from the 0-5.5M salt gradient; lane 17: trypsinised protein.

Fraction Number

MwM 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 MwM

205kDa
116kDa
97kDa
84kDa
66kDa
55kDa
45kDa
36kDa
29kDa
24kDa
20kDa
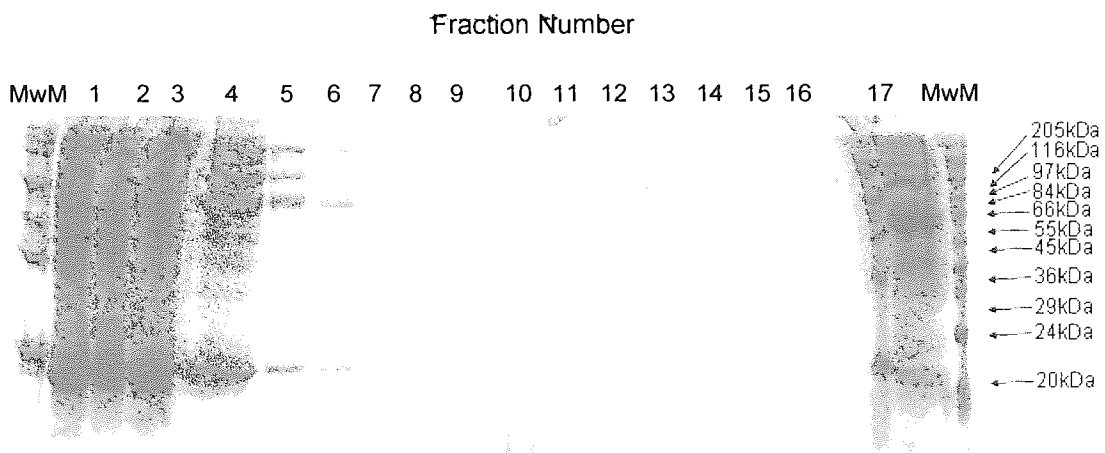
4.2.3.10. *Cloning of BTLCP from Ps. putida NCIB 9872*

Having highlighted a potential TGase protein arising in *Ps. putida* extensive database searches were carried out to identify similar proteins. Ginalski have highlighted a number of organisms predicted to be bacterial transglutaminase-like cysteine proteinases (Ginalski, et al. 2004). One organism highlighted as encoding such a protein was *Ps. putida* (Accession Number; NP_742333). Searching the non-redundant database revealed no sequences resembling the protein in question. Due to purification difficulties surrounding the protein responsible, a molecular approach was adopted to potentially identify and clone the relevant region of DNA with a view to confirming TGase activity.

The completed genome of *Ps. putida* was searched for the presence of a TGase-like protein. From the publicly accessible sequence (Accession number: AAN68294), primers were designed (Table 2.3) to enable cloning of gene identified, based on conceptual translation of the completed genome sequence from *Ps. putida* (Nelson 2002). The genetic sequence was amplified by PCR with subsequent enzyme restriction allowing ligation into pET22b(+) (Novagen, UK) (Figure 4.19A and B, respectively). The insert was confirmed by DNA sequencing (Appendix V) and expressed via IPTG induction. No expressed protein, corresponding to the correct molecular weight, was observed by SDS-PAGE. In addition, no TGase activity was detected.

Figure 4.19; Amplification and Cloning of TGase-like DNA from Ps. putida 9872

TGase-like protein was amplified by PCR using primer Ps_1 and Ps_2 (Table 2.3). In the presence of 2 mM $MgCl_2$ reaction conditions were as follows; 94°C for 4 min followed by 30 cycles of 94°C for 30s, 56°C for 30s and 72°C for 1 min. The fragment was subcloned into pGEM-T and ligated into pET22b(+) according the method outlined in Section 2.11.9. The presence of the gene of interest was confirmed by restriction analysis, with *Xho*I and *MSc*I, and subsequent agarose gel electrophoresis and ethidium bromide staining.
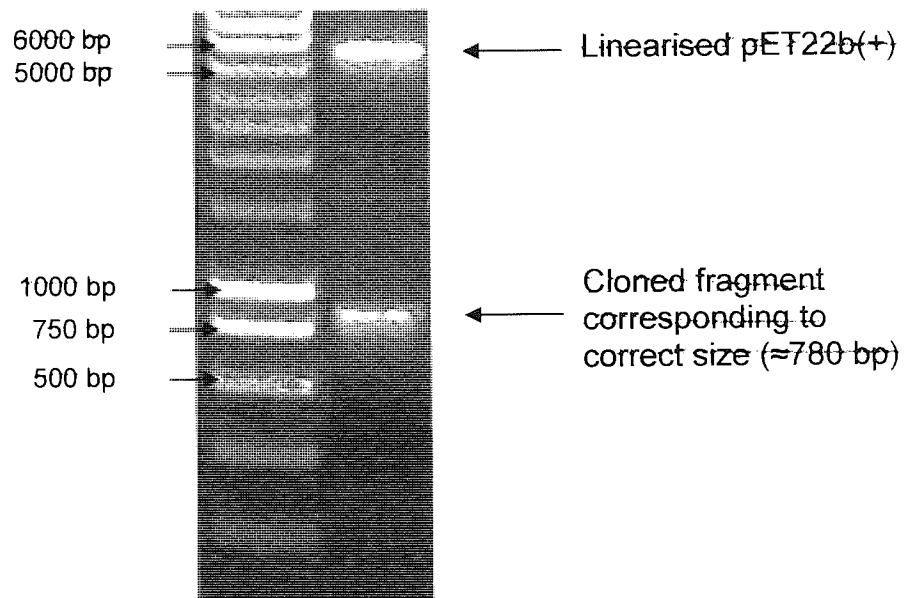
4.2.3.11. *Degenerate PCR to Identify Homologous Regions to* Streptomyces *sp. TGases*

Degenerate primers were designed based on highly conserved flanking regions of the active site of *Streptomyces sp.* TGases (Figure 4.20). The results of PCR using genomic DNA purified from *Ps. putida* NCIB 9872 are shown in Figure 4.21. Primer combination F1+R1 gave a weak band of the correct size (≈450 bp), with the other combinations failing to give satisfactory results. The band was purified and cloned into pGEM-T for DNA sequencing (Appendix VI).

Figure 4.20; Streptomyces *Degenerate Primers and Expected Product Sizes*

Multiple alignment of TGases from *S. cinnamoneus* CBS 683.68 (Accession No. CAA70055), *S. platensis* (Accession No. AAS84612) and *S. mobaraensis* IFO13819 (Accession No. AF531437) showing the regions of homology (highlighted) that were used to design the degenerate primers for PCR (A) with expected product sizes shown (B).

(A)
Forward Primers

Reverse Primers

(B)

Primer Combination	Expected Product Size (bp)
F1 + R1	450
F1 + R2	680
F2 + R1	375
F2 + R2	620

The obtained sequence shared no similarity to the corresponding region of *Streptomyces* TGase. The sequence obtained was used to search the translated protein database (non-redundant) (Figure 4.22). The sequence, obtained by degenerate PCR, shared 63% similarity with D-tyrosyl-tRNA (Tyr) deacylase from the genome of *Ps. putida* KT 2440, as well as sharing similarity to other D-tyrosyl-tRNA (Tyr) deacylase enzymes originating from other bacterial species. Interestingly higher similarity was observed between D-tyrosyl-tRNA (Tyr) deacylase enzymes originating from *Polaromonas sp.,* *Rhodoferax ferrireducens*, *Azoarcus sp., Ralstonia metallidurans* and *Burkholderia xenovorans* exhibiting similarities of 73%, 73%, 70%, 69% and 74%, respectively. As such, further investigations into a possible TGase from *Ps. putida* ceased.

Figure 4.21; Degenerate PCR Analysis of Ps. Putida NCIB 9872

Genomic DNA was isolated from *Ps. putida* and used as the starting material for PCR. Degenerate primers (Table 2.3) designed to *Streptomyces* TGase were used in all viable combinantions. PCR was carried out in the presence of 3 mM $MgCl_2$ for the following times: 94°C for 1 min, 30 cycles of 94°C for 30s, 50°C for 30s and 72° for 30s. After amplification, products were analysed by agarose gel electrophoresis (2% w/v) and bands visualised under UV light by ethidium bromides staining.
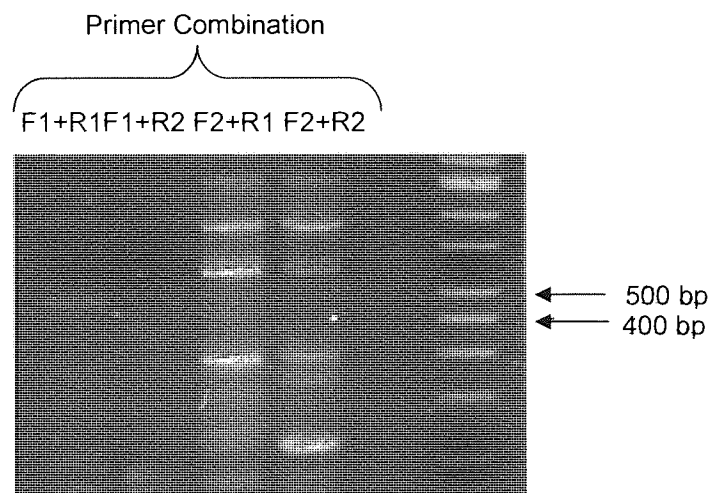
Figure 4.22; Result of tBLASTx Search using DNA Sequence generated from Degenerate PCR from Ps. putida NCIB 9872

The DNA sequence derived from degenerate PCR from *Ps. putida* (See Figure 4.21) was used to perform a tBLASTx search of bacterial sequences in the NCBI database.

The query sequence represents the amino acid sequence derived from the cloned DNA sequence from degenerate PCR (see Appendix VI). Subject sequence is derived from the genome of *Ps. putida* KT2440 (Accession number: AE015451 region: 5728952-5729389). Identical amino acids coloured red.

```
Query   377        SLQDIGGGLLVVSQFTLAADTRGGNRPSFTAAAAP

                   SL D GGGLL VSQFTLAADTR G RPSF  AA P

Sbjct   5729197    SLKDVGGGLLLVSQFTLAADTRNGMRPSFSTAAPP


Query   263        DEGRRLYEYFVDQARLAHAQVQTGEFAADMQVHLV

                      G  L  Y   QA    A V  GF  ADMQVHLV

Sbjct 5729083      ALGAELFDYLLQQAKAQYADVASGRFGADMQVHLV


Query   168        NDGPVTIPMRM   135

                   NDGPVT

Sbjct   5728988    NDGPVTFMLQI   5728955
```

4.3. Discussion

4.3.1. Database Searching and Screening for Microbial TGase

Searching both protein and nucleotide databases for TGase homologues, originating from bacteria, revealed that very few organisms comprise a sequence with significant similarity to either *Bacillus* or *Streptomyces* TGases. This is not completely unexpected when considering the relative uniqueness of microbial TGase isoforms.

Interestingly no gene or protein with significant similarity to *Streptomyces* TGase was identified in *Streptomyces coelicolor*, considered by many to be the model Streptomycete (Messer and Zakrzewska-Czerwinska 2002). This would infer that either TGase has originated from an organism not initially considered a member of the Streptomyces genus or conversely that *S. coelicolor* has been assigned incorrectly, assuming possession of TGase is a prerequisite for *Streptomyces sp.*. Since TGase was not detected in all *Streptomyces* strains examined (Section 4.2.3.2) the former is more likely. Isolates containing TGase sequences shared a high degree of amino acid similarity to the query sequence (TGase from *S. mobaraensis*), greater than 79% (Figure 4.1). This shall be discussed in more depth in Chapter 5.

A number of sequences were highlighted in related *Bacillus sp.* as similar to *B. subtilis* TGase (Figure 4.2). It would appear that the sequences do not share the degree of similarity attributable to their *Streptomyces* counterparts. Interestingly, a potential TGase was highlighted in *Polaromonas sp.*, though the protein encoded shares only 26% similarity.

To conclude, the sequences of microbial TGase isoforms are sufficiently unique such that database searches fail to identify any potential TGase proteins/genes in other organisms. It is possible, therefore, that alternative, potentially novel enzyme isoforms reside in organisms not scheduled for genome sequencing. Therefore screening for enzyme activity offers the best strategy for identifying new microbial TGase derivatives.

4.3.2. Identification of TGase from Micro-Organisms

Gene homology and literature searches were of little use in identifying microbial TGase homologues and as such alternative methods were investigated. Large scale searching for microbial TGase has been carried out previously (Motoki, et al. 1992; Andou, et al. 1993; Bech, et al. 1996) with investigators utilising different TGase detection methods, hydroxamate activity and incorporation of radio labelled putrescine into casein (Lorand, et al. 1972; Folk and Chung 1985). These rely on the enzymes' ability to incorporate amines into casein, though the sensitivity of the hydroxamate assays limits it usefulness in screening low TGase producing organisms (Section 3.2.2). Similar assays, depending on the incorporation of an amine substrate into casein, were used to screen micro organisms for TGase activity. An alternative, more specific assay was used to both identify and confirm enzymes capable of forming the ε-(γ-glutamyl) lysine bridge (peptide cross-linking assay).

4.3.2.1. TGase from Fungal Isolates

The fungi constitute an independent group of organisms that shares some characteristics with higher organisms such as chitinous structures, storage of glycogen, and mitochondrial UGA coding for tryptophan (Barr 1992). The branch linking fungi and animals has been supported by 18S rDNA sequence analysis (Bruns, et al. 1991). The Fungal life cycle comprises germination from a spore followed by a period of growth, with sporulation completing the cycle. Of the fungal strains screened, apparent TGase was detected in *A. parasiticus* and two *A. terreus* species early after inoculation (see Section 4.2.3.1). Fungal isolates were cultured in a nutrient rich medium that resulted in a high background reading in the hydroxamate assay, hampering the assay, a similar observation was observed during the purification of gpl TGase (see Section 3.2.1). As such, the growth medium was changed in favour of potato dextrose medium, however the high background remained. It is known that fungal strains secrete a large amount of proteolytic enzymes in response to germination signals (Bergkvist 1963; Yanagita and Nomachi 1967) and it is possible that the presence of such proteins interferes with the

hydroxamate assay. Therefore confirmation of TGase activity was sought by partial purification from the culture medium of *A. terreus* IFO7079. Protein eluted was analysed by both ELISA and Western blotting using antibodies against both eukaryotic and microbial TGase. Enzyme activity was not confirmed and further analyses suspended.

4.3.2.2. TGase from Actinomycetes

A number of Actinomycetes have been screened by European partners as part of the EC HIPERMAX project, using the hydroxamate assay. Species appearing positive were sent to Aston University for confirmatory testing. Upon receipt strains were cultured and analysed by both hydroxamate and biotin cadaverine assays. TGase activity was confirmed in a number of species by biotin cadaverine. The ELISA based assays were favoured for confirmatory testing since a large number of Actinomycetes produce pigments that may have influenced the hydroxamate assay. Also the ELISA based assays were found to be both more reliable and sensitive. This may perhaps explain why only a limited number of species were identified as exhibiting TGase activity be the hydroxamate method.

TGase activity was detected in a number of *Streptomyces* species as well as in *Saccharomonospora viridis*. More detailed biochemical and physiological analysis has been carried that is presented and discussed in Chapter 5.

4.3.2.3. TGase from Members of the Bacilliaceae

TGase was initially described in the organism *B. subtilis* (Kobayashi, et al. 1998c) and to date has only been identified in a few other related species through genome sequencing (Rey, et al. 2004). To date only the TGase found within *B. circulans* (de Barros Soares, et al. 2003) has been published. Although *Bacillus sp.* may be separated into seven groups *B. circulans* could not be included in any of the groups (Xu and Cote 2003). This in part, may explain the diversity of the TGase encoded since it possesses a higher molecular weight (45 kDa) than TGase from *B. subtilis* despite appearing to carry out a similar function to its counterpart in *B. subtilis* (de Barros Soares,

et al. 2003). Whereas the role of Streptomyces TGase remains unknown, the function of *Bacillus sp.* TGase has been established where it has been shown to play a role in the later stages of sporulation (Kobayashi, et al. 1998b; Ragkousi and Setlow 2004). TGase activity detected from *B. polymyxa* was observed at a different time point upon comparison with other species screened perhaps suggesting a different affinity for the culture medium or a different rate of sporulation.

Attempts to purify TGase from the spore coat surface of *B. subtilis* were hampered by the apparent low levels of enzyme. It was anticipated that the gene encoding *B. subtilis* TGase would be cloned and expressed in *E. coli*, a method carried out successfully by Kobayashi and co-workers (Kobayashi, et al. 1998a). Despite using very similar systems only a limited amount of success was obtained. TGase protein was expressed in only one system; reasons for this may include codon usage differences between *Bacillus* and the expressing strain, though expression of a potentially toxic gene product may also be responsible; however the protein expressed exhibited no enzymatic activity as determined by three separate TGase assays (biotin cadaverine, peptide cross-linking and radio-labelled putrescine assays). One can only assume that the problems encountered during this study, with regards to protein purification and expression, are the reasons why *Bacillus sp.* TGase has not reached the marketplace to be exploited as an alternative to *Streptomyces sp.* TGase.

4.3.2.4. TGase from other Micro-organisms

The majority of micro-organisms selected and screened were found to be TGase negative with the exception of *Ps. putida*. This reflects the difficulty in predicting the likely origins of alternative microbial TGase sources. Other difficulties include the selection of medium and growth conditions since these may influence protein expression.

This study would suggest that microbial TGase is indeed restricted to a limited number of organisms (*Bacillus* and *Streptomyces*). Whilst every attempt was made to completely evaluate the TGase assays for the detection of physiological levels of TGase, it remains possible that the assays were not

sensitive enough to detect TGase from alternative species. However, it should be noted that TGase was easily detected from *B. subtilis* in accordance with published work (Kobayashi, et al. 1998b; de Barros Soares, et al. 2003). TGase activity has not previously been described in the organisms screened and there is no existing evidence that these organisms produce TGase, despite reports of bacterial homologues to mammalian TGases (Makarova, et al. 1999; Ginalski, et al. 2004). At the outset it was apparent that organisms exhibiting a sporulation phase were likely to produce a TGase protein since both *Bacillus subtilis* and *Streptomyces mobaraensis* undergo forms of sporulation and produce TGase (Ando, et al. 1989a; Kobayashi, et al. 1998c). Therefore one may expect other members of the genus to exhibit TGase activity as well as other spore forming micro-organisms, including fungi. However this was not the case, since TGase activity was confined to members of the *Bacilliaceae* and Actinomycetes.

Whilst no TGase activity was detected in the majority of organisms screened, it is possible that the proteins are in fact cysteine proteases that do not possess TGase activity or indeed do not cross-link proteins by a TGase type reaction, though carrying out a similar function. It is also possible that TGase proteins exist but are expressed at specific time points or at times of specific challenge (environmental or physiological) whereby they are active against a narrow substrate range as is the case with the toxins CNF and DNT (Schmidt, et al. 1998; Horiguchi 2001). In such instances the assays used would be unable to detect them.

4.3.2.4.1. Potential TGase from Ps. putida

The Pseudomonads represent a metabolically diverse group of organisms. This metabolic versatility as well as the need to constantly adapt to varying environmental conditions suggests that these bacteria possess sophisticated mechanisms which govern their gene expression. Of the organisms screened TGase activity was detected in two out of three strains of *Ps. putida*, perhaps reflecting a degree of intra-species differentiation, since members of the genus show a large degree of both genetic and metabolic diversity (Barrett

and Bell 2006). Maximal activity was detected towards the later stages of exponential growth indicating a potential role in cellular maturation. Further analysis revealed the protein responsible to behave in a different manner to known TGases when assessed in the presence of amine inhibitors (Figure 4.13). However, apparent TGase activity was susceptible to the synthetic inhibitor R281 suggesting the protein to be a TGase. Further confirmation was gained by suppression of enzyme activity following incubation with N-ethylamine suggesting a cysteine residue was, in part, responsible for the activity observed.   The enzyme responsible was partially purified and deemed to possess a molecular weight of between 20-50 kDa (Figure 4.15). Purification of the enzyme responsible proved difficult, perhaps due to the low physiological amounts and instability of the enzyme present. Use of the R281-coupled Sepharose proved unhelpful in purification of the protein responsible. Using this method it was apparent that the resin bound more than one protein. Therefore, it is possible that the inhibitor R281 is capable of binding to surface exposed cysteine residues in cellular proteins as well as those within the active site of TGases. This may limit the use of such a method in the identification of novel TGase isoforms.

Database searches highlighted the presence of a potential TGase protein from the published genomic sequence. Further analysis of this sequence revealed the presence of similar domains found throughout the eukaryotic TGases. The protein was successfully cloned, though no protein expression or activity could be detected. This highlights the difficulty of computer programs to accurately predict both coding regions as well as ascribing biological functions, based on protein homology to existing proteins. Such methods are also hampered since computer programs used to search for microbial TGase homologues use mammalian TGase as a reference. It is unlikely for any potential microbial TGase to share appreciable homology to their eukaryotic counterparts, as is the case with both *Bacillus* and *Streptomyces* derived TGases.

Using degenerate primers, designed to amplify conserved regions of *Streptomyces* TGases, a cross reacting band was amplified from the genomic DNA purified from *Ps. putida* 9872 corresponding to that amplified

from *Streptomyces* sp.. After the fragment was sequenced and subcloned, database searches revealed the similarity to the enzyme D-tyrosyl-tRNA (Tyr) deacylases originating from the organism *Ps. putida* KT2440 (see Section 4.2.3.10). Physiologically, D-TyrtRNATyr deacylases specifically recycle misaminoacylated D-Tyr-tRNATyr residues by hydrolyzing the ester link between D-Tyr and tRNA (Soutourina, et al. 1999). The mechanism has been postulated to aid in the protection of cells against the toxicity of D-amino acids produced by endogenous metabolism (Soutourina, et al. 1999). They are also capable of using D-Asp, D-Trp, D-Ser, D-Leu, D-Gln, D-Phe and D-Gly as substrates also (Lim, et al. 2003).

Analysis of the literature surrounding D-TyrtRNATyr deacylase enzymes provides possible explanations to the results observed during bacterial screening. The presence of amines histamine and putrescine were found to illicit a dose dependent increase in apparent TGase activity observed from cell lysates of *Ps. putida* NCIB 9872 (see Figure 4.13). The addition of 1mM spermidine, structurally related to putrescine and histamine, has been shown to stimulate a 3-fold increase in D-Tyr-tRNATyr deacylases hydrolysis (Soutourina, et al. 1999). This would suggest that ionic strength would improve the rate of reaction of both deacylase and the potential enzyme observed in *Ps. putida*.

Though the exact active site has yet to be determined, the crystal structure of D-TyrtRNATyr deacylases from *Haemophilus influenzae* and *E. coli* have been solved, and been shown to share 67% identity (Ferri-Fioni, et al. 2001; Lim, et al. 2003). Analysis of the 3D structure reveals the presence of an oxyanion hole, located adjacent to Thr-80 (a proposed residue involved in nucleophilic attack of substrate), which serves to stabilize the negatively charged transition states (Lim, et al. 2003). The presence of the oxyanion hole is reminiscent of the catalytic machinery of other hydrolytic enzymes including the thiol proteases, to which TGases are distantly related, and the serine-lactamases (Kraut 1977; Herzberg and Moult 1987). Similar to Gln-78 of deacylase, either a glutamine or asparagine amide group participates in the formation of the oxyanion hole in papain and subtilisin, respectively. If we consider the deacylase proposed reaction mechanism (Figure 4.23), it bears

a striking similarity to that of mTGase, albeit in reverse. The deacylase proposed mechanism follows the general base mechanism observed in other hydrolytic enzymes and involves the nucleophilic attack of a carbonyl atom in the substrate molecule with formation of an acyl enzyme intermediate before release of the hydrolysis product. The formation of such an acyl enzyme intermediate is mirrored by TGase enzymes.

Figure 4.24; Proposed Catalytic Mechanism of D-Tyr-tRNATyr Deacylase

Based on the crystal structure of *H. influenzae* deacylase (Lim, et al. 2003).
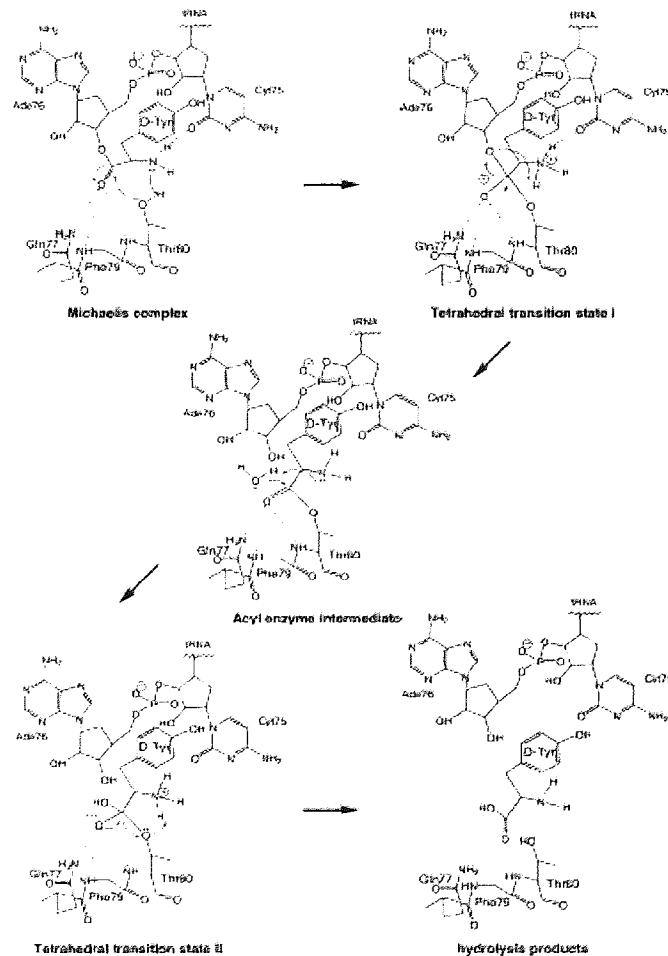
From the data presented it would appear that the apparent TGase activity observed was that catalyzed by the endogenous deacylase enzyme derived from *Ps. putida*. Though the TGase family has obvious differences in substrate specificity when compared to the papain family of cysteine proteases, considerable kinetic and mechanistic similarities exist (Parameswaran, et al. 1997). Most notably both operate via an acylation-deacylation pathway with a cysteine in the catalytic centre. TGases do however, show exceptional specificity for lysine side chains in proteins eventually leading to the formation of protein cross-links. From the observations made during the program of work undertaken, and the similarities drawn between D-TyrtRNATyr deacylase and TGase discussed previously it is possible that the deacylase family may also have evolved from an ancestral protease, a distant relative of the TGase superfamily and since they share a similar mechanism of action the enzymes may have evolved multiple functions, as has the TGase superfamily, accounting for the apparent TGase activity observed. Though the requirement of a TGase enzyme in *Ps. putida* remains to be determined, one may speculate a role in maintaining cellular integrity given the recovery of maximum activity during the transition between exponential and stationary phase of growth. However, it should be noted that the link between the two enzymes made on the basis of such conclusions may be tenuous. From the similarities observed, it would be interesting to confirm or deny the ability of deacylases to take part in additional enzyme catalyzed reactions and compare them to the reactions catalyzed by the TGase superfamily. If proven it would imply some degree of evolutionary relationship between the deacylase and bacterial TGases and may present an industrial alternative to *Streptomyces sp.* TGase provided suitable expression could be attained.

Chapter 5

Studies on TGases of *Streptomyces* and Related Genera

5.1. *Introduction*

The family *Streptomycetaceae* contains the genus *Streptomyces* which itself comprises species formerly classified as *Chainia, Elytrosporangium, Kitasatoa, Kistaosporia, Actinosporangium* and *Streptoverticillium*. These species had been associated with the *Streptomyces* genus based on morphological and chemotaxonomic grounds (Embley and Stackebrandt 1994). Nucleic acid sequencing and pairing studies confirmed that they could not be separated from the Streptomyces according to Schleifer (Schleifer and Stackebrandt 1983) and as such were reclassified as members of the *Streptomyces*. To date a large number of *Streptomyces* species have been described (>3000) as a result of their biotechnological importance as producers of antibiotics. Producers of novel bioactive compounds were described as new species and patented leading to over classification of the genus.

In addition to numerical classification, based on phenotypic traits described by Williams (Williams et al. 1983), a number of additional chemotaxonomic and molecular methods are now used to improve *Streptomyces* taxonomy, these include; cell wall composition (Lechevalier and Lechevalier 1970); phage typing (Wellington and Williams 1981); DNA-DNA hybridisation (Labeda 1992); comparison of 16S rRNA and 23S rRNA (Stackebrandt et al. 1991); comparison of ribosomal protein patterns (Ochi 1989) and low frequency restriction fragment analysis (Beyazova and Lechevalier 1993). However, it should be noted that many of the enumerated methods have yet to be carried out simultaneously. Indeed work carried out by Manfio (Manfio et al. 2003) highlighted that all tests are not consistent in identifying any given *Streptomyces sp.*.

Studies carried out by Williams (Williams et al. 1983) identified 23 major clusters (4 or more strains), 20 minor clusters (2-4 strains) and 25 clusters containing a single member within the *Streptomyces* genus. In attempts to further define the genus, analyses were carried out on the ribosomal protein AT-L30 (Ochi 1995b), with such protein analyses being previously proven to be effective in the taxonomic evaluation of other eubacteria (Ochi 1994;

1995a). The results of such analysis generally agreed with the phenotypic study undertaken by Williams (Williams et al. 1983), however a number of disagreements were also noted, disagreements that would need to be clarified by alternative methods (DNA relatedness analysis) if they were to be used in unison for the identification and classification of *Streptomyces*.

Transglutaminase of *Streptomyces* origin was first described in the culture filtrate of S-8112, a variant of *S. mobaraensis* (Ando et al. 1989a). Based on the phenotypic study carried out by Williams (Williams et al. 1983) *S. mobaraensis* belongs to Cluster F58 and shares a high degree of similarity to *S. lavendulae*, and is closely related to *S. griseocarneus* and *S. cinnamoneus* (Cluster F55).

Having screened more than 5000 organisms Andou *et al.* (Andou et al. 1993) detected TGase enzyme activity in the filtrate of the organism *Streptomyces* S-8112. Since its identification, extracellular TGase has been described in only a few other *Streptomyces sp.* namely; *Streptomyces griseocarneus* and *Streptomyces cinnamoneus* subsp. *cinnamoneus* (Motoki et al. 1992); *Streptomyces sp.* and *S. lavendulae* (Andou et al. 1993); *S. ladakanum* (Tsai et al. 1996); *S. lydicus* and *S. platensis* (Bech et al. 1996); and more recently *S. hygroscopicus WSH03-13* (Cui et al. 2006) and *S. fradiae* (Liu et al. 2006a). Of the publicly accessible sequences, TGases from the aforementioned *Streptomyces* share more then 79% and more than 82% identity at the protein and nucleotide level respectively.

As detailed in Chapter 4, TGase was detected in the culture supernatant of a number of *Streptomyces* isolates. The ability of these enzymes to introduce isopeptide cross-links into proteinaceous substrates makes them ideal candidates for a number of industrial/commercial processes (Section 1.5.2.1).

It is the focus of this chapter to concentrate on the purification of TGase proteins from *Streptomyces* and related genera. Subsequent analysis of enzyme properties and functionality will thereby assess their commercial usefulness and potential for industrial application.

5.2. *Results*
5.2.1. *Analysis of TGase Proteins*

TGase is initially transcribed as an inactive zymogen that is secreted from the cell and proteolytically processed by a number of interdependent proteases to release the active, mature enzyme (Zotzel et al. 2003a; Zotzel et al. 2003b) (Section 1.4.2.3). A number of isolates from the *Streptomyces* genera as well as *Saccharomonospora viridis* have been shown to exhibit TGase activity (Section 4.2.3.2). It is not known when, during their life cycle, TGase activity is maximal or the quantity of active enzyme produced.

Streptomyces and related isolates were grown on GYM medium (30 °C with shaking at 225 rpm) and sampled from Day 3 to Day 10. Growth curves were constructed by the determination of wet cell weight (Section 2.4.5). Additionally, cell-free samples were analysed by both the hydroxamate assay, for the determination of enzyme activity, and probed by Western blotting using the polyclonal antibody raised against TGase from *S. mobaraensis* (N-Zyme Biotech, Germany) (Sections 2.5.1 and 2.9.3 respectively). From the resulting data, (Figure 5.1 A-D) it may be seen that all isolates analysed first produced a higher molecular weight cross-reacting ($\approx$42 kDa as analysed by SDS-PAGE) band corresponding to the approximate molecular weight of Pro-TGase. In each case the levels of Pro-TGase decreased with culture age with a concomitant increase in the levels of a smaller molecular weight band (mature TGase). The appearance of this smaller molecular weight band corresponded with an increase in TGase enzyme activity, determined by hydroxamate. These findings are in agreement with those put forward by Pasternack (Pasternack et al. 1998) who described the sequential secretion of Pro-TGase and subsequent proteolytic activation of the zymogen. The existence of mature TGase in the culture medium appeared to be relatively short before the enzyme was broken down by inherent protease action. Maximal TGase activity was detected at the transition between the late exponential and stationary phases of growth, determined by wet cell weight and as such supports a role for TGase in the differentiation process.

In addition to the presence of inactive and active isoforms of TGase, a strong cross-reacting band was observed ($\approx$10 kDa) that was found to be an artefact of the purification procedure from which the antibody was produced (personal communication R. Pasternack (Pasternack 2005)).

Figure 5.1; Analysis of TGase Producing Streptomyces Cultures

Streptomyces strains grown on agar plates were used as inoculum for TGase producing ability. The spores were inoculated into 250ml GYM medium and incubated at 30°C with shaking (225 rpm). Aliquots of culture broth were removed daily, cells removed by centrifugation at 10,000 rpm for 5 mins and used for the determination of wet cell weight as described in Section 2.4.5. Clarified culture supernatant was analysed for TGase activity (i) using the hydroxamate assay and also by Western blotting (ii). In each case equal amounts of proteins were separated by SDS-PAGE, transferred to nitrocellulose membrane and probed with anti-mtg (N-Zyme Germany) at a dilution of 1:10,000 and incubated overnight at 4°C with gentle shaking. After washing, primary antibody was detected by the appropriate HRP conjugated secondary antibody at a dilution of 1:5,000 after which cross-reacting bands were detected by chemiluminescence as described in Section 2.9.3.1.

(A)S. baldaccii

i. Enzyme Activity and Cell Weight with Time of Growth
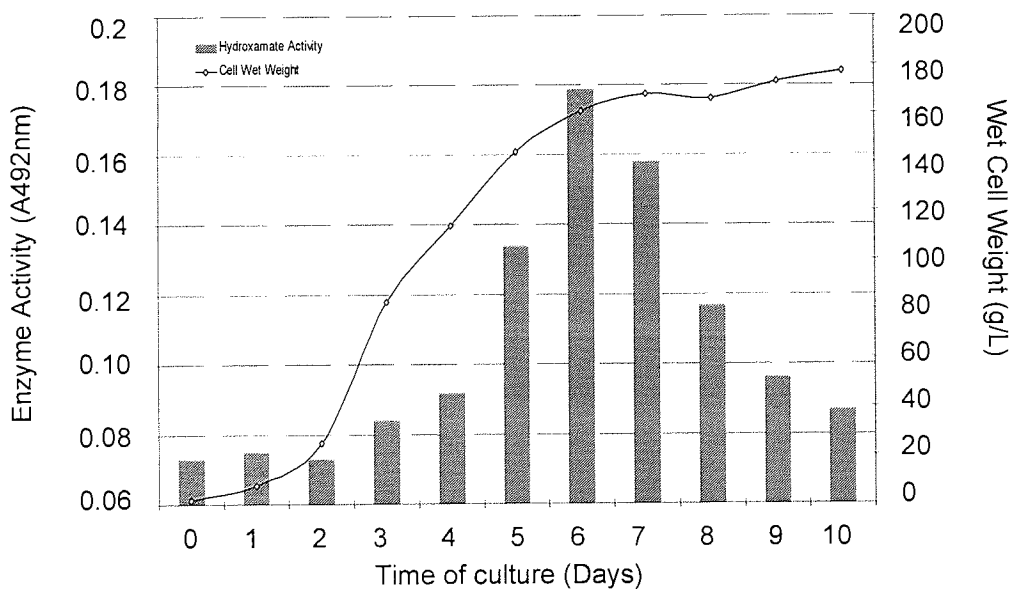
ii. Western Blot to determine TGase Protein in Culture Supernatants
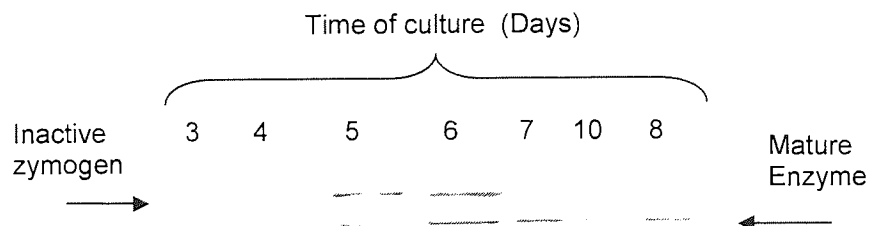
(B) S. viridis

i. Enzyme Activity and Cell Weight with Time of Growth
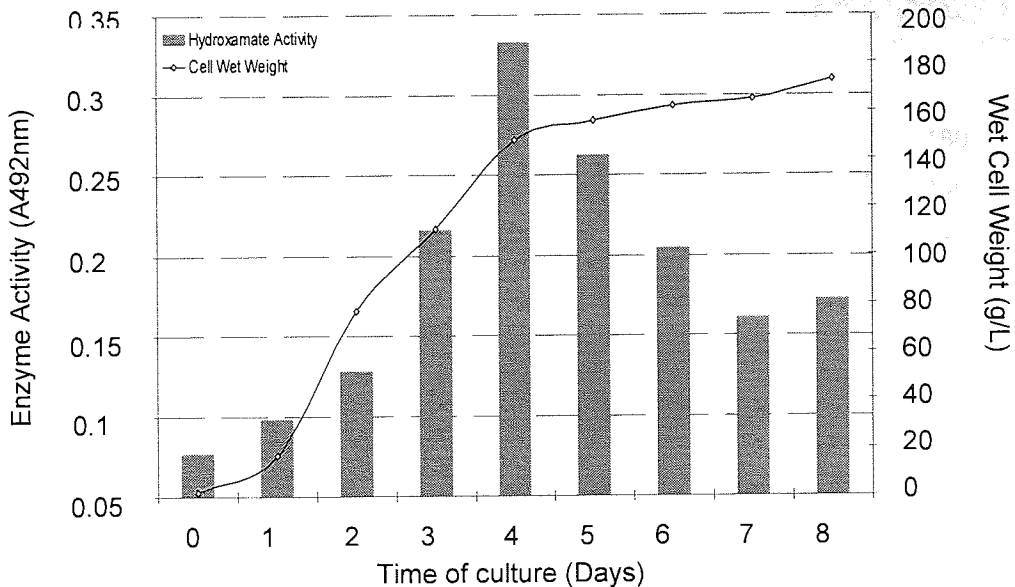
ii. Western Blot to determine TGase Protein in Culture Supernatants
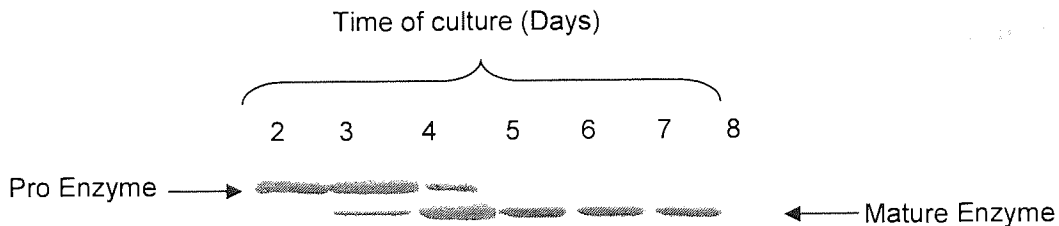
(C) S. platensis

i. Enzyme Activity and Cell Weight with Time of Growth
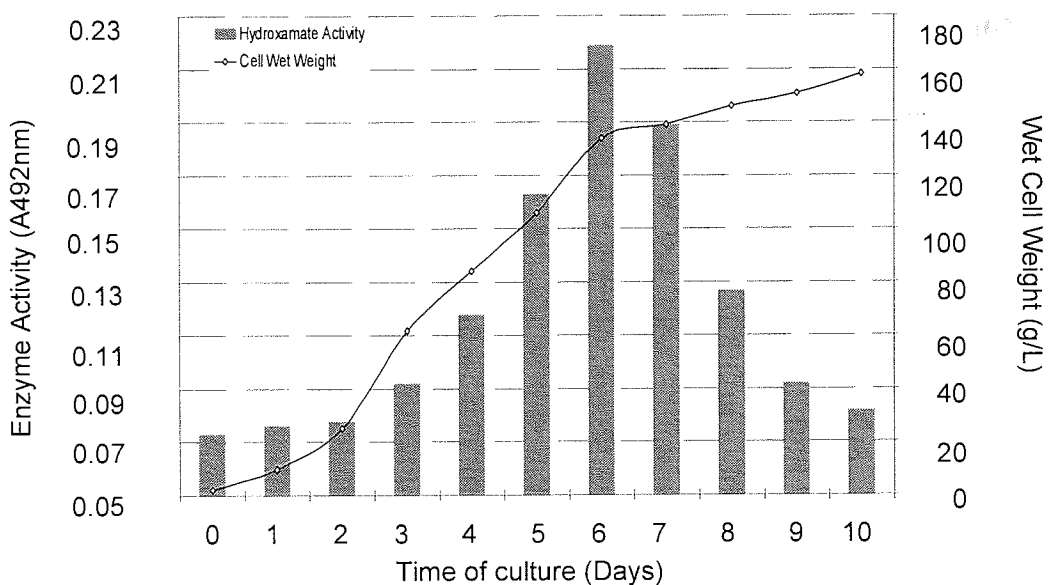
ii. Western Blot to determine TGase Protein in Culture Supernatants
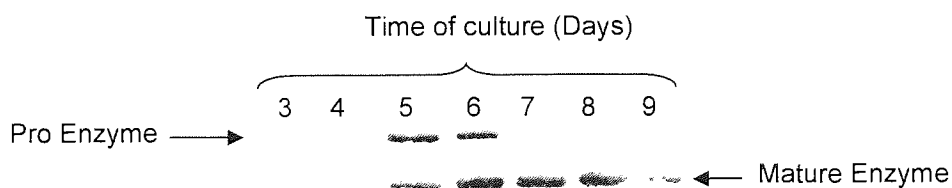
(D) S. paucisporogenes

i. Enzyme Activity and Cell Weight with Time of Growth
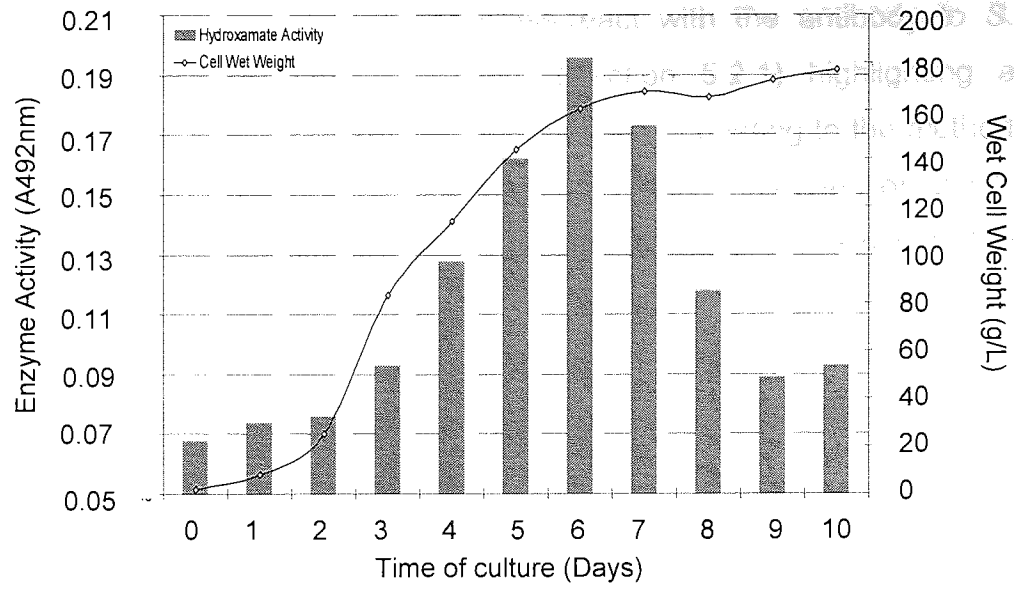
5.2.2. Purification of TGases from Streptomyces and Related Genera

Screening the *Streptomyces* and related genera highlighted the presence of an endogenous TGase in a number of isolates including *S. mobaraensis*, *S. platensis*, *S. paucisporogenes*, *S. baldaccii* and *Saccharomonospora viridis*. The TGases encoded were found to cross-react with the antibody to *S. mobaraensis* TGase (N-Zyme, Germany) (Section 5.2.1) highlighting a degree of structural similarity. Proteins were purified according to the method described by Duran (Duran et al. 1998) with modifications (Section 2.8.1) (Figure 5.2 A-D). The gelatin sepharose purification step was eliminated in favour of an additional cation exchange chromatography step using MonoS to achieve an improved separation of contaminants. Proteins purified in such a manner were used for biochemical characterisation and functional assays herein described.

Of the purified enzymes, highest yields (enzyme protein/L) came from *S. viridis* (9 mg/L) followed by *S. platensis* (4.2/mg/ml), *S. paucisporogenes* (3.6 mg/L) and *S. baldaccii* (2.9 mg/L). Additionally, the proteins shared similar specific activities as assayed by the incorporation of biotinylated cadaverine into N,N`-dimethylcasein.

Figure 5.2; Purification of TGase from Streptomyces and Related Genera

Cultivated cells were separated from the culture medium by centrifugation (10,000 x g for 5 min) and subsequent filtration (through Whatman No.1) with the resulting filtrate being applied to an SP-Sepharose column. After washing protein was eluted with a linear gradient (0-1M) of NaCl (Section 2.8.1.1). Fractions of high specific absorbance were pooled to which was added solid NaCl to a final concentration of 3M. Protein was applied to a phenyl sepharose column with protein eluted by a decreasing NaCl gradient (3-0M). Fractions exhibiting high specific absorbance were pooled and extensively dialysed against 20mM sodium phosphate pH 6. A final clean up step was performed to remove unwanted protein utilising cation exchange chromatography (Section 2.8.1.4) with an increasing NaCl gradient from 0-0.5 M.

Total protein absorbance (A750nm) and specific TGase absorbance (TGase activity [A450nm]/protein absorbance [A750nm]) of fractions obtained from I; cation exchange chromatography, II; hydrophobic interaction chromatography and III; cation exchange chromatography; IV; SDS-PAGE analysis of purified proteins, representative of each purification stage, was carried as described in the Materials and Methods using 12% acrylamide gels and V; the purification table generated where the total activity is given in units where1 unit is defined as the amount of enzyme catalyzing the formation of one micromole of hydroxamic acid per minute under the described reaction conditions. Specific activity was calculated from the total activity divided by total protein (determined by the Lowry assay). In each case the hypothetical NaCl gradient is indicated.

(A) TGase derived from S. baldaccii NRRL-3500

I. Post Cation Exchange Chromatography (SP Sepharose)
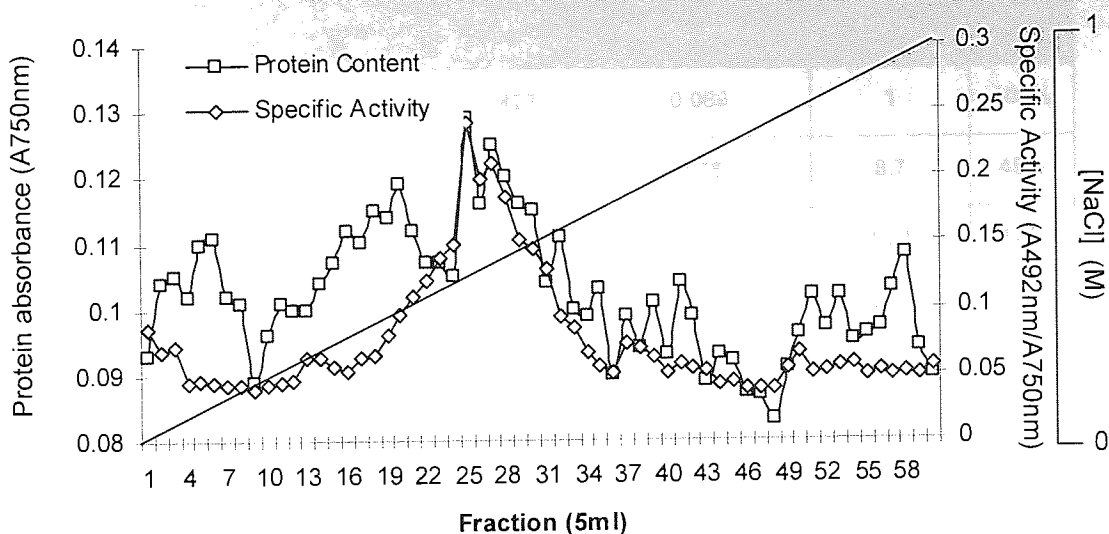
II. Post Hydrophobic Interaction Chromatography (Phenyl Sepharose)
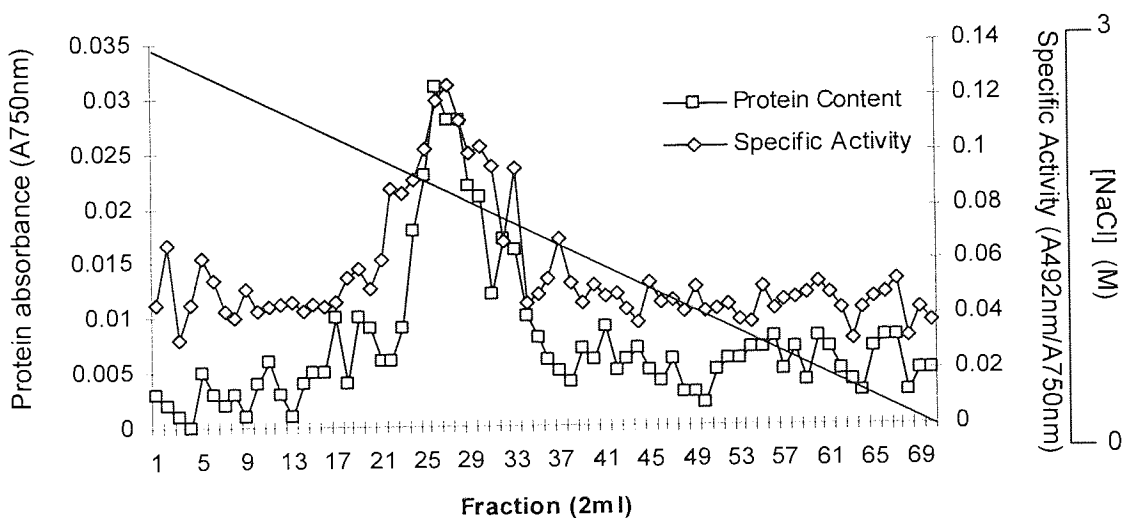
III. Post Cation Exchange Chromatography (Mono S)
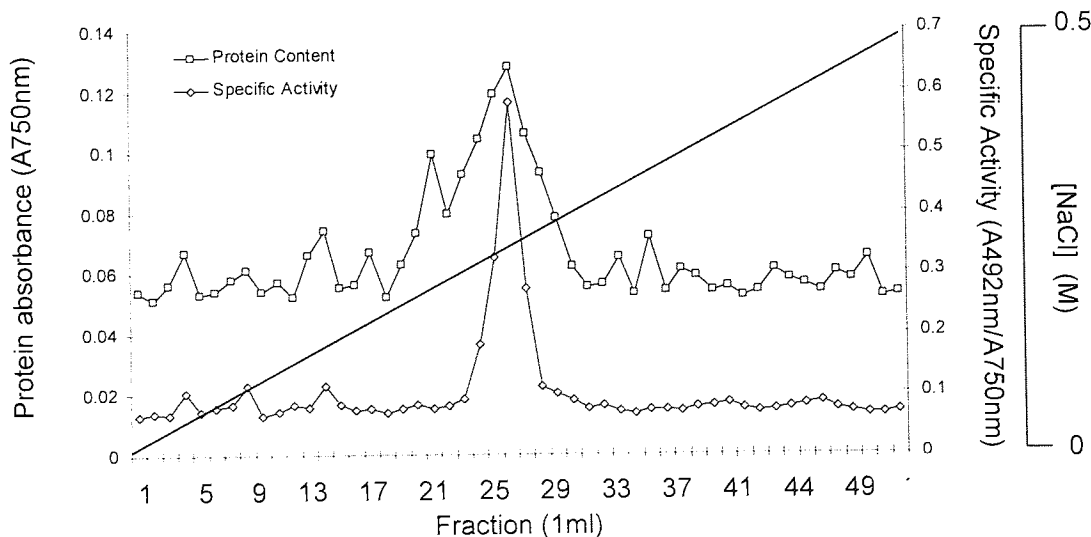
IV. Purification Table for TGase Derived from *S. baldaccii*

Purification Stage	Total Activity (U)	Total Protein (mg)	Specific Activity (U/ mg protein)	Purification (Fold)	Yield
Crude lysate	127	1427	0.089	1	100%
Cation exchange (SP Sepharose)	63	81	0.78	8.7	49%
HIC (Phenyl Sepharose)	21	18	1.17	13.1	17%
Cation exchange (MonoS)	4.8	2.9	1.53	17.1	4%

V. SDS-PAGE Analysis of Samples from the Different Purification Stages of TGase from *S. baldaccii*.

For the purpose of analysis, 10 µg of protein from pooled ion exchange and hydrophobic interaction fractions were separated by SDS-PAGE using 12% polyacrylamide gels. A total of 5 µg of pure TGase was analysed.

Mw- molecular weight markers, IX- post ion exchange, HIC post hydrophobic interaction chromatography, IX2 post ion exchange chromatography (MonoS).

160
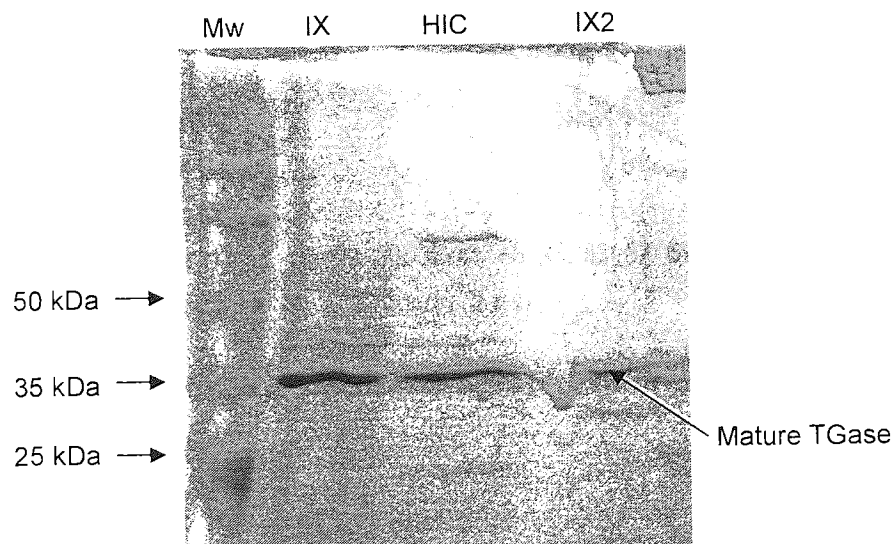
(B) TGase derived from Saccharomonospora viridis ATCC 15386

I. Post Cation Exchange Chromatography (SP Sepharose)
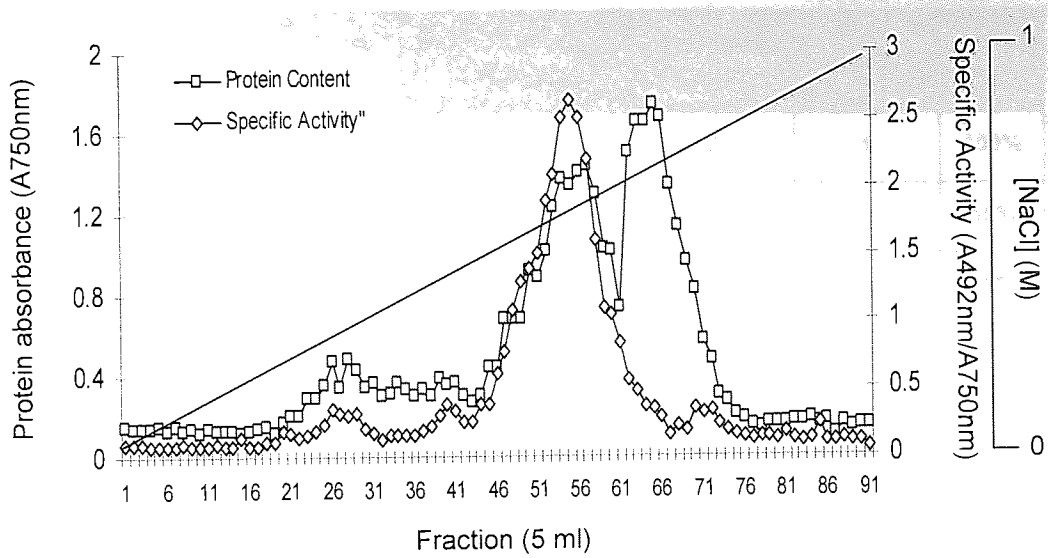
Fraction (5 ml)

II. Post Hydrophobic Interaction Chromatography (Phenyl Sepharose)

Fraction (2 ml)

III. Post Cation Exchange Chromatography (Mono S)

Fraction (1ml)
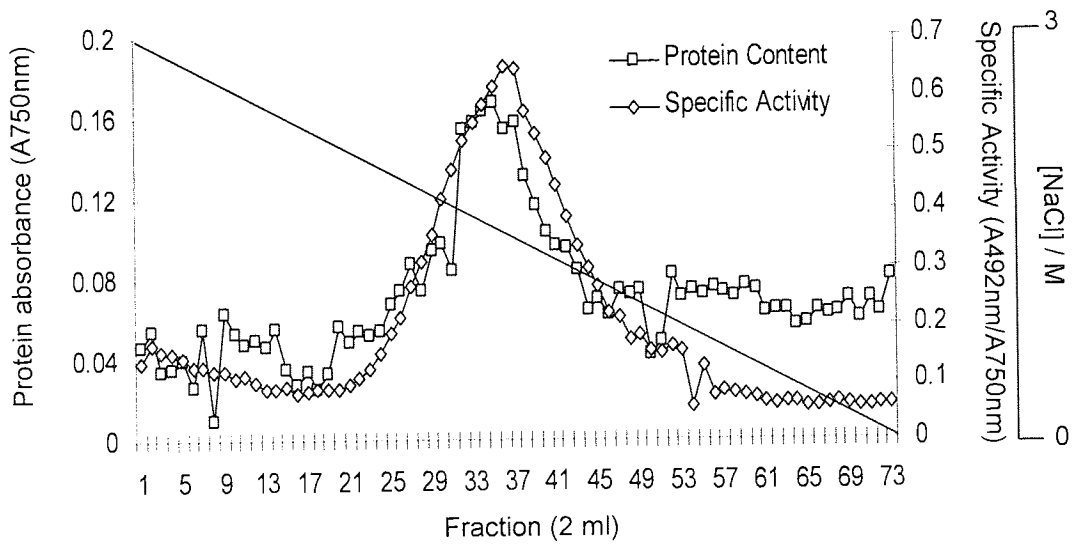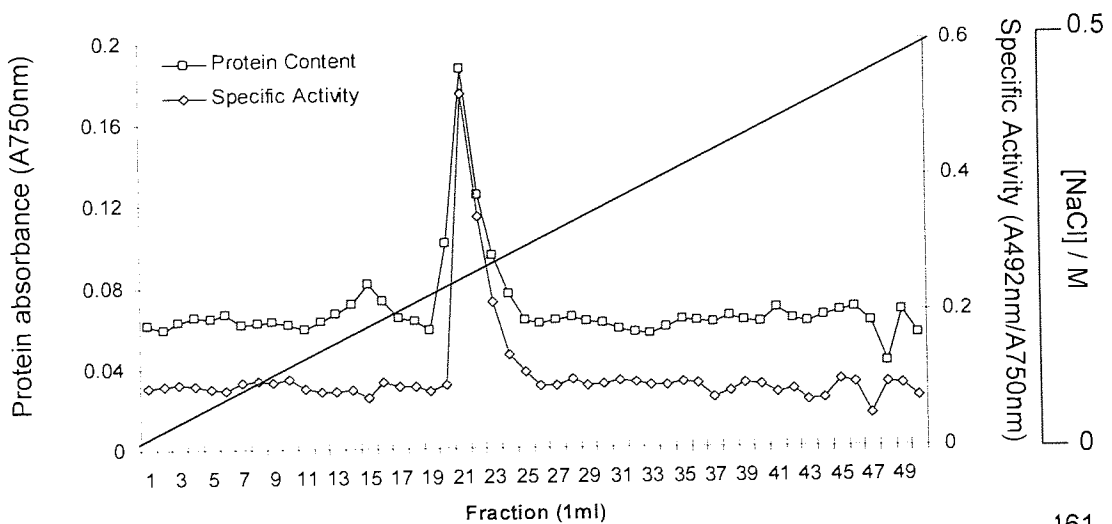
IV. Purification Table for TGase derived from *Saccharomonospora viridis*

Purification Stage	Total Activity (U)	Total Protein (mg)	Specific Activity (U/mg protein)	Purification (Fold)	Yield
Crude lysate	202	1535	0.132	1	100%
Cation exchange (SP Sepharose)	112	134	0.84	7.6	55%
HIC (Phenyl Sepharose)	50	46	1.1	8.3	25%
Cation exchange (MonoS)	15	9	1.63	12.3	7%

V. SDS-PAGE Analysis of Samples from the Different Purification Stages of TGase from *Saccharomonospora viridis*

Protein equating to 20, 10 and 5 µg from the ion exchange, hydrophobic interaction and MonoS chromatography (IX2), respectively was visualised by SDS-PAGE using 12% polyacrylamide gels.

Mw- molecular weight markers, IX- post ion exchange, HIC post hydrophobic interaction chromatography, IX2 post ion exchange chromatography (MonoS).
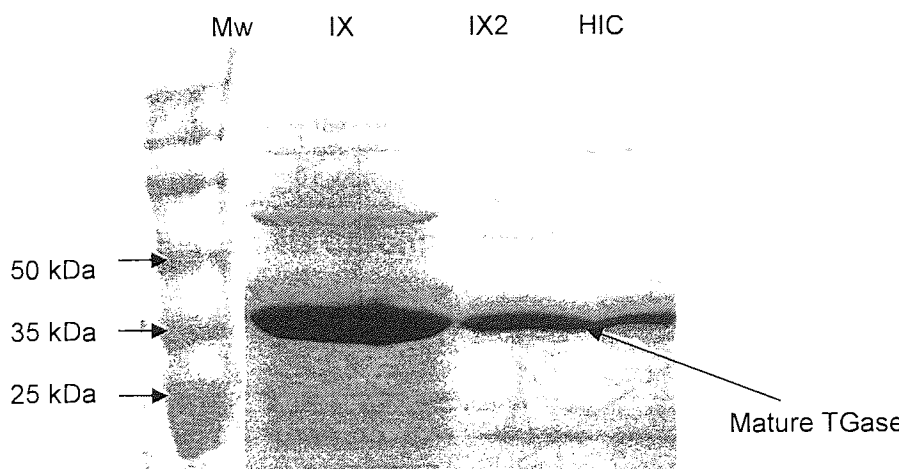
(C) TGase Derived from S. paucisporogenes ATCC 12696

I. Post Ion Exchange Chromatography (SP Sepharose)

Fraction (5 ml)

II. Post Hydrophobic Interaction Chromatography (Phenyl Sepharose)

Fraction (2 ml)
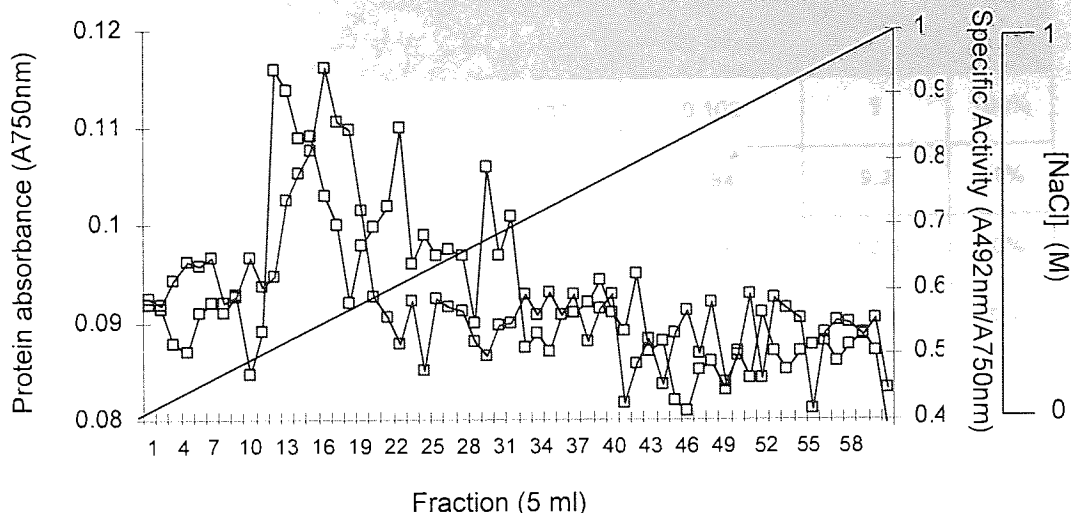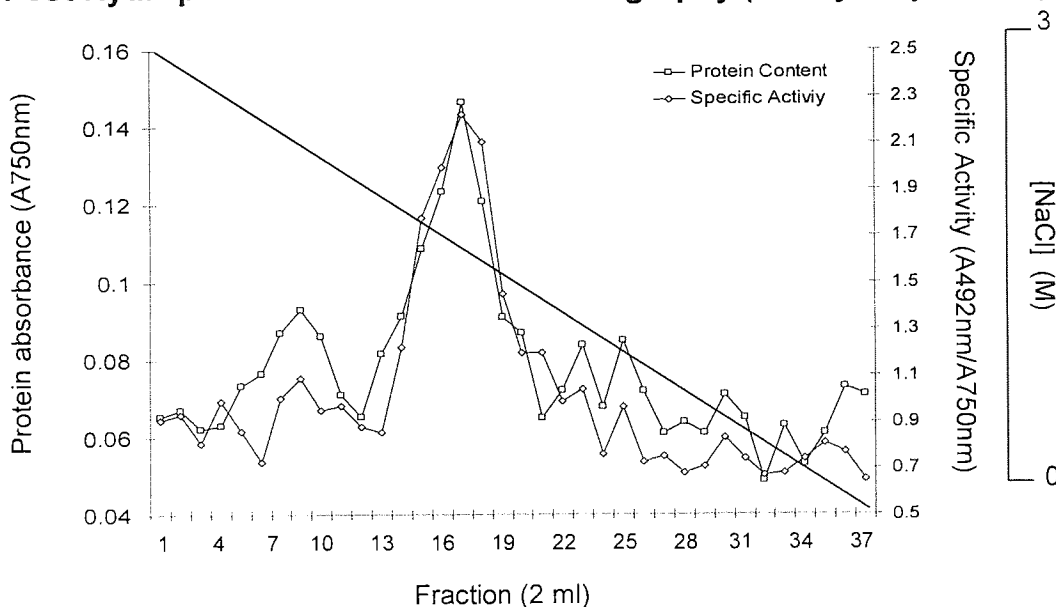
III. Post Gelatin Sepharose 4B Chromatography

Fraction (1ml)
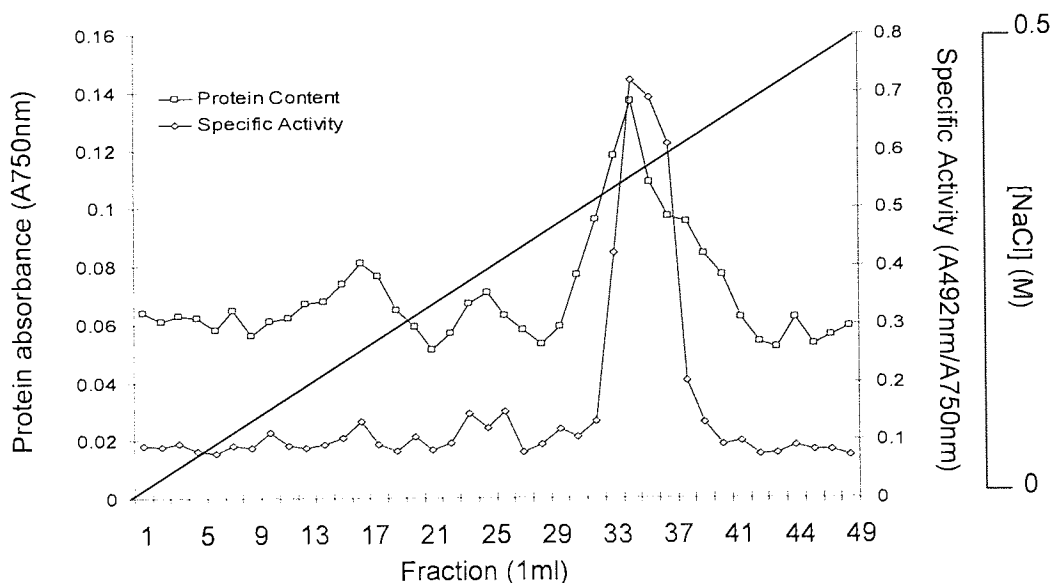
IV. Purification Table of TGase Derived from S. *paucisporogenes*

Purification Stage	Total Activity (U)	Total Protein (mg)	Specific Activity (U/mg protein)	Purification (Fold)	Yield
Crude lysate	116	1137	0.102	1	100%
Cation exchange (SP Sepharose)	82	87	0.94	9.2	71%
HIC (Phenyl Sepharose)	24	19	1.26	12.3	20%
Cation exchange (MonoS)	5.6	3.6	1.56	15.2	5%

V. SDS-PAGE Analysis from the Different Purification Stages of TGase from S. *paucisporogenes*

From the active pooled fractions of each purification round 10 µg of protein was analysed by SDS-PAGE using 12% polyacrylamide gels.

Mw- molecular weight markers, IX- post ion exchange, HIC post hydrophobic interaction chromatography, IX2 post ion exchange chromatography (MonoS). Figure depicts results from a single purification.
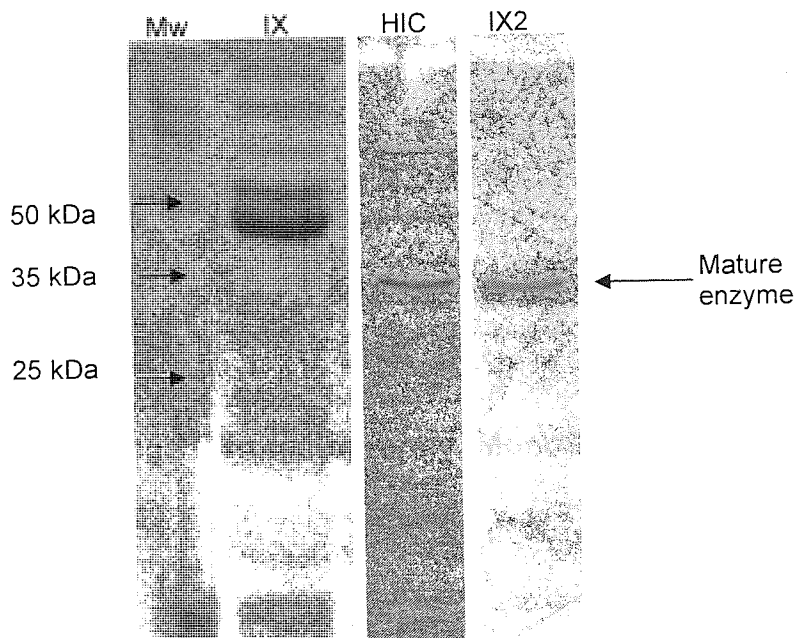
(D) TGase Derived from S. platensis ATCC 13865

I. Post Ion Exchange Chromatography (SP Sepharose)
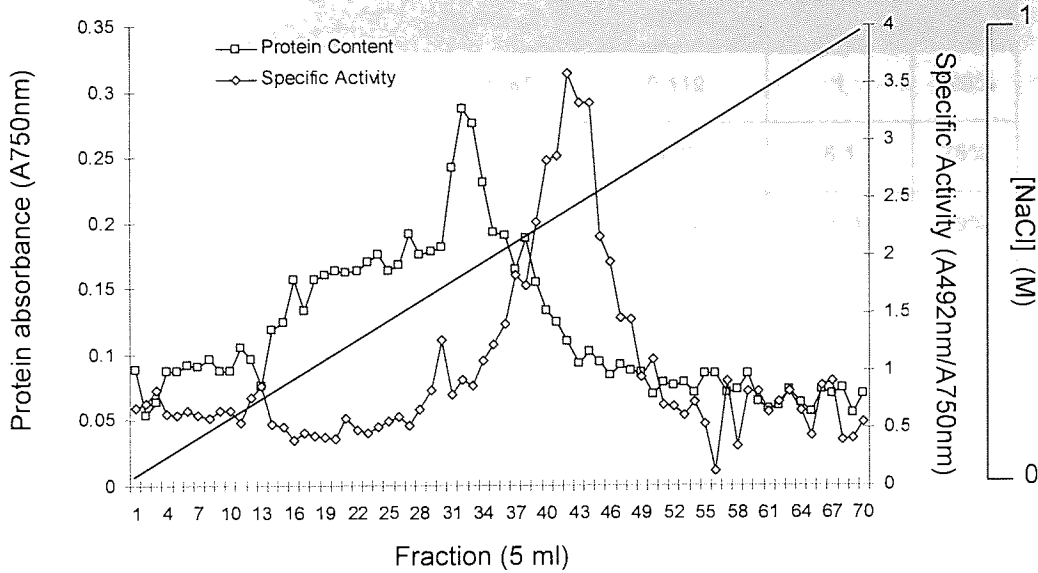
II. Post Hydrophobic Interaction Chromatography (Phenyl Sepharose)
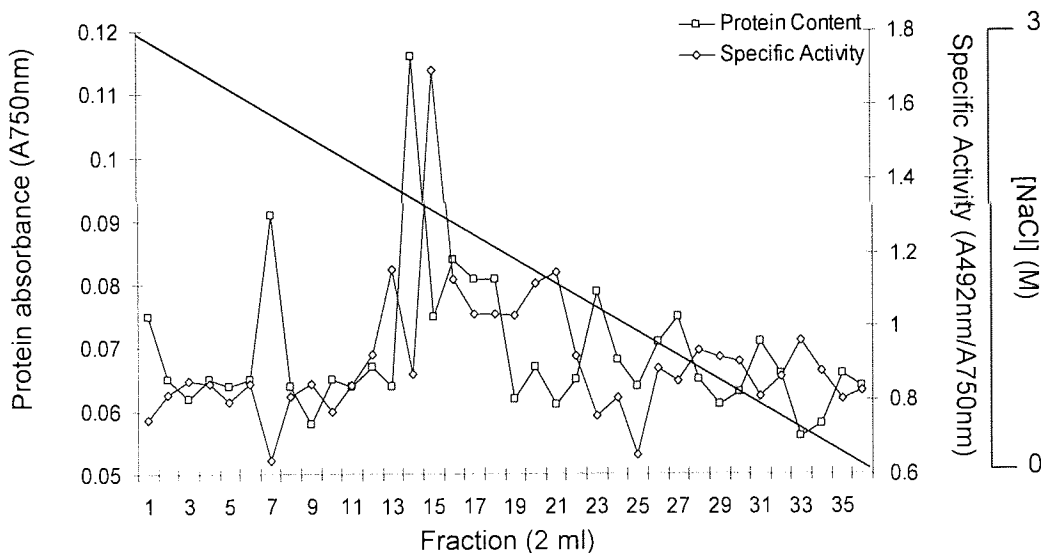
III. Post Cation Exchange Chromatography (MonoS)
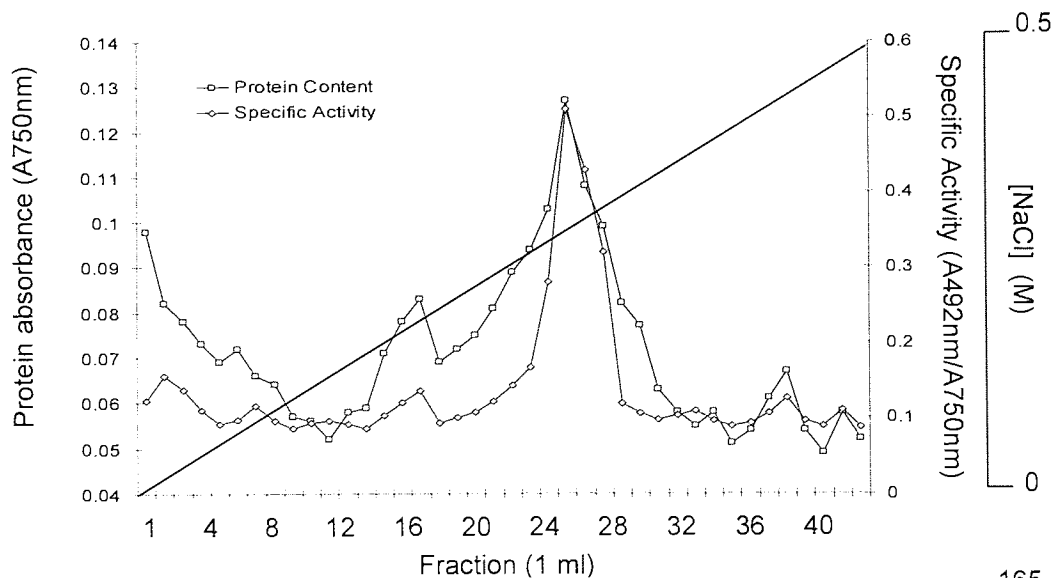
IV. Purification Table of TGase Derived from *S. platensis*

Purification Stage	Total Activity (U)	Total Protein (mg)	Specific Activity (U/mg protein)	Purification (Fold)	Yield
Crude lysate	163	1368	0.119	1	100%
Cation exchange (SP Sepharose)	122	167	0.73	6.1	75%
HIC (Phenyl Sepharose)	31	25	1.24	10.4	19%
Cation exchange (MonoS)	6.7	4.2	1.61	16.6	5%

V. SDS-PAGE Analysis from the Different Purification Stages of TGase from *S. platensis*

A total of 10 µg of protein was analysed from each purification stage with the exception of the MonoS step (IX2), from which 5 µg was analysed.

Mw- molecular weight markers, IX- post ion exchange, HIC post hydrophobic interaction chromatography, IX2 post ion exchange chromatography (MonoS). Figure depicts results from a single purification.
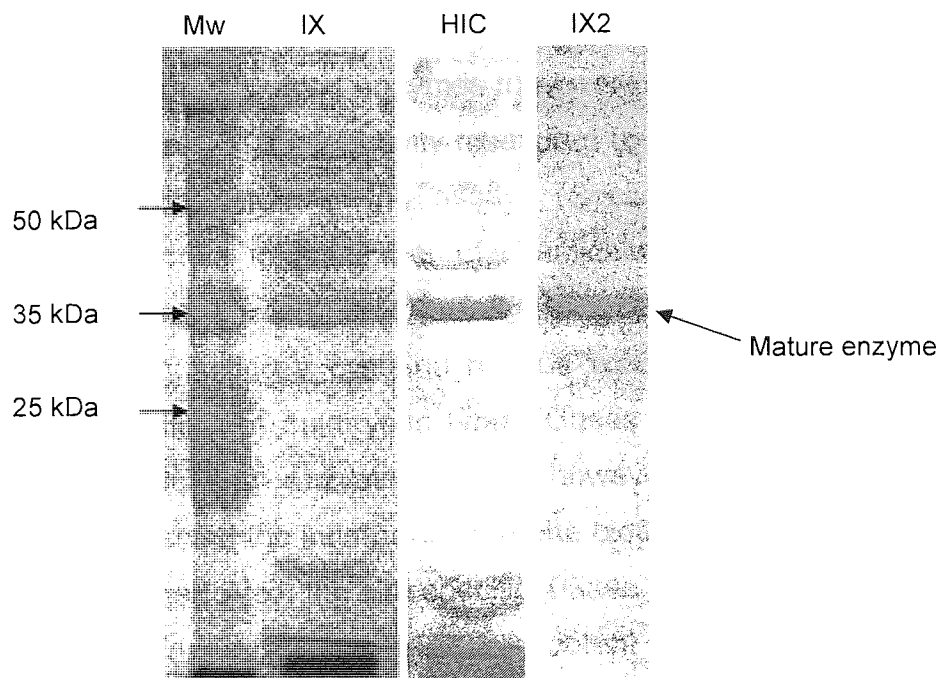
5.2.3. Biochemical Characteristics of Purified TGases

To determine the variability amongst purified TGase proteins their biochemical characteristics were determined in an attempt to highlight potentially divergent forms with altered phenotypic characteristics. Initially, assays involved the pH and temperature dependency of the enzymes. In addition, their activity profile in the presence of a number of inhibitors was also examined.

TGase from *S. mobaraensis* possesses pH and temperature optima of between 6 - 7 and 45 °C – 55 °C respectively (Kanaji et al. 1993). To determine pH and temperature optima for the purified TGases enzyme activity was measured by both the biotin cadaverine incorporation and the synthetic peptide cross-linking assays. Isolates exhibited similar temperature optima which were found to be ≈40 °C as determined by both the biotin cadaverine and peptide cross-linking assay (Figure 5.3a). Purified enzymes also shared similar activities under differing pH conditions with the exception of TGase from *S. baldaccii* that appeared more stable in lower pH conditions, retaining a higher proportion of its activity (Figure 5.3b).

To further characterise the purified enzymes they were treated with a variety of inhibitors and remaining TGase activity measured by the biotin cadaverine incorporation or peptide cross-linking assays (Sections 2.5.2 and 2.5.3 respectively). In response to a number of chemicals all TGase enzymes behaved in a similar manner (Table 5.1). Enzyme activity was susceptible to N-ethylmaleimide, suggesting a cysteine residue was involved in catalytic activity. Known activators of eukaryotic type TGases (DTT and Ca^{2+}) had little effect on the microbial enzymes tested, however, both primary and secondary amines affected the incorporation of the biotinylated substrate in a dose dependent manner. Also, the site specific TGase suicide inhibitor R281 markedly reduced TGase activity with R283, a potent inhibitor of eukaryotic TGases, having little or no effect.

Figure 5.3; Effect of Temperature and pH on the Activity of Purified

TGase Enzymes

Equal amounts of purified TGase proteins were subjected to both temperature (A) and pH (B) dependency utilising the biotin cadaverine TGase assay. Reactions were performed in triplicate on at least three separate occasions. Values are expressed as a percentage of the maximum reading (A450nm) recovered from individual enzyme assays.

(A)

Reactions were initiated by the addition of purified TGases and incubated at the required temperature for 1 hour before being processed and measured for enzyme activity as described in Section 2.5.2.

(B)
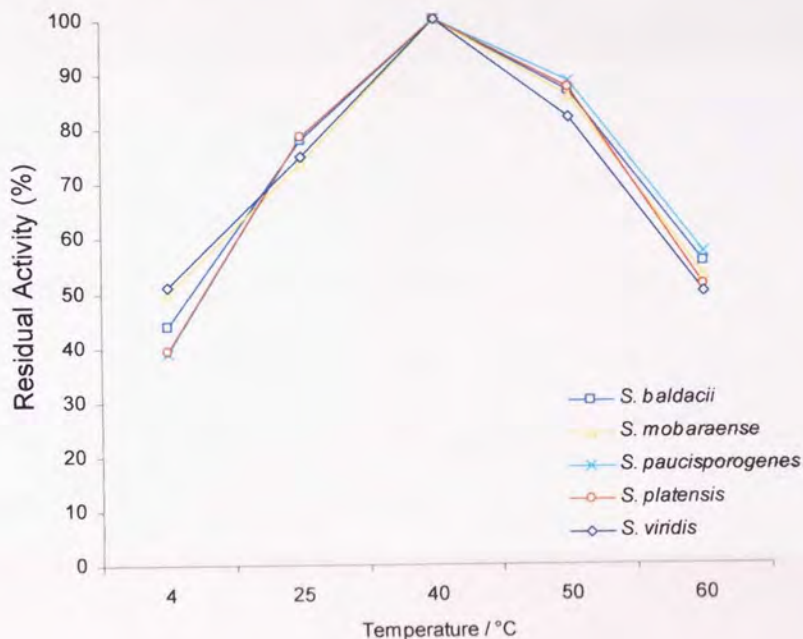
Biotin cadaverine was added to tris buffer (Tris.HCl for pH 4-7 and tris base for pH 7-10) of the relevant pH prior to incubation with equal amounts of purified TGase proteins as described in Section 2.5.2. The amount of TGase activity was then measured at the different pH values.
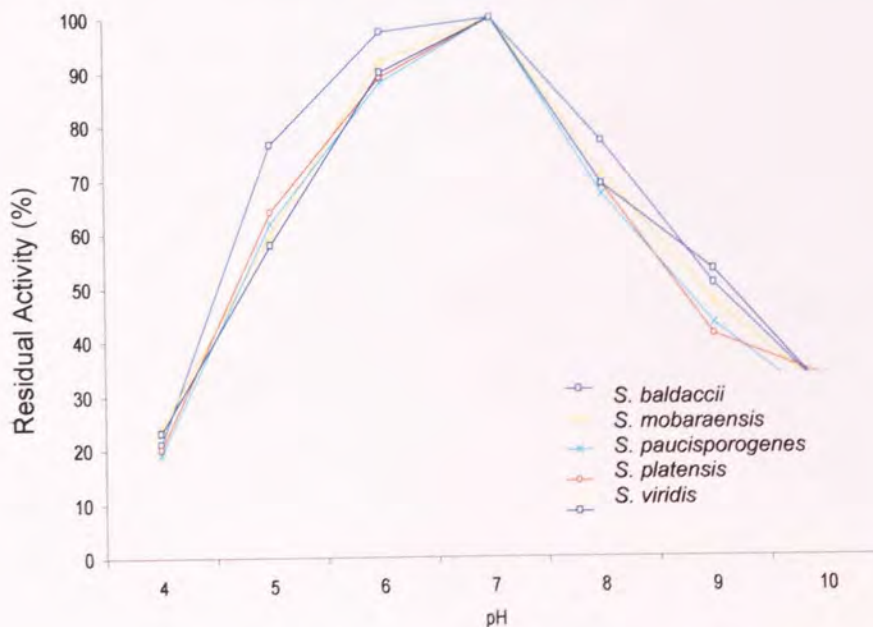
Table 5.1; The Effect of Various Chemicals on Streptomyces sp. TGases

Equal activity of purified TGases was incubated with the named chemicals at a concentration of 2 mM unless stated otherwise. After 1 hour incubation, at room temperature, remaining TGase activity was assayed by measuring the incorporation of biotin cadaverine into N`N`dimethyl casein as described in Section 2.5.2. 100% enzyme activity is equivalent to 5 unit of enzyme activity determined by the hydroxamate assay.

Chemical	Residual Activity (%)				
	S. mobaraensis	S. baldaccii	S. paucisporogenes	S. platensis	S. viridis
None	100	100	100	100	100
Cystamine	26	24	22	24	23
Hydroxylamine	24	22	24	24	23
R281 (250 µM)	28	30	33	42	31
R283 (1 mM)	96	92	95	94	97
CaCl$_2$	99	102	100	101	104
NEM	31	30	40	34	28

5.2.4. *Structural and Sequence Relationships between TGases of*

Streptomyces *and Related Genera*

5.2.4.1. *Degenerate Primer Design*

Streptomyces sp TGase nucleotide sequences publicly available, via the NCBI database, were aligned using MegAlign (DNAStar). Conserved regions encompassing the TGase active site were identified (Figure 5.4) and used to produce 2 degenerate primer pairs with minimum redundancy (Table 5.2). These primers were used to amplify the active site region by PCR (Section 2.11.6).

Figure 5.4; Conserved Regions Used for Degenerate Primer Design

Multiple alignment of TGases from *S. cinnamoneus* CBS 683.68 (Accession No. CAA70055), *S. platensis* (Accession No. AAS84612) and *S. mobaraensis* IFO13819 (Accession No. AF531437) showing the regions of homology (highlighted) that were used to design the degenerate primers for PCR.

Forward Primers

Reverse Primers

5.2.4.2. *Identification of TGase Active Site Domains*

Genomic DNA was isolated from organisms exhibiting TGase activity and used as the starting material for PCR. To aid primer binding the annealing temperature was reduced to 52 °C. All four primer combinations were used to identify products of the expected sizes listed in Table 5.2.

Table 5.2; Expected Degenerate PCR Product Size

Primer Combination	Expected Product Size (bp)
F1 + R1	450
F1 + R2	680
F2 + R1	375
F2 + R2	620

As expected, bands corresponding to the correct sizes were obtained from *S. baldaccii* NRRL B-3500, *S. platensis* ATCC 13865, *S. paucisporogenes* ATCC 12696 and *Saccharomonospora viridis* ATCC 15386, as shown in Figure 5.5, A-C. Despite TGase activity being detected in *S. albidoflavus* TUB B-706 on a number of occasions no bands were detected by degenerate PCR (Figure 5.5, D). Further optimisation of the reactions conditions failed to produce bands of relevant sizes despite a F2 primer specific product being amplified of ≈ 300 bp.

Figure 5.5; Amplification of TGase Active Site Regions by Degenerate PCR

Degenerate primers were designed around highly conserved regions, based on amino acid alignments of known *Streptomyces* derived TGase proteins. The primers were used to amplify the relevant coding region from TGase harbouring strains by PCR (as described in Section 2.11.6), decreasing the annealing temperature and increasing the magnesium concentration accordingly to aid primer binding. Cycling conditions used for PCR were: 94°C for 3 min followed by 25 cycles of 94°C for 30s, 50°C for 30s and 72°C for 30s.

PCR was initiated with the relevant primer pairs and the products analysed by gel electrophoresis using 1.5% agarose gels as described in Section 2.11.4.

A. *S.platensis* ATCC 13865
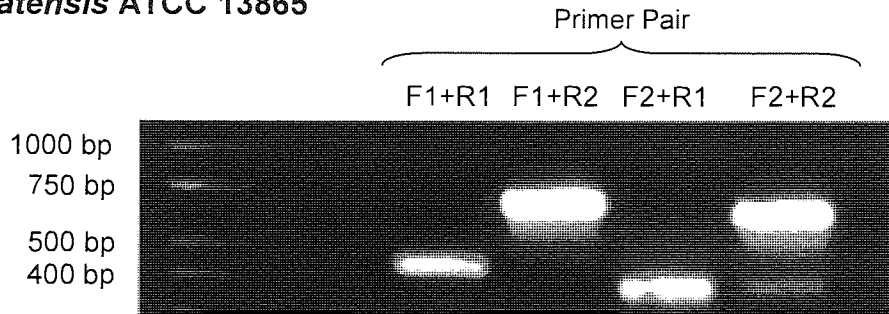
B. *S. baldaccii* NRRL 3500

C. *S. paucisporogenes* (S.p) and *Saccharomonospora viridis* (S.v) using degenerate primers F1 and R1
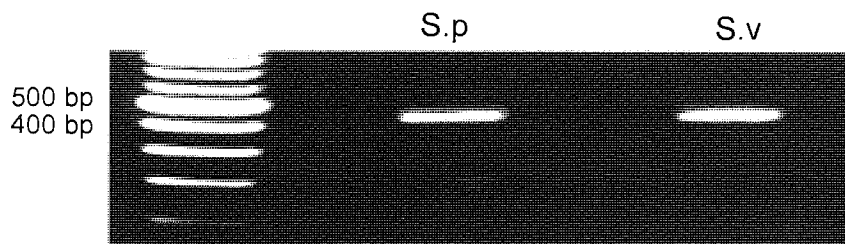
D. *S.albidoflavus* TUB B-706

172
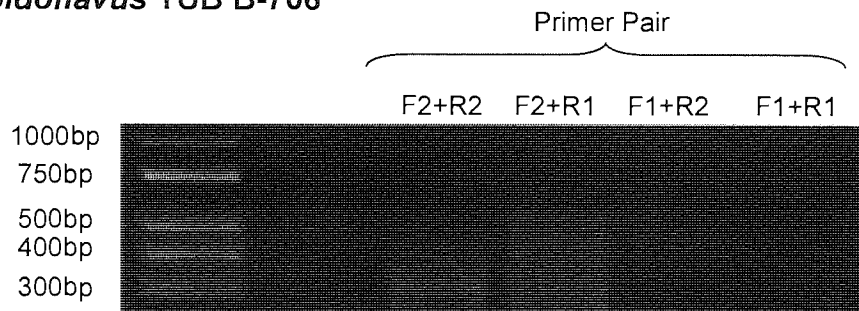
For the purpose of sequence alignments, bands obtained from the primer pairing F1 and R1 were excised and subcloned into pGEM-T (Promega) for DNA sequencing. Aligned sequences (Appendix VII) revealed high similarity ($\geq$ 90%) to the TGase sequence of *S. mobaraensis*. The TGase sequences from *S. baldaccii* showed the most variation (90% similarity) upon comparison. It was also revealed that the TGases encoded within *S. paucisporogenes* ATCC 12696 and *S. platensis* ATCC 13865 were >98% similar to each other and that TGase from *Saccharomonospora viridis* ATCC 15386 exhibited more than 99% similarity with TGase from *S. mobaraensis*, perhaps suggesting that either the organism has been miss-classified or that TGase has spread throughout the *Streptomyces* and related genera by horizontal transfer, an opinion that shall be discussed later in more detail.

5.2.5. *Potential Application of Purified TGases*

The commercial applications of TGases predominantly centres around their ability to modify proteins, this in part, is via the introduction of ε-(γ-glutamyl) lysine cross links between substrate proteins (Griffin et al. 2002). Therefore, if new TGases are to be of use in alternate industrial applications it is necessary to determine their ability to introduce said cross-links into a range of substrate proteins. Purified TGase enzymes from *S. baldaccii* NRRL B-3500, *S. platensis* ATCC 13865, *S. paucisporogenes* ATCC 12696 and *Saccharomonospora viridis* ATCC 15386 were assessed for cross-linking ability into both sodium caseinate and gelatin. The data obtained were then compared to that from commercially available preparations of the enzyme. Such preparations are marketed by Ajinomoto Co. Inc., Japan (*S. mobaraensis*) and Yiming Fine Chemicals Co. Ltd., China (organism not disclosed).

It can be seen that as incubation time increases the amount of aggregated, high molecular weight material increases for each enzyme analysed (Figure 5.6). There is also a clear concomitant reduction in the amount of substrate protein. Overnight incubation with TGase could possibly saturate the substrate protein with ε-(γ-glutamyl) lysine cross-links, making quantitation

difficult. For such quantification, incubations were performed for 4 hours prior to enzymatic digestion and determination of ε-(γ-glutamyl) lysine cross-links.

Analysis of protein cross linking data (Figure 5.7) revealed that TGase from *S. baldaccii* behaved in a similar manner to known TGases. TGases from *S. platensis* and *S. paucisporogenes* were found to behave identically; additionally they were able to incorporate a greater proportion of isopeptide cross-links per milligram protein as compared to TGase from *S. mobaraensis* and *S. viridis* (also found to encode identical proteins). TGase marketed by Yiming Fine Chemicals Co. Ltd performed very poorly with casein as substrate and didn't perform to the same level in the presence of gelatin as observed with the other mTGase preparations from *S. mobaraensis*, *S. baldaccii* and *S. platensis*.

Figure 5.6; Cross-Linking of Model Substrates by Microbial TGases

Equal amounts (50U), based on activity values determined by the hydroxamate assay, of each investigated enzyme were added to known amounts of sodium caseinate (1.5 mg/ml in PBS pH 7.4) and incubated for set time periods at 37 °C to allow cross-linking to occur. At 0, 5 min, 15 min, 30 min, 1 hour, 2 hours, 4 hours and 16 hours 30 µg protein was removed and analysed by SDS-PAGE (10% acrylamide), for polymer/aggregate formation. In each incidence the hashed line indicates the division between stacking and resolving gels.

(A) S. mobaraensis
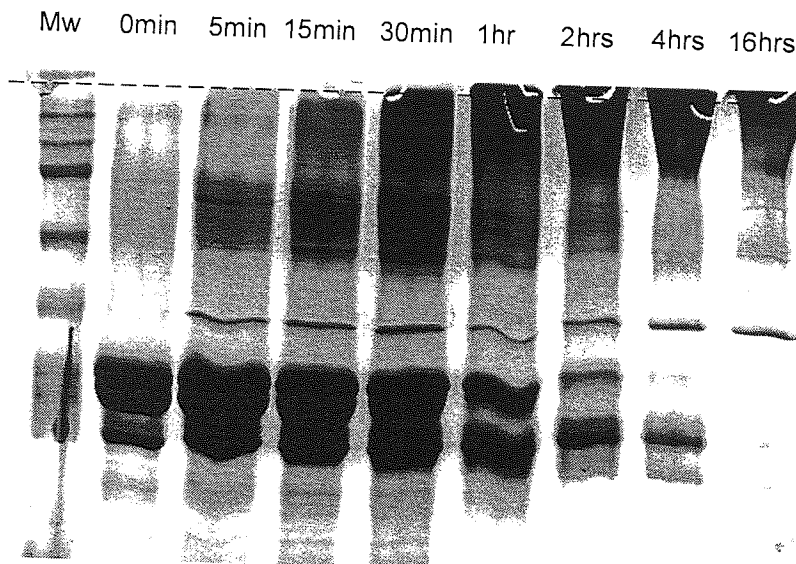
(B) S. baldaccii

(C) *S. paucisporogenes*

Mw   0min   5min   15min   30min   1hr   2hrs   4hrs   16hrs

(D) *S. viridis*
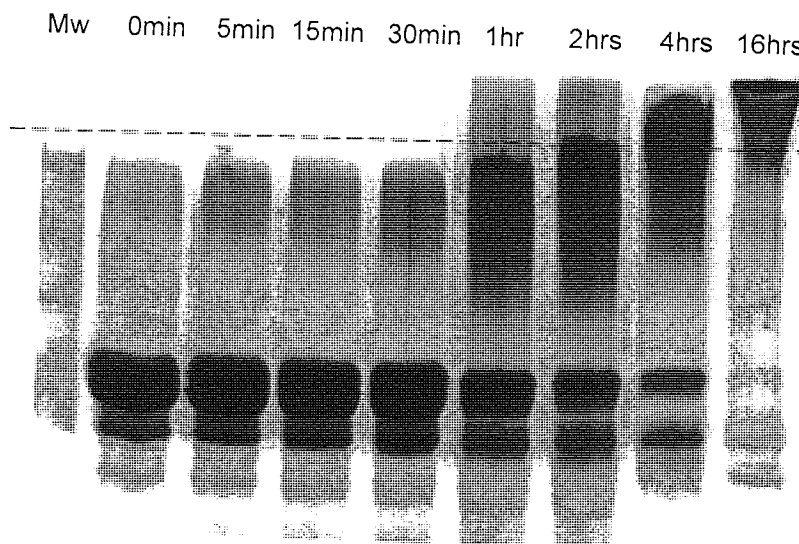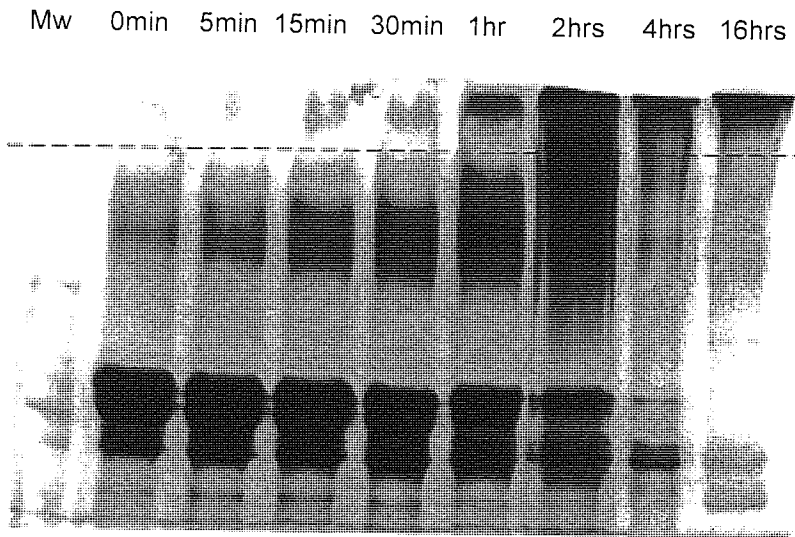
Mw   0min   5min   15min   30min   1hr   2hrs   4hrs   16hrs
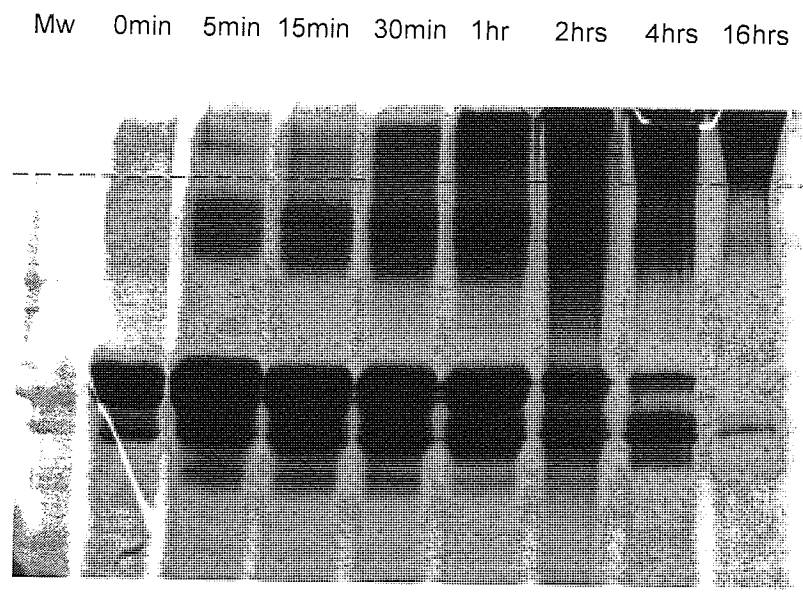
Figure 5.7; Analysis of TGase Mediated Protein Cross-linking

Cross linked casein (A) and gelatin (B) were proteolytically digested to release the ϵ(-γ glutamyl) lysine isodipeptide which was assayed by ion exchange chromatography as described in Section 2.10.1. Graphs represent data from two individual experiments.

(A)

(B)
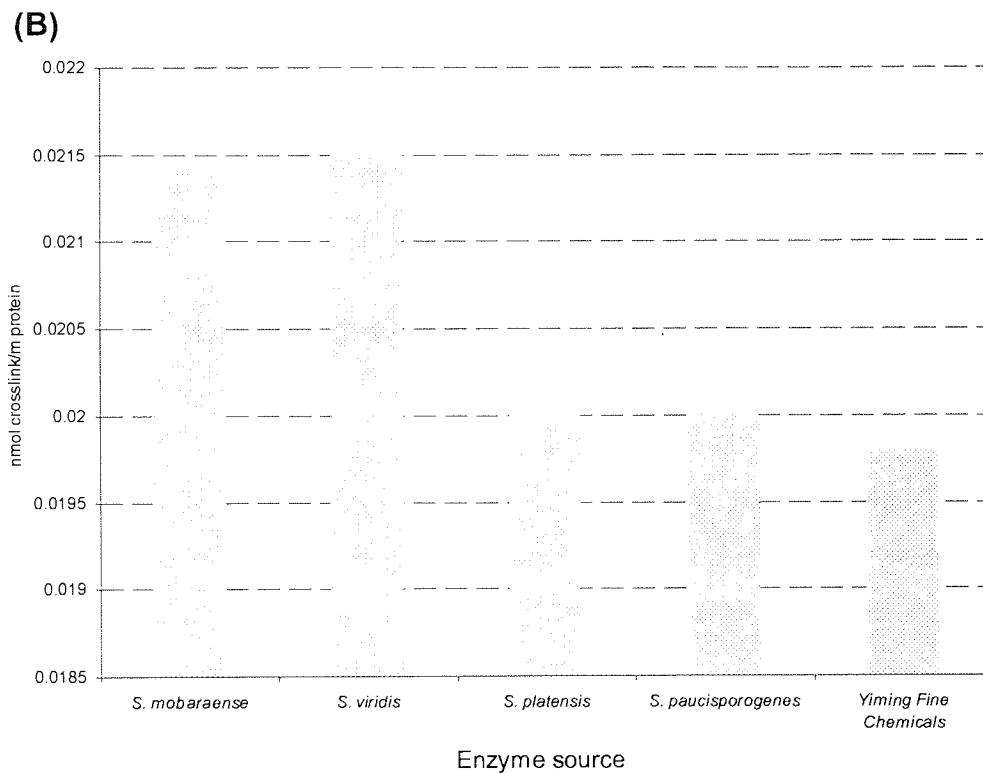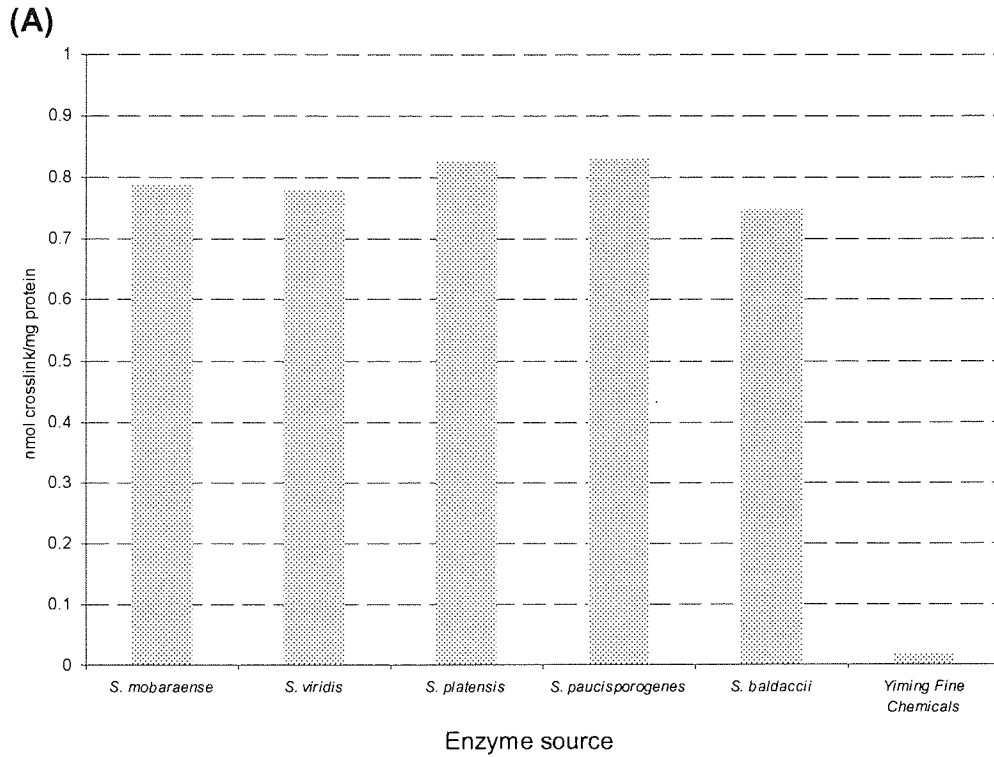
5.2.6. *TGase Derived from* S. baldaccii *NRRL B-3500*

TGase derived from *S. baldaccii* NRRL B-3500 has been shown to exhibit somewhat different enzymatic characteristics to the commercially available enzyme (*S. mobaraensis*). Degenerate PCR has highlighted the enzyme to be the most divergent with respect to its nucleotide sequence surrounding the active site. As such, the DNA sequence encoding TGase from *S. baldaccii* NRRL B-3500 was determined by PCR (Figure 5.8) (Personal communication, M. Paloheimo, ROAL, Finland). As expected, the leader sequence (pro-region) of *S. baldaccii* TGase showed a high degree of variability in both nucleotide and amino acid sequence compared with that of *S. mobaraensis* TGase. The gene showed a codon usage typical of actinomycete genes (Wright and Bibb 1992); as reflected by the high preference for G or C in the third base position of the codons.

The complete nucleotide coding sequence of *S. baldaccii* TGase (determined by M. Paloheimo, Roal, Finland) was found to share 84% identity with TGase from *S. mobaraensis* (Accession No. AF531437). Interestingly, the TGase encoded from *S. baldaccii* shares a higher identity with the TGases described in the organisms *S. cinnamoneus* (Accession No. AB085698) and *S. platensis* (Accession no. AY555726) 91% and 87% (Appendix VIII). It was found that from the deduced amino acid sequence, *S. baldaccii* TGase (mature enzyme) shared 80% identity with TGase from *S. mobaraensis* (Accession No. AF531437), 92% with TGase from *S. cinnamoneus* (Accession No. AB085698) and 85% with *S. platensis* (Accession No. AY555726).

To ensure the protein purified was encoded by the amplified DNA (Figure 5.8) peptide mass mapping was carried out (University of Helsinki) (Figure 5.9). The results of which confirmed that the protein under investigation was encoded by the nucleotide sequence amplified.

Complete TGase from *S. baldaccii* NRRL B-3500 was calculated to have an isoelectric point of 7.94 and a molecular weight of approximately 46.46 kDa (analysis carried out using DNAStar, Protean™). It is anticipated that cleavage between residues 87 and 88 by the endogenous TAMEP would release the mature enzyme, as has been previously reported in *S. mobaraensis* (Zotzel et al. 2003a). The mature enzyme was calculated to have a pI 7.61, with an approximate molecular weight of 37.65 kDa. In comparison, TGase from *S. mobaraensis* in its full form has a pI 6.98 and an approximate molecular weight of 45.68 kDa, and in its mature form has a pI 7.26 with an approximate molecular weight of 37.86 kDa, its molecular weight is similar to that described for *S. baldaccii* NRRL B-3500 TGase (analysis carried out using DNAStar, Protean™).

Figure 5.8; DNA Sequence and Deduced Amino Acid Sequence of TGase derived from S. baldaccii NRRL B-3500

Sequence determined by M. Paloheimo, Roal, Finland. Active site cysteine indicated in shaded area. * denotes the stop codon.

```
  1   ATGTACAAACGTCGGAGAATTCTCACCTTCGCCACTGCGGGTGCAGTCATATGC
  1     M   Y   K   R   R   R   I   L   T   F   A   T   A   G   A   V   I   C

 55   ACCGCCGGACTCATCCCGTCGGTCAGCCAGGCCGCCGGCAGCGGCGACAGGGGA
 19     T   A   G   L   I   P   S   V   S   Q   A   A   G   S   G   D   R   G

 99   GAGAAAGGGTCCTACGCCGAAACGCACGGCCTGACGGCGGATGACGTCAAGAAC
 37     E   K   G   S   Y   A   E   T   H   G   L   T   A   D   D   V   K   N

153   ATCAACGCACTCAACGAGAGCGCCCTGACTCCGGAGCAACCCGGCAGGTCACCA
 56     I   N   A   L   N   E   S   A   L   T   P   E   Q   P   G   R   S   P

207   GGGGAATTGCCTCCGAGTGCCAGCCCGTCCTCCCGGGCCCCCGGCTCTACCGAC
 74     G   E   L   P   P   S   A   S   P   S   S   R   A   P   G   S   T   D

261   GACAGGGAAACCCCTCCGGCCGAGCCGCTCGAAAGGATGCCTGATGCGTACCGG
 92     D   R   E   T   P   P   A   E   P   L   E   R   M   P   D   A   Y   R

315   GCCTACGGAGGCAGGGCCACTACGGTCATCAACAACTACATACGCAAGTGGCAG
110     A   Y   G   G   R   A   T   T   V   I   N   N   Y   I   R   K   W   Q

369   CAGGTCTACAGTCACCGCGACGGCAAGAAGCAGCAAATGACCGAAGAGCAGCGA
128     Q   V   Y   S   H   R   D   G   K   K   Q   Q   M   T   E   E   Q   R

423   GAAAAGCTGTCCTACGGTTGCGTCGGCGTCACCTGGGTCAATTCGGGTCCCTAC
146     E   K   L   S   Y   G   C   V   G   V   T   W   V   N   S   G   P   Y

477   CCGACGAACAAATTGGCGTTTGCGTCCTTCGACGAGAACAAGTACAAGAACGAC
164     P   T   N   K   L   A   F   A   S   F   D   E   N   K   Y   K   N   D

531   CTCAAAAACATCAGCCCCCGACCCGGGGAAACGCGGGCAGAGTTCGAGGGCCGC
182     L   K   N   I   S   P   R   P   G   E   T   R   A   E   F   E   G   R

585   ATCGCCAAGGACAGTTTCGACGAGGAGAAGGGCTTCAAGCGGGCGCGTGATGTG
200     I   A   K   D   S   F   D   E   E   K   G   F   K   R   A   R   D   V

639   GCGTCCATCATGAACAAGGCCCTGGAAAGCGCCCACGACGAGGGGGCGTACATC
218     A   S   I   M   N   K   A   L   E   S   A   H   D   E   G   A   Y   I

693   GACAACCTCAAGACGGGGCTCACGAACAACAATGACGCTCTGCTCTACGAGGAC
236     D   N   L   K   T   G   L   T   N   N   N   D   A   L   L   Y   E   D

747   AGCCGATCGAACTTCTACTCGGCACTGAGGAACACACCGTCTTTCAAGGAAAGA
254     S   R   S   N   F   Y   S   A   L   R   N   T   P   S   F   K   E   R

801   GACGGAGGCAACTACGACCCGTCCAAGATGAAGGCGGTGATCTACTCGAAGCAC
272     D   G   G   N   Y   D   P   S   K   M   K   A   V   I   Y   S   K   H

855   TTCTGGAGCGGGCAGGACCAGCGGGGCCCTTCTGACAGGAGGAAGTACGGCGAT
290     F   W   S   G   Q   D   Q   R   G   P   S   D   R   R   K   Y   G   D
```

```
909  CCGGAAGCCTTCCGCCCCGCCCAGGGTACCGGCCTGGTCGACATGTCGAAGGAC
308   P   E   A   F   R   P   A   Q   G   T   G   L   V   D   M   S   K   D

963  AGAAGCATCCCGCGCAGTCCCGCCAACGCCGGCGAAAGCTGGGTCAATTTCGAC
326   R   S   I   P   R   S   P   A   N   A   G   E   S   W   V   N   F   D

1017 TACGGCTGGTTCGGGGCTCAAACGGAAGCGGATGCCGACAAGACGGTATGGACC
344   Y   G   W   F   G   A   Q   T   E   A   D   A   D   K   T   V   W   T

1071 CACGGCGACCACTACCACGCGCCCAATAGTGGCCTGGGCCCCATGCACGTATAC
362   H   G   D   H   Y   H   A   P   N   S   G   L   G   P   M   H   V   Y

1125 GAGAGCAAGTTCCGGAACTGGTCTGCCGGGTACGCGGACTTCGACCGCGGAACC
380   E   S   K   F   R   N   W   S   A   G   Y   A   D   F   D   R   G   T

1179 TACGTGATCACGTTCATACCCAAGAGCTGGAACACCGCCCCCGCCAAGGTGCAG
398   Y   V   I   T   F   I   P   K   S   W   N   T   A   P   A   K   V   Q

1233 CAGGGCTGGCCGTAA
402   Q   G   W   P   *
```

Figure 5.9; Confirmation of S. baldaccii NRRL B-3500 TGase by Peptide Mass Mapping

Sequenced fragments (shown shaded) based on the deduced amino acid sequence of TGase from *S. baldaccii* NRRL B-3500. Active site region indicated in boxed area.

```
  1   GSTDDRETPPAEPLERMPDAYRAYGGRATTVINNYIRKWQQVYSHRDGKK

 51   QQMTEEQREKLS YGCVG VTWVNSGPYPTNKLAFASFDENKYKNDLKNISP

101   RPGETRAEFEGRIAKDSFDEEKGFKRARDVASIMNKALESAHDEGAYIDN

151   LKTGLTNNNDALLYEDSRSNFYSALRNTPSFKERDGGNYDPSKMKAVIYS

201   KHFWSGQDQRGPSDRRKYGDPEAFRPAQGTGLVDMSKDRSIPRSPANAGE

251   SWVNFDYGWFGAQTEADADKTVWTHGDHYHAPNSGLGPMHVYESKFRNWS

301   AGYADFDRGTYVITFLRKSWNTAPAKVQQGWP
```

5.2.7. *Molecular Cloning of TGase from* S. baldaccii *NRRL B-3500*

5.2.7.1. *Expression Vector*

Due to low yields and the expense of conventional purification (from the culture broth), it would be advantageous to have a recombinant enzyme which was readily producible at a more reasonable cost. Expression in *E. coli* is the most simple and commonly used system and as such TGase from *S. baldaccii* NRRL B-3500 was to be amplified by PCR, cloned and expressed as a GST fusion protein utilising the vector pGex2T (Amersham Bioscience, UK).

5.2.7.2. *Preparation of* S. baldaccii *NRRL 3500 TGase Expression Construct*

Primers BalF and BalR (Table 2.3) were designed to amplify the mature region of TGase from *S. baldaccii* NRRL B-3500 (Figure 5.10), based on the nucleotide sequence detailed in Figure 5.8. The coding region was amplified by PCR using an annealing temperature of 56 °C and an extension time of 2 min in the presence of 2 mM $MgCl_2$. The resulting single band, corresponding to the correct size (996 bp) was excised, purified and subcloned into pGEM-T (Promega, UK). Resulting blue colonies were screened for the presence of the desired DNA by restriction analysis of recovered pGEM vector from transformants (Figure 5.11). The DNA encoding *S. baldaccii* TGase was then cloned into the expression vector pGEX2T (Amersham Bioscience, UK) (Figure 5.12).

Figure 5.10; PCR Amplification of mature TGase from S. baldaccii NRRL 3500

Genomic DNA was isolated and used to amplify the mature region of TGase from *S. baldaccii* NRRL B-3500 according to the procedures described in Section 2.11.6 using an annealing temperature of 56 °C and extension time of 90s.

Lane 1; Molecular weight marker, Lane 2; PCR amplified *S. baldaccii* TGase.
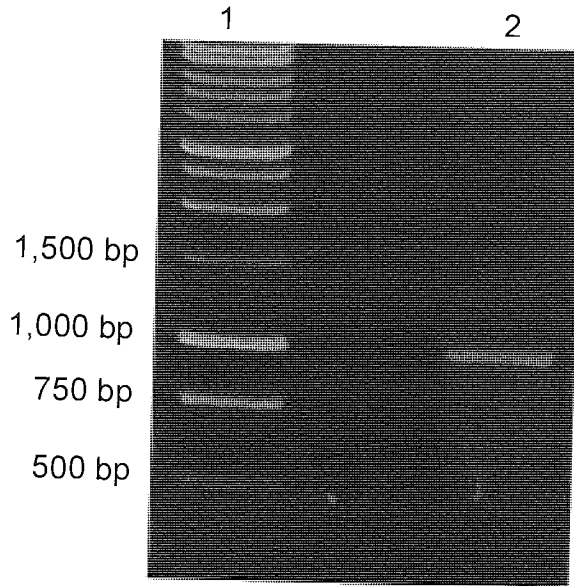
Figure 5.11; Restriction of pGEM-T to Release S. baldaccii TGase

Restriction digest showing *S. baldaccii* TGase insert ($\approx$1,000 bp) and pGEM-T vector ($\approx$3,000 bp). Restriction digestion was carried out as described in Section 2.11.8.

Lane 1; Molecular weight marker, Lane 2; pGEM-T (BaldTG) BamHI: EcoRI
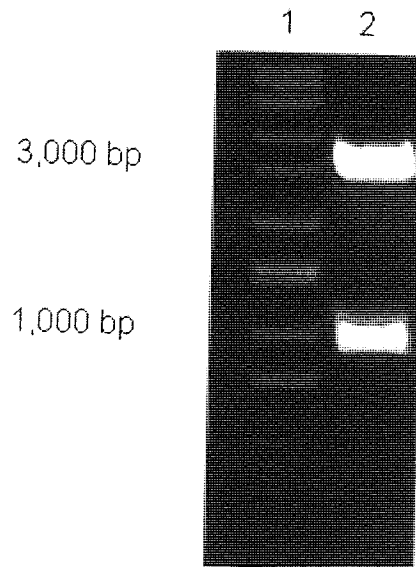
Figure 5.12; Presence S. baldaccii *TGase Gene in pGEXBaldTG*

Post transformation, colonies obtained were screened for the presence of the inserted gene by restriction analysis, using enzymes *EcoR*I and *BamH*I, as described in Section 2.11.8.

Lane 1; Molecular weight marker, Lane 2; Linearised pGEXBaldTG (*BamH*I restricted), Lane 3; pGEXBaldTG *BamH*I: *EcoR*I restricted.
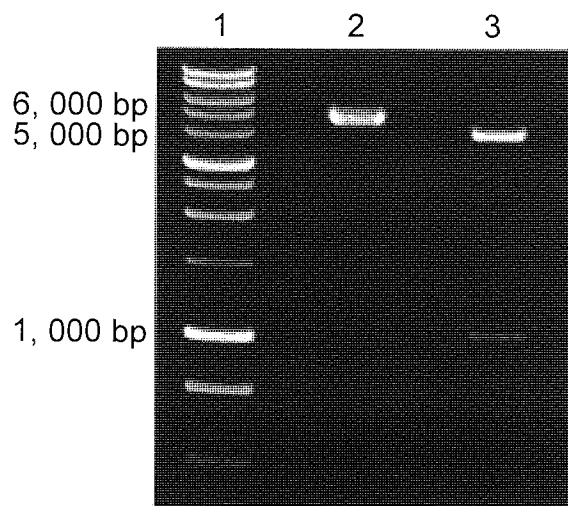
5.2.7.3. Analysis of Recombinant TGase Expression

Plasmid pGEXBaldTG was transformed into *E. coli* Tuner™ (Novagen, UK), allowing IPTG to be titrated to give optimal expression conditions. A series of experiments were performed to determine the optimum conditions for induction with regards to temperature and induction length.

Confirmation of the TGase-GST fusion protein was sought by both activity assays (hydroxamate) and Western blotting, probing with the commercially available anti-body against mTGase (N-zyme Biotech, Germany). Analysis by Western blotting (Figure 5.14) revealed a single band corresponding to that obtained by SDS-PAGE ($\approx$64kDa).

Figure 5.13; SDS-PAGE Analysis of Expressed Recombinant TGase
from **S. baldaccii** *NRRL B-3500*

The plasmid pGEXBaldTG was transformed into *E. coli* Tuner ™ (Novagen, UK) and grown to A_{600} = 0.7 prior to induction, by the addition of IPTG to a final concentration of 1 mM. Aliquots were removed at 1, 2 and 4 hours post induction and analysed by SDS-PAGE using 10% polyacrylamide gels. GST, encoded by the plasmid pGEX2T, was expressed as positive control.

MwM; Molecular weight markers, U; uninduced control, 1,2 and 4 represent hours after induction.
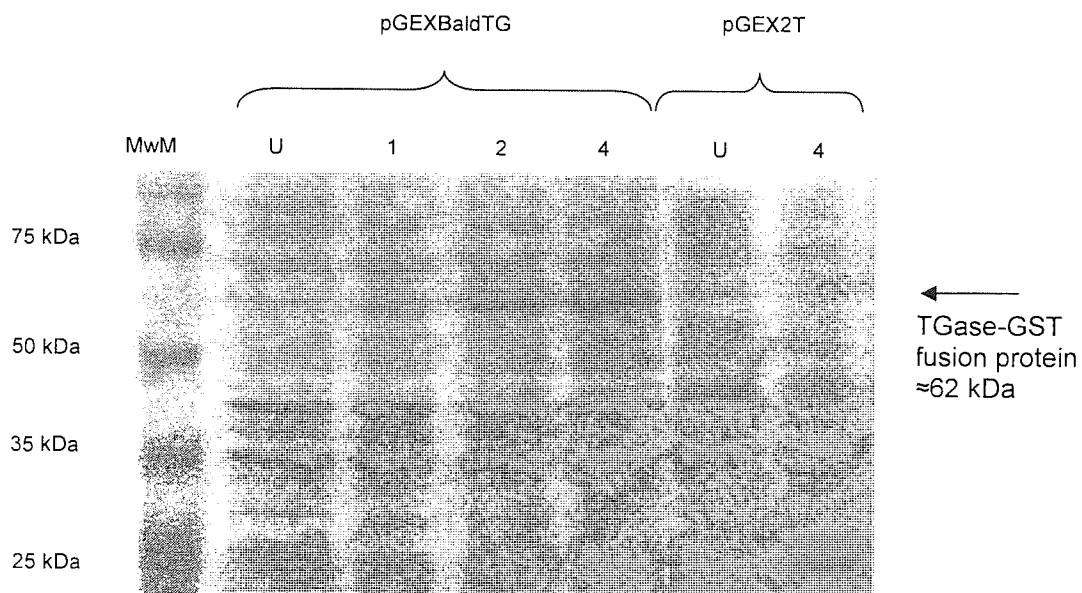
Figure 5.14; Confirmation of Recombinant TGase Production by Western Blotting

Induced *E. coli* containing pGEX2BaldTG was analysed by SDS-PAGE (A) and Western blotting (B) for the presence of recombinant TGase as described in Section 2.9.3. In addition, samples of uninduced *E.coli* pGEXBaldTG (U) were also analysed along with 5 µg *S. mobaraensis* TGase (+ve) and the lysate generated from induced *E. coli* containing pGEX2T (-ve).

(A)
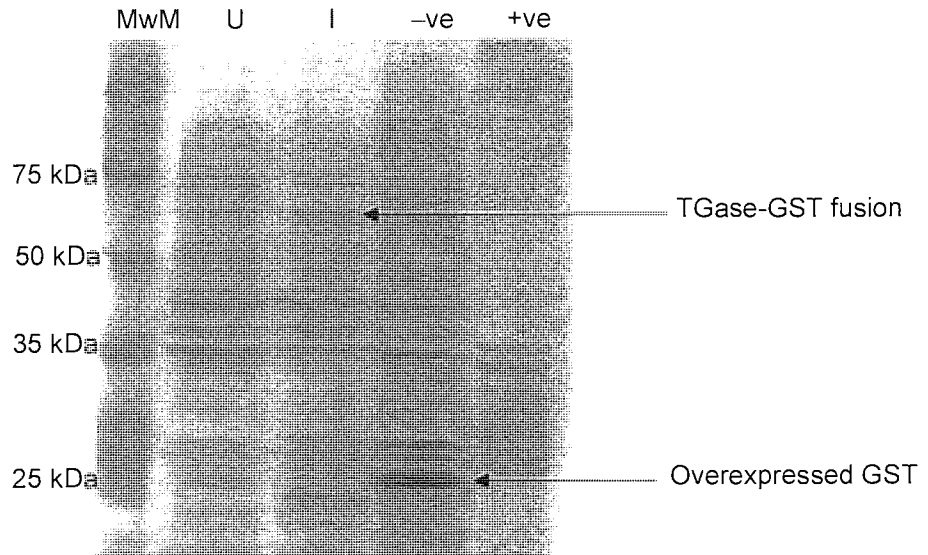
(B)
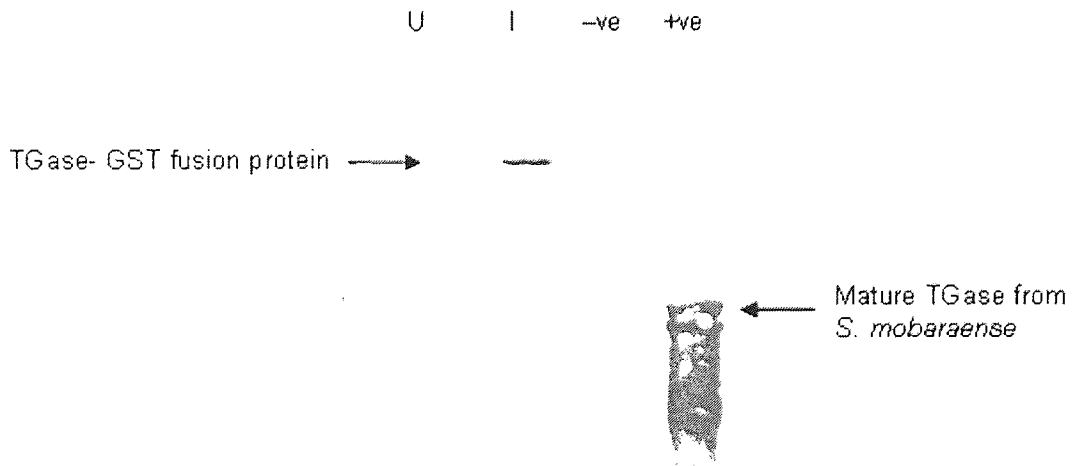
Figure 5.15; Assessment of TGase Activity from Induced E.coli

Harbouring Expression Plasmid pGEXBaldTG

Expressed protein was produced as described in Section 2.11.15. Cells were collected, from 2 ml of induced culture, by centrifugation at 15,000 x g for 2 min and lysed by the addition of Bugbuster (Novagen, UK) according to the manufacturers' instructions. Generated lysates were assayed for their TGase activity by the hydroxamate assay with a 30 min incubation time before quenching.
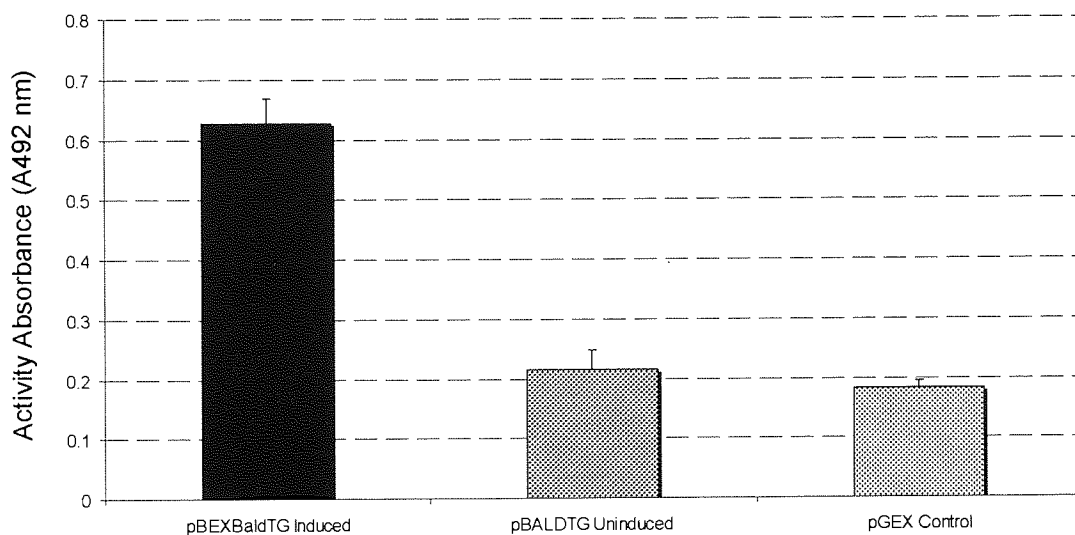
5.2.8. Physiological Studies on TGase from Streptomyces sp.

To date the exact biological role of *Streptomyces sp.* TGase remains to be defined, though it has been postulated to be involved in cellular growth and morphological differentiation (Kashiwagi et al. 2002a). In an attempt to decipher the enzymes' role in cellular growth a number of experiments were carried out.

After the organism *S. baldaccii* had been grown to produce maximum TGase, the fluorescent primary amine substrate monodansyl cadaverine (2 mM) was added and incubated for a further 1 hour at 30°C with shaking at 225 rpm. Cells were collected by centrifugation, washed three times with PBS and mounted on a microscope slide for confocal microscopy. Upon examination, *S. baldaccii* was found to exhibit a large amount of auto-fluorescence; an attribute echoed by other *Streptomyces sp.* examined limiting the usefulness of such a technique to identify TGase substrates in the microbes' cell wall. The large amount of autofluorescence ruled out the use of fluorescent primary amines to investigate amine incorporation into the organism.

To elucidate a role in cellular growth the TGase site specific inhibitor, R281, was used to inhibit any TGase enzyme secreted by the organisms during study. As can be seen in Figure 5.16, R281 was found to have a profound effect on the cells of *S. baldacci,* when compared to the non-TGase producing *S. badius*, with increasing concentrations reducing both cellular growth as well as the appearance of surface colonies. Removal of the inhibitor enabled "normal" growth to resume. This would indeed confirm the function of TGase in morphological differentiation assuming no other toxic effects are occurring.

Figure 5.16; Effect of R281 on Submerged Streptomyces *Cultures*

Equal amounts of S. *baldaccii* (A) and S. *badius* (B) were inoculated into 5 ml GYM medium, in a 6-well plate, supplemented with R281 and grown statically at 30 °C for 5 days with fresh R281 being added every 48 hours. The effect on cellular growth was noted.

(A)
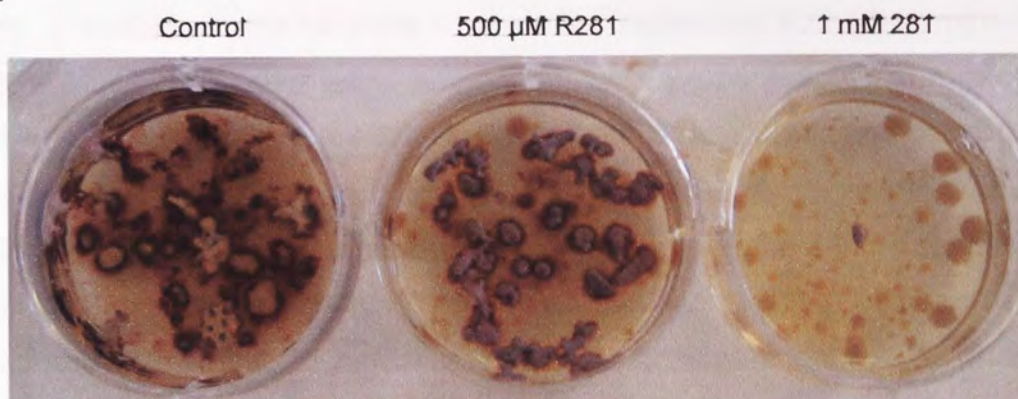
(B)

5.3. *Discussion*
5.3.1. *Identification of TGases from* Streptomyces sp.

TGases were identified and purified from the organisms *S. platensis, S. paucisporogenes* and *S. baldaccii*. Each enzyme was produced as an inactive zymogen that was proteolytically processed to release the active form. Enzymes produced were of a similar molecular size (determined by SDS-PAGE) and possessed similar catalytic properties. Of the purified proteins, the enzyme present in the filtrate of *S. baldaccii* was found to be most diverse with respect to enzyme characteristics (Section 5.2.3), exhibiting better performance at lower pH, though temperature dependence remained similar to other *Streptomyces* derived TGase enzymes. The genetic sequence of *S. baldaccii* TGase was determined (M. Paloheimo, Roal, Finland) and aligned with other TGase-encoding genes and was found to share 80% identity with *S. mobaraensis* TGase. The genetic differences are likely to be responsible for the enzymes altered enzymatic properties. In addition, TGase from *S. paucisporogenes* was found to be identical to that of *S. platensis*, and TGase from *S. viridis* was found to be identical to that from *S. mobaraensis*, as such the enzymes behaved in a similar manner in the tests undertaken.

TGase activity was identified in the organism *S. albidoflavus* (see Section 4.2.3.2). The protein responsible was found to cross-react slightly with the mTGase anti-body at the desired molecular weight. Attempts to further purify and characterise the protein responsible were unsuccessful since apparent enzyme activity disappeared with successive culturing. This could perhaps indicate that the culture medium used for production was sub-optimal, as such; different media were trialled without any effect on TGase productivity being observed. Additionally, analysis of *S. albidoflavus* genomic DNA by degenerate PCR failed to produce any related TGase sequence. Perhaps a distinct TGase was present in the species, however, it was deemed more likely that there was a TGase producing contaminant present in the initial culture which was removed by successive culturing.

5.3.2. *Physiological Studies on TGase from Streptomyces sp.*

In each case maximum TGase activity was recovered towards the end of exponential growth/start of stationary phase. This would suggest a potential role of TGase in cellular differentiation put forward by Kashiwagi (Kashiwagi et al. 2002a). Attempts to identify cellular substrate proteins, by incubating cultures with the fluorescent primary amine substrate, monodansyl cadaverine, with subsequent analysis by confocal microscopy were unsuccessful. *Streptomyces sp.* have been shown to exhibit identical gene expression cascades whether grown in liquid or on solid media (Fernandez et al. 2002). Known TGase producing *Streptomyces sp.*, grown in solid state fermentation, produce higher yields of enzyme when compared to liquid growth. In such situations TGase from *S. platensis* was found to outperform *S. mobaraensis* with regards to enzyme yield (Personal Communication, G.Szakacs, Technical University of Budapest). This could suggest that TGase, despite being produced in liquid culture, is not necessary for liquid growth. TGase activity, therefore, is more likely to be associated with the transition between vegetative and sporulation states observed when isolates are grown on solid phase. It is possible that TGase may play a role in the development of the aerial hyphae by either cross-linking hyphal proteins or by aiding the recruitment of cellular debris for hyphal construction (Chater 1989). Attempts to induce TGase production by encouraging sporulation in liquid medium (by the addition of an iron chelator) proved unsuccessful.

TGase may potentially play a role in mycelial development since mycelial growth was restricted in the presence of varying concentrations of the TGase inhibitor R281. This is an effect that was not seen with non-TGase producing *Streptomyces sp.* (*S. badius*). Investigations into such areas could perhaps elucidate the exact physiological role of TGase. Since TGase activity was only identified in a few *Streptomyces* species the advantage conferred by possessing such an enzyme remains intriguing.

5.3.3. *Ubiquitous Distribution of mTGase Amongst* Streptomyces?

The *Streptomyces* genera incorporates not only the *Streptomyces* but also closely related species (defined on taxonomical grounds) including; *Chainia; Elytrosporangium; Kitasatoa; Kistaosporia; Actinosporangium* and *Streptoverticillium* (Embley and Stackebrandt 1994). The genus is separated into a number of clusters (major and minor) and overall members of the genus share greater than 70% similarity by the S_{SM} coefficient (Williams et al. 1983). Despite this high degree of similarity between members and the large amount of Actinomycetes screened as part of this study, TGase was only recovered from a limited number of species.

If we consider TGase producing *Streptomyces sp.*, there would appear to be a somewhat scattered distribution of TGase throughout a number of cluster groups (Figure 5.17). There would appear to be a more localised distribution amongst the whorl forming streptomycetes (formerly *Streptoverticillium*) perhaps suggesting the origin of TGase, prior to its lateral transfer. If this is the case then one might expect TGases from the former Streptoverticillia to be the most divergent, assuming a constant and equal selective pressure and mutation rate. In the case of *S. baldaccii* and *S. cinnamoneum* (both formerly *Streptoverticillium*) this would appear to be the case since the enzymes encoded share the least similarity with that from *S. mobaraensis*. The differences in TGase sequences supplies evidence that although the *Streptoverticillium* are closely related to the *Streptomyces* (Hatano 2003) they still possess certain phenotypic traits that may be unique to their genus. One might therefore expect all former Streptoverticillia to possess TGase activity though this is not the case. It is therefore likely that during evolution TGase function has evolved from an ancestral protease (Kashiwagi et al. 2002a) in a member of the Streptoverticillia and over time, through horizontal transfer between related species, the gene has slowly disseminated.

TGase activity was not observed in all Streptomyces nor in every former Streptoverticillia. No gene of significant similarity was observed searching the *S. coelicolor* genome for TGase homologues. One could hypothesise that the enzyme carries out a non-essential function that does not confer a significant

evolutionary advantage; alternatively the organisms may already encode a protein that carries out a similar function by a different pathway or the TGase encoded may act only on specific amino acid residues that are located within defined protein conformations, as is the case with CNF1 and DNT (Lerm et al. 1999) and hence would not be detected using the assays in this work. This may explain the low occurrence of Streptomyces TGase. Conversely, microbial TGase may have evolved relatively recently to fulfil a particular biological function and is in the process of being passed on therefore representing evolution in action. Further study is needed on this unique enzyme to elucidate its biological and evolutionary significance.

Figure 5.17; Simplified Dendrogram Showing the Relationships between Some of the Major and Minor Clusters Defined at the 63% Similarity Level in the S_J/UPGMA Analysis

Adapted from Williams (Williams 1983). Clusters comprising organisms exhibiting TGase activity are indicated by a red arrow. Former *Streptoverticillium* highlighted in blue.
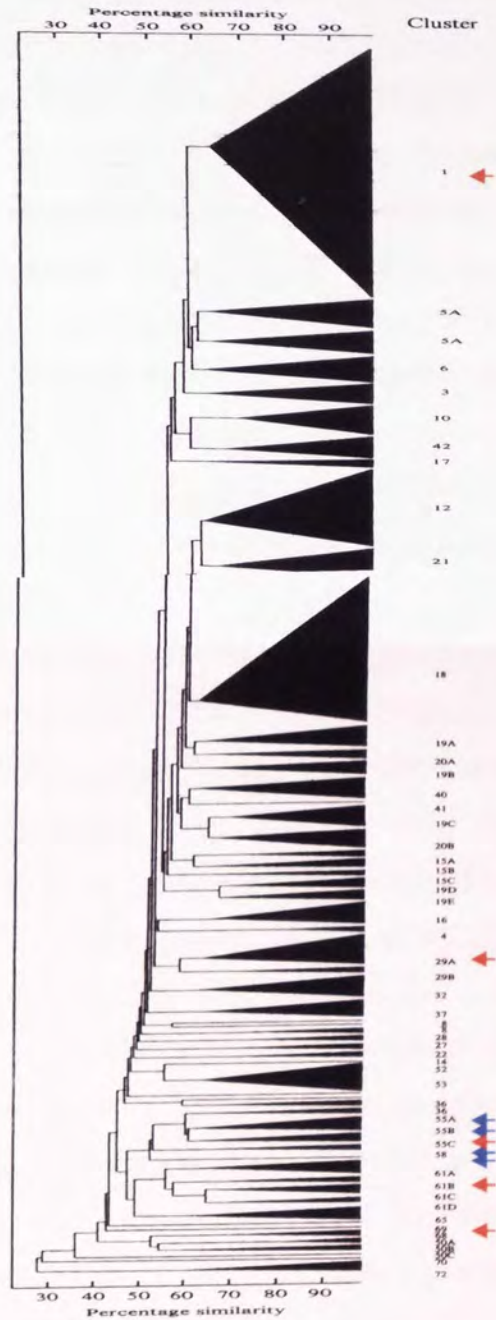
5.3.3.1. *Recovery of TGase from Closely Related Genera*

TGases were purified from a number of closely related Streptomyces species which were found to exhibit a high degree of identity with TGase from *S. mobaraensis*. TGase has been purified from the family Pseudonocardiaceae, from the organism *Saccharomonospora viridis* (Figure 5.2B). The TGase encoded therein was found to be identical to that from *S. mobaraensis*, analysed by degenerate PCR spanning the active site region. The overall genetic sequence was found to exhibit greater then 99% identity to that of *S. mobaraensis* TGase (personal communication M. Paloheimo). At the protein level TGase production proceeded as follows; the inactive zymogen was sequentially secreted followed by subsequent activation to release mature TGase that showed full enzymatic activity. Further analysis confirmed that the protein was an identical size to that from *S. mobaraensis* and shared many of the same characteristics; including pH and temperature optima. *S. viridis* TGase behaved identically to *S. mobaraensis* TGase in the presence of the substrates casein, gelatin and BSA (Figure 5.7) analysed by the ability of the enzymes to introduce intra and intermolecular cross-links.

Initially it was thought that the organism *S. viridis* had been generally mis-classified. However; on the basis of differential growth characteristics and colony morphology the organisms were found to be distinct. Members of the genus *Saccharomonospora* have been found to display only small differences (<1%) in the 5` terminal region of 23S rRNA with representatives of the *Streptomyces* genera (Ruan et al. 1994). It would appear likely, therefore, that *S. viridis* has acquired the TGase gene through horizontal transfer of genetic information, probably from *S. mobaraensis*, though it must be noted that the exact donor organism is unknown. The minimal differences in genetic sequence would indicate that this process has occurred relatively recently. Several mechanisms exist by which prokaryotes may acquire genetic information; conjugal transfer, phage-mediated insertion and by uptake of exogenous DNA from the external environment (Ochman et al. 2000; Redfield 2001). Such horizontal gene transfer plays a major role in organism evolution, especially towards the early stages (Feng et al. 1997). Only genes that confer an evolutionary advantage may persist in the recipient organism (Ochman and 2005) though the exact advantage conferred by

possessing TGase remains undetermined especially as is isn't identified in all *Streptomcyes* sp..

5.3.4. *Applications of New TGase Enzymes*

In many industrial processes microbial TGase is used as a means of crosslinking protein rich substrates (Section 1.5.2). To date microbial TGase is marketed by Ajinomoto Co. Ltd. (Japan), and Yiming Fine Chemicals Co. Ltd. (China). Comparison of the enzyme preparations revealed that the Ajinomoto Co. Ltd enzyme performs substantially better than its counterpart manufactured by Yiming Fine Chemicals Co. Ltd. by introducing a larger proportion of ε-(γ-glutamyl) lysine cross-links per milligram protein substrate (sodium caseinate and gelatin) (Figure 5.7). Interestingly, TGases purified from *S. platensis* and *S. paucisporogenes* (found to produce identical proteins) were found to be introduce a higher proportion of ε-(γ-glutamyl) lysine cross-links into casein (1.05 fold increase) though the enzyme incorporated less cross-links into other protein substrates, namely gelatin. Despite the increase in cross-link being small, applied on an industrial scale, the application of TGase from *S. platensis* or *S. paucisporogenes* would equate to a substantial saving in enzyme preparation and purification costs, though this would depend on the enzymes affinity for the substrate protein and of course the occurrence on the protein. The enzymes purified may also exhibit greater performance in the presence of alternative protein substrates though this waits to be investigated.

TGase has been purified from the organism *S. baldaccii*. This enzyme has been shown to possess slightly different enzyme characteristics whilst still incorporating comparable levels (to *S. mobaraensis*) of ε-(γ-glutamyl) lysine cross-links into proteins. Such characteristics may confer an industrial advantage over similar forms of the enzyme when applied to distinct processes thereby making TGase from *S. baldaccii* more cost effective. A critical aspect of using enzymes in industrial processes is their mass production. To date Streptomyces TGase has been expressed in a number of homologous and heterologous systems with differing degrees of success (Section 1.5.3). Most recently TGase from the organism *S. fradiae* has been

purified and mass produced by both expression in the original strain (Liu et al. 2006b) and by purification from inclusion bodies produced in *E. coli* (Liu et al. 2006a) with the former yielding a higher amount of active enzyme, twice that originally produced by the organism. Such methods could be utilised to increase the yield of enzyme, though the conventional fermentation and purification processes would ultimately escalate costs in comparison to a recombinant enzyme. Also a recombinant enzyme has the distinct advantages of ease of handling, simple growth medium with easily reproducible scale-up possibilities. The downstream purification procedures are also cheaper and quicker resulting in less enzyme being lost as a result of degradation and loss during the purification procedure.

The mass production of TGase, at a reasonable cost, would appear to be a major problem for industrial applications. Since yields obtained from conventional fermentation are relatively low, it would be advantageous to have an easily producible recombinant enzyme. As such, TGase from *S. baldaccii* was initially cloned and expressed as a fusion protein with glutathione-S-transferase (see Section 5.2.7.3). Growth of *E. coli* avoids the problems that may arise with complex dimorphic cultures. The protein expressed was found to exhibit TGase activity demonstrating the ease with which recombinant TGase could be manufactured for use in industry with the associated reduced handling and purification costs, though the process may require some optimisation. An enzyme with such characteristics may therefore offer a viable alternative to the commercially available brand provided sufficient yields may be harvested.

Chapter 6

General Discussion

6.1. Introduction

Until recently, guinea-pig liver and recombinant Factor XIIIa have been the sole sources of commercial TGase with Factor XIIIa now being produced recombinantly in the methylotrophic yeast (Park et al. 2002), though it is only commercially available in small quantities. However, since the enzymes are expensive and time-consuming to purify, their use on an industrial scale has been hampered (Berovici et al. 1987). The discovery of microbial TGase has provided an alternative and has revolutionised many industrial processes. Microbial TGase can be easily obtained from the culture broth and has been shown to have many applications in the food processing and textiles industries (Nielsen 1995). As such a number of patents have been filed in the area which comprehensively cover TGase obtained from microorganisms belonging to the genus *Streptoverticillium* (Andou et al. 1993; Bech et al. 1996).

From the outset of this study, the onus was to identify and characterise novel microbial protein cross-linking enzymes (TGases) that could be used as biocatalysts in a number of new and existing industrial sectors. The work presented here may be divided into three main categories: assay development; screening of micro-organisms using gene homology and random screening approaches; and investigating the TGase enzymes identified in members of the *Actinomycete* family by biochemical and genetic methods, including cloning of relevant genes identified and expression in *E. coli*.

6.2. Assay Development

The biochemical assays chosen for use in this study surrounded two of the three TGase-catalysed reactions, amine incorporation using the hydroxamate (Folk and Chung 1985), biotin-cadaverine incorporation (Slaughter et al. 1992) and [^{14}C] putrescine incorporation assays (Lorand et al. 1972) and cross-linking activity utilizing the synthetic peptide cross-linking assay (Trigwell et al. 2004). The hydroxamate assay and [^{14}C] incorporation into N,N'dimethylcasein have been used previously to identify novel TGase

enzymes with limited success (Motoki et al. 1992; Bech et al. 1996). In this study the hydroxamate assay was found to be the least sensitive perhaps as a function of the reaction mechanism surrounding detection and subsequent release of the iron complex formed during quenching. The assay was also prone to variability upon comparison to alternative TGase assays (the biotin-cadaverine (Slaughter et al. 1992) and peptide cross-linking (Trigwell et al. 2004) assays); and was influenced by the presence of high amounts of protein. This lack of sensitivity is perhaps the reason why so few TGase producing organisms were identified as part of the initial screen. The [^{14}C] putrescine assay has also been used to identify microbial TGase comprising organisms based on the amine incorporating ability of mTGase (Bech et al. 1996; Kobayashi et al. 1998c). As with many biochemical assays there are specificity issues surrounding the choice of substrates since a number of other enzymes may catalyze an identical reaction skewing apparent TGase identification. Suzuki et al. identified a number of proteases that may also catalyse the incorporation [^{14}C] putrescine into N,N' dimethylcasein (Suzuki et al. 1997). Acceleration in the rate of radiolabelled putrescine incorporation was observed using purified chymotrypsin, subtilisin and papain. It is likely that the aforementioned proteases would also catalyze the incorporation of biotin cadaverine into N,N' dimethylcasein also.

Bearing in mind the lack of specificity and sensitivity of individual assays it may be advantageous to use a multitude of biochemical assays for future reference. Alternatively, specific substrates could be manufactured, based on known TGase substrates, with modifications limiting the numbers of available glutamine and lysine residue available for reaction. Such a system would limit the numbers of false positives and false negatives observed whilst searching out novel TGase enzymes. However, novel TGases may not possess identical substrate specificities so a degree of flexibility in the design of such synthetic substrates would need to be taken into account.

Though not available at the beginning of this study, the peptide cross-linking assay solves the specificity issue surrounding the amine incorporation assays. The peptide cross-linking assay was deemed the most sensitive assay of those assessed and when used in the presence of a variety of

inhibitors proved a valuable tool for the confirmation of TGase activity. It may be possible to improve the peptide cross-linking assay by altering the synthetic peptide accordingly thereby enabling a somewhat limited study of the enzymes' substrate specificity. Alternatives to the peptide cross-linking assay exist (Kusch et al. 2006) whereby changing the acyl acceptor peptide would enable in-depth studies into the exact substrate requirements of TGases to be carried out.

It was envisaged that using assays directed towards two TGase catalysed reactions [amine incorporation and peptide cross-linking (Slaughter et al. 1992; Trigwell et al. 2004)], in tandem, would improve on existing and past methodologies that relied on sole TGase assays for the identification of TGase enzymes (hydroxamate or [^{14}C] putrescine assays (Motoki et al. 1992; Bech et al. 1996)). By increasing the number of assays used one is increasing the confidence of detection. Targeting two TGase catalysed reactions may enable more efficient TGase detection, though this depends on the enzymes physiological environment, which ultimately influences the reaction catalysed.

6.3. Inhibition Profiling for TGase Activity

Inhibition profiling was used to confirm results gained from the TGase assays described previously. The effect of both competitive primary amine substrates and site specific inhibitors, on the action of mammalian and microbial TGases was assessed and used as a template on which to gauge the responses of apparent TGase positive microbes. Of the inhibitors used, histamine and putrescine were commonly occurring competitive substrates known to inhibit mammalian and microbial TGases. As all TGases should possess a cysteine residue, incubating with N-ethylmaleimide therefore should abolish TGase activity by covalently modifying the cysteine residue. The synthetic inhibitor R283, whilst found to be a potent inhibitor of mammalian type TGases, had little effect on mTGase and was therefore left out of subsequent screening for the identification of novel microbially derived TGases. It was initially thought that both R281 and R283 targeted and bound the TGase active site cysteine thereby extinguishing TGase activity. R283

failed to inhibit mTGase, suggesting it operates by a different mechanism to R281. Since gpl-TG comprises eighteen cysteine residues it could be suggested that R283 may inhibit gpl-TGase by binding to surface exposed cysteine residues thereby disrupting protein structure or preventing access to the enzyme active site. One hypothesis as to why R283 failed to inhibit mTGase concerns the orientation of the active site. Though the secondary structures of the mTGase and FTG active site regions superimpose well (Kashiwagi et al. 2002a), the positions of the histidine and asparagine/aspartate residues in the mTGase active site are reversed, resulting in altered substrate specificities (Kashiwagi et al. 2002a). This may result in weaker interactions between R283 enabling the molecule to be easily displaced by the substrate. This difference in inhibition profile further highlights the differences at the protein level between the mammalian and microbial TGase isoforms. Although it is predicted that novel microbial TGases should behave in a similar manner to known TGases it is possible that they may exhibit differences in their enzymatic characteristics. Therefore apparent TGases may not conform rigidly to the mTGase inhibition profile; however they should not be discounted assuming they are detectable using the assays described here.

6.4. TGase from Micro-organisms

Although a number of micro-organisms have been postulated to express a TGase enzyme (Makarova et al. 1999), to date microbial TGase has only been described in a limited number of micro-organisms all of which are members of the *Streptomyces* (Motoki et al. 1992; Washizu et al. 1994; Bech et al. 1996) and *Bacillus* (Kobayashi et al. 1998c; de Barros Soares et al. 2003) families. Computational analysis of genes possessing a known biological function has identified a large number of bacterial genes, as potentially containing a TGase catalytic core domain though none have been confirmed experimentally (Ginalski et al. 2004). These hypothetical genes have been identified using mammalian TGase and since mTGase is substantially different to its eukaryotic counterparts, it may be questionable that such hypothetical classification may accurately predict gene product function. Genomic sequencing of micro-organisms appears biased towards

bacteria of medical significance. As such database searching would fail to identify discrete relationships between environmental organisms due to the lack of genetic information surrounding the organisms concerned.

Gene homology searches, however, highlight a small number of organisms belonging to the *Bacillus* and *Streptomyces* genus that contain microbial TGase (Figure 4.1 and 4.2). The process was limited to the sequenced genes present in the database (NCBI), a large proportion of which have been ascribed putative biological functions. The crystal structure of *Bacillus* TGase is yet to be solved. The 3-D structure of *Streptomyces* TGase is unique (Kashiwagi et al. 2002a) and thus provides little help in searching conserved domain databases for potential TGase homologues. To conclude, if a putative TGase exists in the known genome sequences it is likely to encode another novel microbial TGase isoform, adding to the evolutional complexity of this ever increasing enzyme family.

6.4.1. EC Project HIPERMAX

The EC funded HIPERMAX Project (Project No. NMP-3-CT-2003-505790) is concerned with the identification of novel proteins which may be used as applied biocatalysts. Such proteins include microbially derived TGases. As such a large proportion of higher bacteria and fungi were screened by industrial partners for TGase activity, by hydroxamate. As highlighted previously the hydroxamate assays has proven to be least sensitive of the assays used and was perhaps responsible for the detection of low numbers of TGase containing organisms. It may be of interest to reassess the organisms screened using the more sensitive and specific assays. Using such alternative assays would reduce the recovery of high background readings taken using the hydroxamate assay postulated to be brought about by the large about of protein present that was deemed to interfere with the assay.

6.4.2. Screening Micro Organisms

To identify microbes comprising novel TGase activity, screening was initiated using the biotin cadaverine and peptide cross-linking assays (Slaughter et al. 1992; Trigwell et al. 2004). Organisms were selected using the insights gained from database and previous TGase searches (Motoki et al. 1992; Andou et al. 1993; Bech et al. 1996). Screening of organisms was limited to those available from local culture collections, representing a somewhat limited source of environmentally diverse microbes. It would be advantageous to screen a wider variety of organisms for the presence of TGase, such organisms would include extremophiles such as *Sulfolobus acidocaldarius, Methaopyrus* and *Pyrodictium occultum* for example.

Though TGase from *Streptomyces sp.* is extracellular (Kanaji et al. 1993) little is known of its potential localisation in new organisms. By understanding the role played by microbial TGase it may be possible to develop a more targeted approach to identifying microbial homologues. Additionally, much is known about the optimum medium requirements and fermentation conditions for maximum TGase production in *Streptomyces* (Zhu et al. 1996; Junqua et al. 1997; Yan et al. 2005), little is known about the optimal medium conditions for expression of a novel TGase making choosing an appropriate medium difficult. However organisms closely related to either *Bacillus* of *Streptomyces* would possibly require similar medium constituents since closely related members of the species have been shown to express TGase under similar physiological conditions. Given the incredible metabolic diversity of some, altering such conditions to optimise TGase production/identification could represent an incredibly time consuming and potentially fruitless task though may uncover a novel TGase enzyme.

6.4.3. TGase from the Bacilliaceae

To date work has been completed on the exact physiological role and genetic control of *Bacillus* derived TGase (Suzuki et al. 2000; Ragkousi and Setlow 2004; Monroe and Setlow 2006). Physiological levels of TGase were

detectable from a number of *Bacillus* species using both the biotin cadaverine and peptide cross-linking assays increasing confidence in the ability of the assays to detect novel TGase activity from other micro-organisms. In all cases TGase activity was attributable to the later stages of sporulation (those stages involved with cortex formation) in accordance to work undertaken by Kobayashi and co-workers (Kobayashi et al. 1998b).

It is likely that TGase performs an identical role in all *Bacillus sp.*. though a large scale comparison has yet to be completed. Since *Bacillus* are closely related to *Clostridium*, on the basis of *hrcA* comparison (a unique heat shock regulatory gene) (Ahmed et al. 1999), one could postulate that members of the Clostridia also possess a TGase that carries out a similar role to TGase from *Bacillus* sp.. Though the mechanism of sporulation is similar in both *Bacillus* and *Clostridium* the initiation of sporulation is quite different (Brun and Shimkets 2000). Whereas *Bacillus sp.* undergo sporulation in response to nutrient starvation, clostridia use sporulation to escape life-threatening conditions generated by a fermentative metabolism. This metabolism causes a decrease of external and internal pH which dissipates the proton gradient across the cell membrane and initiates sporulation (Durre et al. 2002). Sporulation in Clostridia also appears to be coupled to other regulatory networks such as solventogenesis (Durre and Hollergschwandner 2004). It should be noted that the Clostridia are strict anaerobes and any physiological studies are likely to be difficult.

The difficulties encountered during this study, relating to efficient preparation of *Bacillus sp.* TGase, are perhaps the reasons why *Bacillus* TGase has yet to be applied on an industrial scale. Provided efficient production of the enzyme could be mastered, *Bacillus* TGase may present new exploitable alternatives to *Streptomyces sp.* TGase by its altered substrate specificities and reaction rates. If TGases could be purified from the more extreme members of the *Bacillus* genus, and provided suitable production could be attained, it would enable the best TGase, with respect to cross-linking ability, to be used for the specific reaction requirements.

6.4.4. TGase from Pseudomonas putida

Of the micro-organisms selected for screening, only two strains of *Pseudomonas putida* appeared to express a potential TGase enzyme, towards the stationary phase of growth, suggesting a role in the maintenance of cellular integrity during the transition between the exponential and stationary phases of growth (Section 4.2.3.5). This transition is environmentally relevant, since cells pass from a phase of unrestricted growth to a situation of nutrient limitation and diverse stress. The attribution of such a role has been postulated on the basis of the potential role of mTGase in cellular and morphological differentiation (Kashiwagi et al. 2002a). TGase activity was observed in only two out of three *Pseudomonas putida* strains assessed perhaps reinforcing the great diversity observed throughout the *Pseudomonas* genus.

Apparent TGase activity was confirmed by using the peptide cross-linking assay in the presence of TGase inhibitors as well as identifying the presence of the ε(-γ glutamyl)lysine bond in the lysates generated, serving as an indication of TGase activity. The enzyme responsible for the apparent TGase activity observed was D-tyrosyl-tRNA tyr deacylase, based on degenerate PCR results, biochemical observations and literature searches. From this a number of feasible conclusions may be made: the organism (*Ps. putida*) encodes a putative TGase enzyme that possesses different biochemical properties to other known TGase enzymes; the assays used for detection of novel TGases enabled the detection of alternate enzymes capable of catalysing a similar reaction; *Ps. putida* encodes other proteins that are capable of performing TGase type reactions under distinct conditions, of these the latter would appear most likely based on the evidence presented. The methodology, incorporating specific TGase assays and inhibitors, was designed to limit the number of false positives observed. It should be noted, however, that the inhibitors may have many other substrates *in vivo* and would not only limit the extent of TGase catalysed incorporation. The suicide inhibitor, which was initially designed against mammalian–type TGases, was shown to bind many cellular proteins (Section 4.2.3.9) which would have affected the assay accordingly.

From the data generated it would appear that some strains of *Ps. putida* contain an enzyme capable of forming the ε(-γ glutamyl)lysine bond which may represent a novel lineage in the TGase evolutionary tree, though more work would be required to substantiate the claim. It is also possible that TGase activity has been acquired, by horizontal transfer, a protein carrying out a similar function during the evolutionary process.

6.5. TGase from Actinomycetes

As mentioned previously, a number of *Streptomyces sp.* including *S. baldaccii, S. platensis* and *S. paucisporogenes,* and *Saccharomonospora viridis* were identified as possessing TGase activity, which was confirmed on the basis of inhibition profiling. Given the high similarity between purified TGases it is likely that the zymogen is processed by a co-secreted endogenous protease as part of an identical activation pathway to that described by Zotzel and co-workers for the processing of *S. mobaraensis* TGase (Zotzel 2003a; b). It maybe that the endogenous proteases would also share a high degree of similarity between the *Streptomyces* species identified. Interestingly, TGase was not identified in all *Streptomyces* as one might expect from this closely related family (Stackebrandt et al. 1981) perhaps suggesting separate lineages of organisms within the genus bringing into question the classification of the species, a commonly occurring problem. This would therefore suggest that either the evolution or dissemination of TGase amongst the *Streptomyces* is in its early stages or conversely that, since the selective pressure in no longer present, the TGase protein is evolving to fulfil a different enzymatic niche or is sequentially being lost. Another possibility is that TGase has not originated from *Streptomyces* and that it passed to the *Streptomyces* genus from a closely related ancestor, though further screening would be required to substantiate this. It is apparent that TGase is spreading throughout the Actinomycetes perhaps through a process of horizontal transfer.

6.5.1. Characteristics of Streptomyces Derived TGase

The biochemical characteristics of the enzymes purified appeared remarkably similar to both each other and the commercially available TGase

derived from *S. mobaraensis* (Ando et al. 1989a). In the case of the *Streptomyces* this was not totally unexpected since they are closely related (Stackebrandt et al. 1981). *Streptomyces* derived TGases shared a similar molecular weights and optimal temperature (37 °C) and pH optima (pH 7) to *S. mobaraensis* TGase, with the exception of *S. baldaccii* TGase which retained a greater proportion of activity at lower pH perhaps indicative of the different environmental habitats that the organisms inhabits. It may be reasonable to assume that more diverse *Streptomyces* sp., that inhabit more extreme environments, may possess TGase enzymes with differing catalytic properties, organisms may include many thermophilic and psycrophilic strains for example *S. thermogriseus* (Xu et al. 1998) isolated from various hot springs and soils and *S. beijiangensis* (Li et al. 2002) as well as other extreme Actinomcyetes. Such a hypothesis may also hold true for the Bacillus family.

Further comparisons between the purified enzymes highlighted that the TGases purified from *S. platensis* and *S. paucisporogenes* were capable of incorporating a higher amount of protein cross-links into casein than the commercially available TGase, making them more cost effective in certain instances for example in the modification of milk proteins (Vasbinder et al. 2003).

6.5.2. Physiological Role of Streptomyces TGase

To date, the physiological role of mTGase has yet to be identified though It is thought to be involved with mycelial growth and morphological differentiation (Kashiwagi et al. 2002a). In this study maximum TGase activity was recovered at the transition between the late exponential and stationary phases of liquid growth. However, higher yields were obtained from solid medium growth (personal communication G. Szakacs, Technical University of Budapest) suggesting that TGase is actively involved in hyphal formation. Given that *Streptomyces* sp. exhibit identical gene expression cascades regardless of the growth medium and given that shaking liquid culture inhibits sporulation (Fernandez et al. 2002) it maybe that mTGase is indeed involved with hyphal and spore formation. This notion is supported by work carried

out with the synthetic TGase inhibitor R281. Incubation with the inhibitor was found to suppress normal growth. The method of applying TGase specific inhibitors in a physiological setting to assess the involvement of a TGase enzyme in cell wall formation may help uncover the exact physiological role of TGase. By understanding the exact occurrence of mTGase it will no doubt aid the identification of TGase from alternative organisms that grow in a similar way to the *Streptomyces*.

6.5.3. Industrial Applications of Streptomyces TGase

Microbial TGase was detected in *Streptomyces baldaccii*. The gene encoding the active protein was cloned and expressed as a functional GST fusion protein enabling mass-production of the enzyme at relatively low cost, compared to conventional purification techniques. Currently, new techniques have been applied to the mass-production of mTGase that may be applicable to the mass-production of other microbial TGase isoforms. Liu and co-workers (Liu et al. 2006b) have described a technique whereby an expression plasmid encoding the endogenous TGase is transformed into the original strain (*Streptomyces fradiae*) enabling a two-fold increase of TGase. Whilst yields may be elevated, the purification costs would remain high.

TGase from *S. baldaccii* provides an alternative to *S. mobaraensis* TGase (Ajinomoto Co. Ltd., Japan) for the modification of substrate proteins. At the genetic level it shares 80% identity to its *S. mobaraensis* counterpart, with the differences likely to give rise to different functional characteristics, with regards to substrate specificity, thereby opening new avenues for the investigation of protein modification by TGase. It might be possible to further enhance enzyme activity and substrate binding by altering the N-terminus amino acids of TGase (not of primary importance for the global fold in mTGase)(Kashiwagi et al. 2002a) if required. This process has been applied previously to *S. mobaraensis* TGase with great success (Shimba et al. 2002). Additionally, a number of techniques exist for expanding an enzymes catalytic range including gene shuffling and directed evolution, techniques that have been used successfully with other enzymes. Such techniques have been used to improve the thermostability of maltogenic amylase from *Bacillus*

thermoalkalophilus (Tang et al. 2006) as well as improving the thermostability and catalytic activity of alpha-aspartyl dipeptidase from *Salmonella typhimurium* (Kong et al. 2001). Improvements to alpha-aspartyl dipeptidase were made by rounds of error-pone PCR followed by DNA shuffling that generated mutant strains harbouring amino acid substitutions near to both the catalytic centre and substrate binding pocket. The mutations were responsible for the catalytic improvements observed.

It is apparent that a number of TGases exist within the *Streptomyces* genus that may be of industrial use. Although patent restrictions currently apply to *Streptomyces* derived TGases, the direct commercial use of *S. baldaccii* TGase is not possible, though this does not prevent the development of new TGase-catalysed applications. If such applications are developed, the processes themselves would be patentable and of considerable economic value to the inventor.

References

Aeschilmann, D., and Paulsson, M. 1994. Transglutaminase: Protein cross-linking enzymes in tissues and body fluids. *Thrombosis and Haemostasis* **4:** 402-415.

Aeschilmann, D., Koeller, M., Allen-Hoffman, B., and Mosher, D. 1998. Isolation of a cDNA encoding a novel member of the transglutaminase gene family from human keratinocytes. Detection and identification of transglutaminase gene products based on reverse-transcription-polymerase chain reaction with degenerate primers. *Journal of Biological Chemistry* **273:** 3452-3460.

Ahmed, S., Selvapandiyan, A and Bhatnagar, R. 1999. A protein-based phylogenetic tree for Gram-positive bacteria derived from *hrcA*, a unique heat-shock regulatory gene. *International Journal of Systematic Bacteriology* **49:** 1387-1394.

Ahvazi, B., Boeshans, K. and Steinert, P. 2004. Crystal structure of transglutaminase 3 in complex with GMP: structural basis for nucleotide specificity. *Journal of Biological Chemistry* **279:** 26716-26725.

Altschul, S., Madden, T., Schäffer, A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. . 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25:** 3389-3402.

Ando, H., Adachi, M., Umeda, K., Matsuurs, A., Nonaka, M., Uchio, R., Tanaka, H. and Motoki, M. 1989a. Purification and characteristics of a novel transglutaminase derived from micro-organisms. *Agricultural and Biological Chemistry* **53:** 2613-2617.

Ando, Y., Imamura, S., Owada, M. K., Kakunaga, T. and Kannagi, R. 1989b. Cross-linking of lipocortin I and enhancement of its Ca2+ sensitivity by tissue transglutaminase. *Biochemistry and Biophysics Research Communications* **163:** 944-951.

Andou, H., Matsuura, A. and Hirose, S. 1993. Process for producing a transglutaminase derived from *Streptomyces. US Patent Number 5,252,469*.

Ash, C., Farrow, J., Wallbanks, S. and Collins M. . 1991. Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative analysis of small-subunit-ribosomal RNA sequences. *Letters in Applied Mirobiology* **13:** 202-206.

Baker, E., and Dodson, E. 1980. Crystallographic refinement of the structure of actinidin at 1.7 angstrom resolution by fast fourier least squares methods. *Acta Crystalography.* **A36:** 559-572.

Bale, M., and Mosher, D. 1986. Thrombospondin is a substrate for blood coagulation factor XIIIA. *Biochemistry* **25:** 5667-5673.

Barr, D. 1992. Evolution and kingdoms of organisms from the perspective of a mycologist. *Mycologia* **84:** 1-11.

Barrett, R., and Bell, G. 2006. The dynamics of diversification in evolving *Pseudomonas* populations. *International Journal of Organic Evolution* **60:** 484-490.

Baumgartner, W., Golenhofen, N., Weth, A., Hiiragi, T., Saint, R. Griffin, M. and Drenckhahn, D. 2004. Role of transglutaminase 1 in the stabilisation of intracellular junctions of the vascular endothelium. *Histochemistry and Cell Biology* **122:** 17-25.

Bech, L., Norrevang, I., Halkier, T., Rasmussen, G., Schafer, T. and Andersen, J. 1996. Microbial Transglutaminase, their production and use. *International Patent; PCT/DK95/003347.*

Bergkvist, R. 1963. The proteolytic enzymes of *Aspergillus oryzae. Acta Chemistry (Scandinavia)* **17:** 1521-1551.

Berovici, D., Gaertner, F. and Puigserver, A. 1987. Transglutaminase catalysed incorporation of lysing oligomers into casein. *Journal of Agricultural Food Chemistry* **35:** 3301-3304.

Bershad, S. 2001. Developments in topical retinoid therapy for acne. *Seminars in Cutaneous Medicine and Surgery* **20:** 154-161.

Beyazova, M., and Lechevalier, M. 1993. Taxonomic utility of restriction endonuclease fingerprinting of large DNA fragments from *Streptomyces* strains. *International Journal of Systematic Bacteriology* **43:** 674-682.

Bloomfield, S., and Arthur. M. 1994. Mechanisms of inactivation and resistance of spores to chemical biocides. *Journal of Applied Bacteriology* **76:** 91-104.

Boeshans, K., Mueser, T., and Ahvazi, B. 2007. A three-dimensional model of the human transglutaminase 1: insights into the understanding of lamellar ichthyosis. *Journal of Molecular Modelling* **13:** 233-246.

Broker, M., Baumi, O., Gottig, A., Ochs, J., Bodenbenner, M. and Agmann, E. 1991. Expression of human blood coagulation protein factor XIIIa in *Saccharomyces cerevisiae*: dependence of the expression levels from host-vector systmes and medium conditions. *Applied Microbiology and Biotechnology* **34:** 756-764.

Brun, Y., and Shimkets, L. 2000. Prokaryotic Development. *Washington, DC, USA: American Society for Microbiology.*

Bruns, T., White, T., and Taylor, J. 1991. Fungal Molecular Systematics. *Annual Review of Ecology and Systematics* **22:** 525-564.

Bures, D., and Goldsmith, L. 1978. Localization of transglutaminase in adult chicken epidermis. *Archives of Dermatology Research* **262:** 329-332.

Buxman, M., and Wuepper, K. 1975. Keratin cross-linking ad epidermal transglutaminase. A review with observations on the histochemical and immunochemical localisation of the enzyme. *Journal of Investigative Dermatology.* **65:** 107-112.

Candi, E., Oddi, S., Terrinoni, A., Paradisi, A., Ranalli, M., Finazzi-Agro, A. and Melino, G. 2001. Transglutaminase 5 crosslinks loricrin, involucrin and small proline-rich proteins *in vitro. Journal of Biological Chemistry* **276:** 35014-35023.

Candi, E., Paradis, A., Terrinoni, A., Pietron, V., Oddi, S., Cadot, B., Jogini, V., Meiyappan, M., Clardy, J., Finazzi-Agro, A., and Melino, G. 2004. Transglutaminase 5 is regulated by guanine-adenine nucleotides. *Biochemical Journal* **381:** 313-319.

Carballo, J., Ayo, J., and Jimenez Colmenero, F. 2006. Microbial transglutaminase and caseinate as cold set binders: Influence of meat species and chilling storage. *LWT* **39**.

Chandrashejar, R., and Mehta, K. 2000. TG-catalysed reactions in the growth, maturation and development of parasitic nematodes. *Parasitology Today* **16:** 11-17.

Chater, K. 1989. Sporulation in Streptomyces. *Regulation of prokaryotic development structural and functional analysis of bacterial sporulation and germination. American society for microbiology.*: 277-299.

Cohen, C., Dotimas, E. and Korsgren, C. 1993. Human erythrocyte membrane protein band 4.2. *Seminars in Hematology* **30:** 119-137.

Cohen, I., Lim, C., Kahn, D., Glaser, T. Gerrard, J. and White, J. 1985. Disulphide-linked and transglutaminase-catalysed protein assemblies in platelets. *Blood* **66:** 143-151.

Cortez, J., Bonner, P. and Griffin, M. 2004. Application of transglutaminases in the modification of wool textiles. *Enzyme and Microbial Technology* **34:** 64-72.

Cortez, J., Bonner, P and Griffin, M. 2005. Transglutaminase treatment of wool fabrics leads to resistance to detergent damage. *Journal of Biotechnology* **116:** 379-386.

Cui, L., Zhang, D., Huang, L., Liu, H., Du, G., and Chen, J. 2006. Stabilization of a new microbial transglutaminase from *Streptomyces hygroscopicus* WSH03-13 by spray drying. *Process Biochemistry* **41**: 1427-1431.

D'Argenio, G., Grossman, A., Cosenza, V., Valle, N., Mazzacca, G. and Bishop, P. 2000. Recombinant factor XIII improves established experimental colitis in rats. *Digestive Diseases and Sciences* **45**: 987-998.

de Barros Soares L, A.F., Zachia Ayub M. 2003. Purification and properties of a transglutaminase produced by a bacillus circulans strain isolated from the amozonian environment. *Biotechnology and Applied Biochemistry* **37**: 295-299.

de Jong, G., Wijnaards, G., Boumans, H., Koppelman, S. and Hessing, M. 2001. Purification and substrate specificity of transglutaminase from blood and *Streptoverticillium mobaraense. Journal of Agricultural Food Chemistry* **49**: 3389-3393.

DeLano, W.L. 2002. The PyMOL molecular graphics system

Della Mea, M., Caparros-Ruiz, D., Claparols, I., Serafini-Fracassini, D. and Rigau, J. 2004. AtPng1p. The first plant transglutaminase. *Biochemical Processes and Macromolecular Structures* **135**: 2046-2054.

Dempfle, C., Harenberg, J., Hochreuter, K. and Heene, D. . 1992. Microtiter assay for measurement of factor XIII activity in plasma. *Journal of Laboratory and Clinical Medicine* **119**: 522-528.

DiGiovanna, J., and Robinson-Bostom, L. 2003. Ichthyosis: etiology, diagnosis, and management. *American Journal of Clinical Dermatology* **4**: 81-95.

Doolittle, R., Watt, K., Cottrell, B., Strong, D. and Riley, M. 1979. The amino acid sequence of the alpha chain of human fibrinogen. *Nature* **280**: 464-468.

Drenth, J., Kalk, K. and Swen, H. 1976. Binding of chloromethyl ketone substrate analogues to crystalline papain. *Biochemistry* **15**: 3731-3738.

Dubbink, H., Verkaik, N., Faber, P., Trapman, J., Schroder, F. and Romijn, J. 1996. Tissue specific and androgen regulated expression of human prostate specific transglutaminase. *Biochemical Journal* **3115**: 901-908.

Duran, R., Junqua, M., Schmitter, J., Gancet, C., and Goulas, P. 1998. Purification, characterisation, and gene cloning of transglutaminase from *Streptoverticillium cinnamoneum* CBS 683.68. *Biochimie* **80**: 313-319.

Durre, P., Bohringer, M., Nakotte, S., Schaffer, S., Thormann, K. and Zickner, B. 2002. Transcriptional regulation of solventogenesis in *Clostridium acetobutylicum. Journal of Molecular Microbiology and Biotechnology* **4**: 295-300.

Durre, P., and Hollergschwandner, C. 2004. Initiation of endospore formation in *Clostridium acetobutylicum. Anaerobe* **10:** 69-74.

Eligula, L., Chuang, L., Phillips, M., Motoki, M., Seguro, K., and Muhlrad, A. 1998. Transglutaminase-induced cross-linking between subdomain 2 of G-actin and the 636-642 lysine-rich loop of myosin subfragment 1. *Biophysical Journal* **74:** 953-963.

Embley, T., and Stackebrandt, E. 1994. The molecular Phylogeny and Systematics of the Actinomycetes. *Annual Review of Microbiology* **48:** 257-289.

Ensign, J. 1978. Formation, properties and germination of actinomycete spores. *Annual Reviews of Microbiology* **32:** 185-219.

Errington, J. 1993. *Bacillus subtilis* Sporulation: Regulation of Gene Expression and Control of Morphogenesis. *Microbiological Reviews* **57:** 1-33.

Fawcett, P., Eichenberger, P., Losick, R., and Youngman, P. 2000. The transcriptional profile of early to middle sporulation in *Bacillus subtilis. Proceedings of the National Academy of Science USA* **97:** 8063-8068.

Fear, J., Jackson, P., Gray, C., Miloszewski, K. and Losowsky. 1984. Localisation of factor Xiii in human tissues using an immunoperoxidase technique. *Journal of biological chemistry* **37:** 560-563.

Feng, D., Cho, G. and Doolittle, R. 1997. Determining divergence times with a protein clock: update and reevaluation. *Proceedings of the National Academy of Science USA* **94:** 13028-13033.

Fernandez, M., and Sanchez, J 2002. Nuclease activities and cell death processes associated with the develpoment of surface cultures of *Streptomyces antibioticus* ETH 7451. *Microbiology* **148:** 405-412.

Ferri-Fioni, M., Schmitt, E., Soutourina, J., Plateau, P., Mechulam, Y., and Blanquet, S. 2001. Structure of Crystalline D-Tyr-tRNATyr Deacylase. A REPRESENTATIVE OF A NEW CLASS OF tRNA-DEPENDENT HYDROLASES *Journal of Biological Chemistry* **276:** 47285-47290.

Fiorentini, C., Donelli, G., Matarrese, P., Fabbri, A., Paradisi,S. and Boquet, P. 1995. *Escherichia coli* cytotoxic necrotizing factor 1: evidence for induction of actin assembly by constitutive activation of the p21 Rho GTPase. *Infection and Immunity* **63:** 3936-3944.

Flanagan, J., and Fitzgerald, R. 2003. Characterisation and quantification of the reaction(s) catalysed by transglutaminase using o-phtaldialdehyde reagent. *Die Nahrung* **47:** 207-212.

Folk, J., and Cole, P. 1966. Identification of a functional cysteine essential for the activity of guinea pig liver transglutaminase. *Journal of Biological Chemistry* **241:** 3238-3240.

Folk, J. 1969. Mechanism of action of guinea pig liver transglutaminase. VI. Order of substrate addition. *Journal of Biological Chemistry* **244:** 3707-3713.

Folk, J., and Finlayson. 1977. The epsilon (gamma-glutamyl)lysine crosslink and the catalytic role of transglutaminase. *Advances in Protein Chemistry* **31:** 11-33.

Folk, J., and Chung, S. 1985. Transglutaminases. *Methods in Enzymology* **113:** 358-357.

Folk, J., Ekmekcioglu, S. and Mehta, K. 2006. Implications of tissue transglutaminase in malignant melanoma. *Molecular Cancer Therapeutics* **5:** 1493-1503.

Fornelli, S. 1993. Enzymatic treatment of protein fibres- state of the art biotechnology. *International Dyer* **178:** 29-33.

Fox, K., and Eder, B. 1969. Comparison of survivor curves of *Bacillus subtilis* spores subjected to wet and dry heat. *Journal of Food Science* **34:** 518-521.

Francis, R., McDonagh, J. and Mann, K. 1986. Factor V is a substrate for the transamidase factor XIIIa. *Journal of Biological Chemistry* **261:** 9787-9792.

Frebort, I., and Adachi, O. 1995. Copper/quinone-containing amine oxidases, an exciting class of ubiquitous enzymes. *Journal of Fermentation and Bioengineering* **80:** 625-632.

Gentile, V., Saydak, M., Chiocca, E., Akande, O., Birckbichler, P., Lee, K., Stein, J., and Davies, P. 1991. Isolation and characterization of cDNA clones to mouse macrophage and human endothelial cell tissue transglutaminases. *Journal of Biological Chemistry* **266:** 478-483.

Gillet, S., Pelletier, J. and Keillor, J. 2005. A direct fluorometric assay for tissue transglutaminase. *Analytical Biochemistry* **347:** 221-226.

Ginalski, K., Kinch, L., Rychlewski, L., and Grishin, N. 2004. BTLCP proteins: a novel family of bacterial transglutaminase-like cysteine proteinases. *Trends Biochemical Science* **29:** 392-395.

Gladek, A., Mordaski, M., Goodfellow, M., and Williams, S. 1985. Ribosomal ribonucleic acid similarities in the classification of *Streptomyces*. *FEMS Microbiology Letters* **26:** 175-180.

Gootenberg, J. 1998. Factor concentrates fro the treatment of factor XIII deficiency *Current Opinions in Hematology* **5**: 372-375.

Greenberg, C., Birckbichler, P., and Rice, R. 1991. Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues. *Faseb Journal* **5**: 3071-3077.

Grenard, P., Bates, M. and Aeschlimann, D. 2001. Evolution of transglutaminase genes: identification of a transglutaminase gene cluster on human chromosome 15q15. Structure of the gene encoding transglutaminase X and a novel gene family member, transglutaminase Z. *Journal of Biological Chemistry* **276**: 33066-33078.

Griffin, M., and Wilson, J. 1984. Detection of epsilon(gamma-glutamyl) lysine. *Molecular and Cellular Biochemistry* **58**: 37-49.

Griffin, M., Casadio, R., and Bergamini, C. 2002. Transglutaminases: nature's biological glues. *Biochemical Journal* **368**: 377-396.

Griffin, M., Coutts, I. and Saint, R. 2006. Dipeptide transglutaminase inhibitors and methods of using the same. *WO2004113363 (A3)*.

Haroon, Z., Hettasch, J., Lai, T., Dewhirst, M., and Greenberg, C. 1999. Tissue transglutaminase is expressed, active, and directly involved in rat dermal wound healing and angiogenesis. *FASEB Journal.* **13**: 1787-1795.

Hatano, K., Nishii, T and Kasai, H. 2003. Taxonomic re-evaluation of whorl-forming *Streptomyces* (formerly *Streptoverticillium*) species by using pheotypes, DNA-DNA hybridisation and sequences of *gyrB*, and proposal of *Streptomyces luteireticuli* (ex Kato and Arai 1957) corrig., sp. nov., nom. rev. *International Journal of Systematic bacteriology* **53**: 1519-1529.

Haywood, G., and Large, P. 1985. The occurrence, subcellular localisation and partial purification of diamine acetyltransferase in yeast *Candida boidinii* grown on spermidine or putrescine as sole nitrogen source. *European Journal of Biochemistry* **148**: 277-283.

Herzberg, O., and Moult, J. 1987. Bacterial resistance to beta-lactam antibiotics: crystal structure of beta-lactamase from *Staphylococcus aureus* PC1 at 2.5 A resolution. *Science* **236**: 694-701.

Higgins, D., Thompson, J., and Gibson T. . 1994. Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Hohl, D. 1993. Expression patterns of loricrin in dermatological disorders. *American Journal of Dermatopathology* **15**: 20-27.

Horiguchi, Y. 2001. *Escherichia coli* cytotoxic necrotizing factors and *Bordetella* dermonecrotic toxin: the dermonecrosis-inducing toxins activating Rho small GTPases. *Toxicology* **39**: 1619-1627.

Ichinose, A., Hendrickson, L., Fujikawa, K. and Davie, E. . 1986. Amino Acid sequence of the a subunit of human factor XIII. *Biochemistry* **25**: 6900-6906.

Iismaa, S., Chung, L., Wu, M., Teller, D., Yee, V. and Graham, R. 1997. The core domain of tissue transglutaminase Gh hydrolyses GTP and ATP. *Biochemistry* **36**: 11655-11664.

Ikura, K., Sasaki, R., and Motoki, M. 1992. Use of transglutaminase in quality-improvement and processing of food proteins. *Comments on Agricultural and Food Chemistry* **2**: 389-407.

Ikura, K., Kita, K., Fujita, I., Hashimoto, H., and Kawabata, N. 1998. Identification of amine acceptor protein substrates of transglutaminase in liver extracts: use of 5-(biotinamido) pentylamine as a probe. *Archives of Biochemistry and Biophysics* **356**: 280-286.

Ikura, K., Minami, K., Otomo, C., Hashimoto, H., Natsuka, S., Oda, K., and Nakanishi, K. 2000. High molecular weight transglutaminase inhibitor produced by a microorganism (*Streptomyces lavendulae* Y-200). *Bioscience, Biotechnology and Biochemistry* **64**: 116-124.

Jetten, A., Kim, J., Sacks, P., Rearick, J., Lotan, D., Hong, W. and Lotan, R. 1990. Inhibition of growth and squamous-cell differentiation markers in cultured human head and neck squamous carcinoma cells by beta-all-trans retinoic acid. *International Journal of Cancer* **45**: 195-202.

Junqua, M., Duran, R., Gancet, C. and Goulas, R. 1997. Optimisation of microbial transglutaminase production using experimental designs. *Applied Microbiology and Biotechnology* **6**: 730-734.

Jurgensen, K., Aeschlimann, D., Cavin, V., Genge, M., and Hunziker, E. B. 1997. A new biological glue for cartilage-cartilage interfaces: tissue transglutaminase. *Journal of Bone Joint Surgery* **79**: 185-193.

Kaetsu, H., Hashiguchi, T., Foster, D. and Ichinose, A. 1996. Expression and release of the a and b subunits for human coagulation factor XIII in baby hamster kidney (BHK) cells. *Journal of Biochemistry* **119**: 961-969.

Kalakoutskii, L. and Agre, N. 1976. Comparative aspects of development and differentiation in actinomycetes. *Bacteriological Reviews* **40**: 469-524.

Kampfer, P., Kroppenstadt, R. and Dott, W. 1991. A numerical classification of the genera *Streptomyces* and *Streptoverticillium* using miniaturized physiological tests. *Journal of General Microbiology* **137:** 1831-1891.

Kanaji, T., Ozaki, H., Takao, T., Kawajiri, H., Ide, H., Motoki, M., and Shimonishi, Y. 1993. Primary structure of microbial transglutaminase from *Streptoverticillium sp.* strain s-8112. *Journal of Biological Chemistry* **268:** 11565-11572.

Karandikar, A., Sharples, G., and Hobbs, G. 1997. Differentiation of streptomyces coelicolor A3(2) under nitrate-limited conditions. *Microbiology* **143:** 3581-3590.

Kashiwagi, T., Yokoyama, K., Ishikawa, K., Ono, K., Ejima, D., Matsui, H and Suzuki, E. 2002a. Crystal structure of microbial transglutaminase from *Streptoverticillium mobaraense*. *JOurnal of biological chemistry* **277:** 44252-44260.

Kashiwagi, T., Shimba, N., Ishikawa, K., Suzuki, E., Yokoyama, K. and Hirayama, K. 2002b. Method of modifying microorganism-origin transglutaminase. *WO0214518*.

Kawai, M., Takehana, S. and Takagi, H. 1997. High-level expression of the chemically synthesized gene for microbial transglutaminase from *Streptoverticillium* in *Escherichia coli*. *Bioscience, Biotechnology and Biochemistry* **61:** 830-835.

Kieser, T., Bibb, M., Buttner, M., Chater, K., and Hopwood, D. . 2000. Practical Streptomyces Genetics. *Norwich, UK. John Innes Foundation*.

Kilic, B. 2003. Effect of microbial transglutaminase and sodium caseinate on quality of chicken doner kebab. *Meat Science* **63:** 417-421.

Kim, S., Chung, S. and Steinert, P. 1995. Highly active soluble processed forms of the transglutaminase 1 enzyme in epidermal keratinocytes. *Journal of Biological Chemistry* **270:** 18026-18035.

Kim, S., Jeitner, T., and Steinert, P. 2002. Transglutaminases in disease. *International Neurochemistry* **40:** 85-103.

Klein, J., Guzman, E. and Keuhn, G. 1992. Purification and partial characterisation of transglutaminase from *Physarum polycehalum*. *Journal of Bacteriology* **174:** 2599-2605.

Kobayashi, K., Yamanaka, S., Tanita, Y., Tsuyoshi, N., Fudo, R., Shinozaki, J., Yokozeki, K and Suzuki, S. 1998c. Microbial process for producing transglutaminase. *AU698813B*.

Kobayashi, K., Hashiguchi, K., Yokozeki, K. and Yamanaka, S. 1998a. Molecular Cloning of the transglutaminase gene from *Bacillus subtilis* and its expression in *Escherichia coli*. *Bioscience, Biotechnology and Biochemistry* **62**: 1109-1114.

Kobayashi, K., Suzuki, S., Izawa, Y., Miwa, K., and Yamanaka, S. 1998b. Transglutaminase in sporulating cells of *Bacillus subtilis*. *Journal of General and Applied Microbiology* **44**: 85-91.

Kong, X., Liu, Y., Gou, X., Zhu, S., Zhang, H., Wang, X., and Zhang, J. 2001. Directed evolution of alpha-aspartyl dipeptidase from *Salmonella typhimurium*. *Biochemical and Biophysical Research Communications* **289**: 137-142.

Korsgren, C., Lawler, J., Lambert, S., Speicher, D. and Cohen, C. 1990. Complete amino acid sequence and honologies of human erythrocyte membrane protein band 4.2. *Proceedings of the National Academy of Science USA* **87**: 613-617.

Kraut, J. 1977. Serine proteases: structure and mechanism of catalysis. *Annual Review Biochemistry* **46**: 331-358.

Kumazawa, Y., Numazawa, T., Seguro, K., and Motoki, M. 1995. Suppression of surimi gel setting by transglutaminase inhibitors. *Journal of Food Science* **60**: 715-717.

Kuraishi, C., Sakamoto, J., Yamazaki, K., Susa, Y., Kuhara, C., and Soeda, T. 1997. Production of restructured meat using microbial transglutaminase without salt or cooking. *Journal of Food Science* **62**: 488-490.

Kurth, and Rogers. 1984. Transglutaminase catalysed crosslinking of myosin to soya protein, casein and gluten. *Jourmal of Food Science.* **49**: 573-576.

Kusch, M., Grundman, C., Keitel, S., Seitz, R. and Konig, H. 2006. A novel assay for Factor XIII based on cross-linking of synthetic peptides: analysis of different substrates. *Blood Coagulation and Fibrinolysis* **17**: 575-580.

Labeda, D. 1992. DNA-DNA hybridisation in the systematics of *Streptomyces*. *Gene* **115**: 249-253.

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

Lauber, S., Krause, I., Klostermeyer, H. and Henle, T. 2003. Microbial transglutaminase cross-links beta-casein and beta-lactoglobulin to heterologous oligomers under high pressure *European Journal of Food Research Technology* **216**: 15-17.

LeBlanc, A., Day, N., Menard, A. and Keillor, J. 1999. Guinea pig liver transglutaminase: A modified purification procedure affording enzyme with superior activity and greater yield. *Protein Expression and Purification* **17**: 89-95.

Lechevalier, M., and Lechevalier, H. 1970. Chemical composition as a criterion in the classification of aerobic actinomycetes. *International Journal of Systematic Bacteriology* **20**: 435-443.

Lee, K., Arnold, S., Birckbitcher, P., Patterson, M., Frais, B., Takeu, Y. and Carter, H. 1993. Site mutagenesis of human tissue transglutaminase cys277 is essential for transglutaminase activity but not for GTPase activity. *Biochim. Biophys. Acta.* **1202**: 1-6.

Lentini, A., Kleinman, H., Mattiolo, P., Autuori-Pezzoli, V., Nicolini, L., Pietrini, A., Abbruzzese, A., Cardinali, M. and Beninati, S. 1998. Inhibition of melanoma pulmonary metastasis by methylxanthines due to decreased invasion and proliferation. *Melanoma Research* **8**: 131-137.

Lerm, M., Schmidt, G., Goehring, U., Schirmer, J., and Aktories, K. 1999. Identification of the region of rho involved in substrate recognition by *Escherichia coli* cytotoxic necrotizing factor 1. *Journal of Biological Chemistry* **274**: 28999-29004.

Li, W., Zhang, L., Xu, P., Cui, X., Lu, Z., Xu, L., and Jiang, C. 2002. *Streptomyces beijeiangensis* sp. nov., a psycrotolerant atinomycete isolated from soil in China. *International Journal of Systematic Bacteriology* **52**: 1695-1699.

Lim, K., Tempczyk, A., Bonander, N., Toedt, J., Howard, A., Eisenstein, E. and Herzberg. O. 2003. A catalytic mechanism for d-Tyr-tRNA tyr deacylase based on the crystal structure of *Hemophilus influenzae* HI0670. *Journal of Biological Chemistry* **278**: 13496-13502.

Liu, S., Cerione, R., and Clardy, J. 2002. Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity. *Proceedings of the National Academy of Science USA* **99**: 2743-2747.

Liu, X., Yang, X., Xie, F. and Qian, S. 2006b. Cloning of transglutaminase gene from *Streptomyces fradiae* and its enhance expression in the original strain. *Biotechnology Letters* **28**: 1319-1325.

Liu, X., Yang, X., Xie, F., Song, L., Zhang, G. and Qian, S. 2006a. On-column refolding and purification of transglutaminase from *Streptomyces fradiae* expressed as inclusion bodies in *Escherichia coli. Protein Expression and Purification*: In Press.

Lorand, L., Campbell-Wilkes, L. and Cooperstein, L. 1972. A filter paper assay for transamidating enzymes using radioactive amine substrates. *Analytical Biochemistry* **50:** 623-631.

Lorand, L., Parameswaran, K., Stenberg, P., Tong, Y., Velasco, P., Jonsson, N., Mikiver, L., and Moses, P. 1979. Specificity of guinea pig liver transglutaminase for amine substrates. *Biochemistry* **18:** 1756-1765.

Lorand, L., and Conrad, S. 1984. Transglutaminases. *Molecular and Cellular Biochemistry* **58:** 9-35.

Lorand, L., and Graham, R. 2003. Transglutaminases: crosslinking enzymes with pleitropic functions. *Nature Reviews: Molecular Cell Biology* **4:** 140-156.

Lowry, O., Rosebrough, N. , Farr, A., and Randall, R. 1952. Protein measurement with the Folin phenol reagent. *Journal of Biological Chemistry* **193:** 265-275.

Lynch, G., Slayter, H., Miller, B. and McDonagh, J. 1987. Characterisation of thrombospondin as a substrate for factor XIII transglutaminase. *Journal of Biological Chemistry* **262:** 1772-1778.

Madi, A., Karpati, L., Kovacs, A., Muszbeck, L. and Fesus, L. 2005. High-throughput scintillation proximity assay for transglutaminase activity measurement. *Analytical Biochemistry* **343:** 256-262.

Makarova, K., Aravind, L. and Koonin, E. 1999. A superfamily of archaeal, bacterial, and eukaryotic proteins homologous to animal transglutaminases. *Protein Science.* **8:** 1714-1719.

Manfio, G., Atalan, E., Zakrzewska-Czerwinska, J., Mordarski, M., Rodriguez, C., Collins, M., and Goodfellow, M. 2003. Classification of novel soil streptomycetes as *Streptomyces aureus* sp. nov., *Streptomyces laceyi* sp. nov. and *Streptomyces sanglieri* sp. nov. *Antonie Van Leeuwenhoek* **83:** 245-255.

Marinello, L., Esposito, C., Caputo, I., Sorrentino, A., and Porta, R. 2003. N-terminus end of rat prostate transglutaminase is responsible for its catalytic activity and GTP binding. *The International Journal of Biochemistry and Cell Biology* **35:** 1098-1108.

Matacic, S., and Loewy, A. 1979. Presence of the epsilon-(gamma-glutamic) lysine crosslink in cellular proteins. *Biochim Biophys Acta* **576:** 263-268.

McBride, M., and Ensign, J. 1987. Effects of Intracellular Trehalose Content on *Streptomyces griseus* spores. *Journal of Bacteriology* **169:** 4995-5001.

McDonnell, G., and Russell, A. 1999. Antiseptics and disinfectants: activity, action and resistance. *Clinical Microbiology Reviews* **12:** 147-179.

Mehta, K., Rao, U., Vickery, A., and Birckbichler, P. 1990. Significance of transglutaminase-catalyzed reactions in growth and development of filarial parasite, *Brugia malayi. Biochemical and Biophysics Research Communications* **173:** 1051-1057.

Mehta, K., Rao, U., Vickery, A., and Fesus, L. 1992. Identification of a novel transglutaminase from the filarial parasite *Brugia malayi* and its role in growth and development. *Molecular Biochemistry and Parasitology* **53:** 1-15.

Messer, W., and Zakrzewska-Czerwinska, J. 2002. *Streptomyces* and *Escherichia coli*, model organisms for the analysis of the initiation of bacterial chromosome replication. *Archivum Immunologiae et therapiae experimentalis* **50:** 393-398.

Mongin, E., Louis, C., Holt, R., Birney, E., and Collins, F. 2004. The anopheles gambiae genome: an update. *Trends in Parasitology* **20:** 49-52.

Monroe, A., and Setlow, P. 2006. Localisation of the transglutaminase cross-liniking sites in *Bacillus subtilis* spore coat GerQ. *Journal of Bacteriology* **188:** 7609-7616.

Motoki, M., Okiyama, A., Nonaka, M., Tanaka, H., Uchio, R., Matsuura, A., Ando, H. and Umeda, K. 1991. Production of novel transglutaminase derived from *Streptoverticillium. JP3043080.*

Motoki, M., Atsushi,O., Masahiko,N., Haruo,T., Ryosuke,U., Akira, M., Hiroyasu, A. and Koichi.U. 1992. Transgultaminase. *USPTO No. 5,156,956.*

Murthy, S., Iismaa, S., Begg, G., Freymann, D., Graham, R. and Lorand, L. 2002. Conserved tryptophan in the core domain of transglutaminase is essential for catalytic activity. *Proceedings of the National Academy of Science USA* **99:** 2738-2742.

Mycek, M., Clarke, D., Neidle, A., and Waelsch, H. 1959. Amine incorporation into insulin as catalyzed by transglutaminase. *Archives of Biochemistry and Biophysics* **84:** 528-540.

Nakachi, K., Swift, G., Wilmot, D., Chapman, C., Baker, S., Powell, M., Furmaniak, J. and Rees Smith, B. 2001. Antibodies to tissue transglutaminase: comparison of ELISA and immunoprecipitation assay in the presence and in the absence of calcium ions. *Clin. Chim. Acta.* **304:** 75-84.

Nakanishi, K., Nara, K., Hagiwara, H., Aoyama, Y., Ueno, H., and Hirose, S. 1991. Cloning and sequence analysis of cDNA clones for bovine aortic-endothelial-cell transglutaminase. *European Journal of Biochemistry* **202:** 15-21.

Nelson, K., Weinel, C., Paulsen, I., Dodson, R., Hilbert, H., Martins dos Santos, V., Fouts, D., Gill, S., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J., Timmis, K., Dusterhoft, A., Tummler, B., and Fraser, C. 2002. Complete genome sequence and comparative analysis of the metabolically versatile Pseudomonas putida KT2440. *Environmental Microbiology* **4:** 799-808.

Nemes, Z., Marekov, L. and Steinert, P. . 1999. Involucrin cross-linking by transglutaminase 1. Binding to membranes directs residue specificity. *Journal of Biological Chemistry* **274:** 11013-11021.

Nicieza, R., Huergo, J., Connolly, B. and Sanchez, J. 1999. Purification, characterisation and role of nucleases and serine proteases in streptomyces differentiation. *Journal of Biological Chemistry* **274:** 20366-20375.

Nielsen, P. 1995. Reactions and potential industrial applications of transglutaminase. Review of literature and patents. *Food Biotechnology* **9:** 119-156.

Noguchi, K., Ishikawa, K., Yokoyama, K., Ohtsuka, T., Nio, N. and Suzuki, E. 2001. Crystal structure of red sea bream transglutaminase. *Journal of Biological Chemistry* **276:** 12055-12059.

Nonomura, K., Yamanishi, K., Yasuno, H., Nara, K. and Hirose, S. 1994. Up-regulation of elafin/SKALP gene expression in psoriatic epidermis. *Journal of Investigative Dermatology.* **103:** 88-91.

Nozawa, H., Mamegoshi, S. and Seki, N. 1997. Partial purification and characterization of six transglutaminases from ordinary muscles of various fishes and marine invertebrates. *Computational Biochemistry, Physiology and Molecular Biology* **118:** 313-317.

Ochi, K. 1995b. Comparative ribosomal protein sequence analyses of a phylogenetically defined genus, *Pseudomonas,* and its relatives. *International Journal of Systematic Bacteriology* **45:** 268-273.

Ochi, K. 1989. Heterogeneity of ribosomal proteins among *Streptomyces* species and its application to identfication. *Journal of General Microbiology* **135:** 2635-2642.

Ochi, K. 1994. Phylogenetic diversity in the genus *Bacillus* and comparative ribosomal protein AT-L30 analyses of the genus Thermoactinomyces and relatives. *Microbiology* **140:** 2165-2171.

Ochi, K. 1995a. A taxonomic study of the genus *Streptomyces* by analysis of ribosomal protein AT-L30. *International Journal of Systematic Bacteriology* **45:** 507-514.

Ochman, H., Lawrence, J. and Groisman, E. 2000. Lateral gene transfer and the nature of bacterial innovation. *Nature* **18:** 299-304.

Ochman, H., Lerat, E and Daubin, V. 2005. Examining bacterial species under the specter of gene transfer and echange. *Proceedings of the National Academy of Science, USA* **102:** 6595-6599.

Page, R. 1996. Treeview: An application to display phylogenetic trees on personals computers. . *Computer Applications in the Biosciences* **12:** 357-358.

Paonessa, G., Metafora, S., Tajana, G., Abrescia, P., De Santis, A., Gentile, V. and Porta, R. 1984. Transglutaminase-mediated modifications of the rat sperm surface *in vitro*. *Science* **226:** 852-855.

Parameswaran, K., Cheng, X., Chen, E., Velasco, P., Wilson, J., and Lorand, L. 1997. Hydrolysis of gamma:epsilon isopeptides by cytosolic transglutaminases and by coagulation factor XIIIa. *Journal of Biological Chemistry* **272:** 10311-10317.

Park, D., Kim, J., Lee, S., and Jeong, J. 2002. Secretory expression of the alpha subunit of human coagulation factor XIII in the yeast *Pichia pastoris*. *Biotechnology Letters* **24:** 97-101.

Pasternack, R., Dorsch, S., Otterbach, J., Robenek, I., Wolf, S., and Fuchsbauer, H. 1998. Bacterial pro-transglutaminase from *Streptoverticillium mobaraense*--purification, characterisation and sequence of the zymogen. *European Journal of Biochemistry* **257:** 570-576.

Pasternack, R. 2005. 8th International conference on protein cross-linking enzymes.

Pedersen, L., Yee, V., Bishop, P., Le Trong, I., Teller, D., and Stenkamp, R. 1994. Transglutaminase factor XIII uses proteinase-like catalytic triad to crosslink macromolecules. *Protein Science.* **3:** 1131-1135.

Pfleiderer, C., Mainusch, M., Weber, J., Hils, M., and Fuchsbauer, H. 2005. Inhibition of bacterial transglutaminase by its heat-treated pro-enzyme. *Microbiological Research* **160:** 265-271.

Pietrasik, Z. 2003. Binding and textural properties of beef gels processed with carrageenan, egg albumin and micrbial transglutaminase. *Meat Science* **63**: 317-324.

Pietrasik, Z., and Jarmoluk. 2003. Effect of sodium caseinate and carrageenan on binding and textural properties of pork muscle gels enhanced by microbial transglutaminase. *Food Research International.* **36**: 285-294.

Pisano, J., Finlayson, J. and Peyton, M. 1968. Cross-link in fibrin polymerized by factor 13: epsilon-(gamma-glutamyl)lysine. *Science* **160**: 892-893.

Ragkousi, K., and Setlow, P. 2004. Transglutaminase-mediated cross-linking of GerQ in the coats of *Bacillus subtilis* spores. *Journal of Bacteriology* **186**: 5567-5575.

Rasmussen, L., Mollgaard, A., Petersen, B., and Sorensen, N. 1996. Method for casein finishing of leather. *WO9413839.*

Redfield, R. 2001. Do bacteria have sex? *Nature Reviews Genetics* **2**: 634-639.

Rey, M., Ramaiya, P., Nelson, B., Brody-Karpin, S., Zaretsky, E., Tang, M., Lopez de Leon, A., Xiang, H., Gusti, V., Clausen, I., Olsen, P., Rasmussen, M., Andersen, J., Jorgensen, P., Larsen, T., Sorokin, A., Bolotin, A., Lapidus, A., Galleron, N., Ehrlich, S., and Berka, R. 2004. Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. *Genome Biology* **5**: R77.

Rice, R., and Green, H. 1977. The cornified envelope of terminally differentiated human epidermal keratinocytes consist of cross-linked proteins. *Cell* **11**: 417-422.

Ruan, J., Lang, Y., Shi, Y., Qu, L and Yu, X. 1994. Chemical and molecular classification of *Saccharomonspora* strains. *International Journal of Systematic bacteriology* **44**: 704-707.

Russell, A. 1990. Bacterial spores and chemical sporicidal agents. *Clinical Microbiology Reviews* **3**: 99-119.

Santangelo, J., Kuhn, A., Treuner-Lange, A., and Durre, P. 1998. Sporulation and time course expression of sigma-factor homologous genes in *Clostridium acetobutylicum. FEMS Microbiology Letters* **161**: 157-164.

Schleifer, K., and Stackebrandt, E. 1983. Molecular systematics of Prokaryotes. *Annual Review of Microbiology* **37**: 143-187.

Schmidt, G., Selzer, J., Lerm, M., and Aktories, K. 1998. The Rho-deamidating cytotoxic necrotizing factor 1 from *Escherichia coli* possesses transglutaminase activity. Cysteine 866 and histidine 881 are essential for enzyme activity. *Journal of Biological Chemistry* **273**: 13669-13674.

Seiler, N. 2004. Catabolism of polyamines. *Amino Acids* **26**: 217-233.

Seitz, J., Keppler, C., Hunteman, S., Rausch, U. and Aumuller, G. 1991. Purification and molecular characterisation of a secretory transglutaminase from coagulating gland of the rat. *Biochim. Biophys. Acta.* **1078**: 139-146.

Serafini-Fracassini, D., DelDuca, S. and Beninati, S. 1995. Plant Transglutaminases. *Phytochemistry* **40**: 355-396.

Serafini-Fracassini, D., and DelDuca, S. . 2002. Biochemistry and function of plant transglutaminases. *Minerva Biotechnology* **14**: 135-141.

Setlow, P. 1988. Resistance of bacterial spores to ultraviolet light. *Comments on Molecular and Cell Biophysics* **5**: 253-264.

Setlow, P. 1999. Bacterial spore resistance. *In G. Storz and R. Hengge-Aronis (ed.), Bacterial Stress responses. American Society for Microbiology, Washington D.C*: 217-230.

Shen, J., Bishop, D., Heine, E. and Hollfelder B. 1999. Some factors affecting the control of proteolytic enzymes reactions on wool. *Journal of Textiles Inst* **90**: 404-411.

Shimba, N., Shinohara, M., Yokoyama, K., Kashiwagi, T., Ishikawa, K., Ejima, D. and Suzuki, E. 2002. Enhancement of transglutaminase activity by NMR identification of its flexible residues affecting the active site. *FEBS Letters* **517**: 175-179.

Siepaio, M., and Meuier, J. 1995. Diamine oxidase and transglutaminase activities in white lupine seedlings with respect to cross-liking of proteins. *Journal of Agricultural Food Chemistry* **43**: 1151-1156.

Simon, M., and Green, H. 1988. The glutamine residues reactive in transglutaminase-catalysed cross-linking of involucrin. *Journal of Biological Chemistry* **263**: 18093-18098.

Slaughter, T., Achyuthan, K., Lai, T., and Greenberg, C. . 1992. A microtiter plate transglutaminase assay utilizing 5-(biotinamido)pentylamine as substrate. *Analytical Biochemistry* **205**: 166-171.

Soutourina, J., Plateau, P., Delort, F., Periotes. A. and Blanquet, S. 1999. Functional characterisation of d-Tyr-tRNA tyr deacylase from *Escherichia coli. Journal of Biological Chemistry* **274**: 19109-19114.

Stackebrandt, E., and Woese, C.R. 1981. Towards a phylogeny of the actinomcyetes and related organisms. *Current Microbiology* **5**: 197-202.

Stackebrandt, E., Witt, D., Kemmerling, C., Kroppenstedt, R. and Liesack, W. 1991. Designation of Streptomycete 16S and 23S rRNA-based target regions for oligonucleotide probes. *Applied Environmental Microbiology* **57:** 1468-1477.

Sugai, M., Hatazaki, K., Mogami, A., Ohta, H., Peres, S., Herault, F., Horiguchi, Y., Masuda, M., Ueno, Y., Komatsuzawa, H., Suginaka, H., and Oswald, E. 1999. Cytotoxic necrotizing factor type 2 produced by pathogenic *Escherichia coli* deamidates a gln residue in the conserved G-3 domain of the rho family and preferentially inhibits the GTPase activity of RhoA and rac1. *Infection and Immunity* **67:** 6550-6557.

Sung, L., Chien, S., Fan, Y., Lin, C., Lambert, K., Zhu, L., Lam, J. and Chang, C. 1992. Human erythrocyte protein 4.2: isoform expression, differential splicing and chromosome assignment. *Blood* **79:** 2763-2770.

Suzuki, S., Kobayashi, K., Yamanaka, S. and Yokozeki, K. 1997. Protease catalyzed reaction of polyamine incorporation into protein. *Bioscience, Biotechnology and Biochemistry* **61:** 1609-1612.

Suzuki, S., Izawa, Y., Kobayashi, K., Eto, Y., Yamanaka, S., Kubota, K., and Yokozeki, K. 2000. Purification and characterization of novel transglutaminase from *Bacillus subtilis* spores. *Bioscience, Biotechnology and Biochemistry* **64:** 2344-2351.

Taguchi, S., Nishihama, K., Igi, K., Taira, H., Motoki, M., and Momose, H. 2000. Substrate specificity analysis of microbial transglutaminase using proteinaceous protease inhibitors as natural model substrates. *Journal of Biochemistry (Tokyo)* **128:** 415-425.

Takagi, T., and Doolittle, R. 1974. Amino acid sequence studies on factor XIII and the peptide released during activation by thrombin. *Biochemistry* **13:** 750-756.

Takehana, K., Washizu, K., Ando, S., Koikeda, K., Takeuchi, H., Matsui, M., Motoki, M. and Takagi, H. 1994. Chemical synthesis of the gene for microbial transglutaminase from *Streptoverticillium* and its expression in *Escherichia coli. Bioscience, Biotechnology and Biochemistry* **58:** 88-92.

Tamaki, T., and Aoki, N. 1981. Crosslinking of aplha 2-plasmin inhibitor and fibronectin to fibrin by fibrin-stabilizing factor. *Biochem. Biophys. Acta.* **661:** 280-286.

Tang, S., Le, Q., Shim, J., Yang, S., Auh, J., Park, C. and Park, K. 2006. Enhancing thermostability of maltogenic amylase from *Bacillus thermoalkalophilus* ET2 by DNA shuffling. *FEBS Journal* **273:** 3335-3345.

Tatusov, R., Fedorova, N., Jackson, J., Jacobs, A., Kiryutin, B., Koonin, E., Krylov, D., Mazumder, R., Mekhedov, S., Nikolskaya, A., Rao, B., Smirnov, S., Sverdlov, A., Vasudevan, S., Wolf, Y., Yin, J., and Natale, D. 1999. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4:** 41.

Tavladoraki, P., Rossi, M., Saccuti, G., Perez-Amador, M., Polticelli, F., Angelini, R. and Fererico, R. 2006. Heterologous expression and biochemical characterisation of a polyamine oxidase from *Arabidopsis* involved in polyamine back conversion. *Plant Physiology* **141:** 1519-1532.

Thomazy, V., and Fesus, L. 1989. Differential expression of tissue transglutaminase in human cells. An immuno-histochemical study. *Cell Tissue Research* **255:** 215-221.

Towbin, H., Staehelin, T. and Gordon, J. 1979. Electrophoretic transfer of proteins from polyacrylamide gells to nitrocellulose sheets: procedure and some applications. *Proceedings of the National Academy of Science USA* **76:** 4350-4354.

Traore, F., and Meunier, J. 1992. Cross-linkig of caseins by human placental factor XIII. *Journal of Agriculture and Food Chemistry* **39:** 1892-1896.

Trigwell, S., Lynch, P., Griffin, M., Hargreaves, A., and Bonner, P. 2004. An improved colorimetric assay for the measurement of transglutaminase (type II) -(gamma-glutamyl) lysine cross-linking activity. *Analytical Biochemistry* **330:** 164-166.

Tsai, G., Lin, S. and Jiang, S. 1996. Transglutaminase from *Streptoverticillium ladakanum* and application to minced fish product. *Journal of Food Science* **61:** 1234-1238.

Tzeng, S., Yin, L., Chen, G. and Jiang, S. 2005. Molecular cloning of microbial transglutaminase from *Streptoverticillium ladakanum* BCRC 12422. *Journal of Fisheries Society* **32:** 53-67.

Vasbinder, A., Rollema, H., Bot, A. and de Kruif, C. 2003. Gelation mechanism of milk as influenced by temperature and pH; studied by the use of transglutaminase cross-linked casein micelles. *Journal of Dairy Science* **86:** 1556-1563.

Verderio, E., Johnson, T. and Griffin, M. 2004. Tissue transglutaminase in normal and abnormal wound healing: review article. *Amino Acids* **26:** 387-404.

Wada, J., Nakamura, A., Masutani, T., Ikura, K., Maki, M. and Hitomi, K. 2002. Identification of mammalian-type transglutaminase in *Physarum polycephalum*. *European Journal of Biochemistry* **269:** 3451-3460.

Walker, K., and Weiss, A. 1994. Characterisation of the dermonecrotic toxin in members of the genus *Bordetella*. *Infection and Immunity* **62:** 3817-3828.

Wang, J., Zhao, M., Yang, X., Jiang, Y., and Chun, C. 2007. Gelation behaviour of wheat gluten by heat treatment followed by transglutaminase cross-linking reaction. *Food Hydrocolloids* **21:** 174-179.

Washizu, K., Ando, K., Koikeda, S., Hirose, S., Matsuura, A., Takagi, H., Motoki, M. and Takeuchi, K. 1994. Molecular cloning of the gene for microbial transglutaminase from *Streptoverticillium* and its expression in *Streptomyces lividans*. *Bioscience, Biotechnology and Biochemistry* **58:** 82-87.

Wellington, E., and Williams, S. 1981. Host ranges of phages isolated to *Streptomyces* and other genera. *Zentbl Bakteriol Hyg I Abt Suppl* **11:** 93-98.

West, J., Logan, R., Hill, P., Lloyd, A., Lewis, S., Hubbard, R., Reader, R., Holmes, G., and Khaw, K. 2003. Seroprevalence, correlates, and characteristics of undetected coeliac disease in England. *Gut* **52:** 960-965.

Wijngaards, G., and Paardekooper, E. 1988. Preparation of a composite meat product by means of enzymatically formed protein gel. *Trends in Modern Meat Technology* **2:** 125-129.

Williams, S., Goodfellow, M., Aldersen, G.,Wellington, E., Sneath, P. and Sackin, M. . 1983. Numerical Classification of Streptomyces related genera. *Journal of General Microbiology* **129:** 1743-1813.

Williams-Ashman, H. 1984. Transglutaminase and the clotting of mammalian seminal fluids. *Molecular and Cellular Biochemistry* **58:** 51-61.

Wilson, S. 1992. Modifying meatproteins via enzymatic crosslinking. *Proceedings of the 27th Meat Industry Research Conference, Hamilton*: 247-277.

Witt, D., and Stackebrandt, E. 1990. Unification of the genera *Streptoverticillium* and *Streptomyces* and amendation of *Streptomyces* Waksman and Henrici 1943, 339[AL]. *Systematic and Applied Microbiology* **13:** 361-371.

Wright, F., and Bibb, M. 1992. Codon usage in the G + C- rich *Streptomyces* genome. *Gene* **113:** 55-65.

Xu, D., and Cote, J. 2003. Phylogenetic relationships between *Bacillus* species and related genera inferred from comparison of 3` end 16S rDNA and 5` end 16S-23S ITS nucleotide sequences. *International Journal of Systematic and Evolutionary Microbiology* **53:** 695-704.

Xu, L., Tiang, Y., Zhang, Y., Zhao, L., and Jiang, C. 1998. *Streptomyces thermogriseus*, a new species of the genus *Streptomyces* from soil, lake and hot-spring. *International Journal of Systematic Bacteriology* **48:** 1089-1093.

Yaffe, M., Beegen, H. and Eckert, R. 1992. Biophysical characterisation of involucrin reveals a molecule ideally suited to function as an intermolecular cross-bridge of the keratinocyte cornified envelope. *Journal of Biological Chemistry* **267:** 12233-12238.

Yan, G., Guocheng, D., Li, Y., Chen, J. and Zhong, J. 2005. Enhancement of microbial transglutaminase production by *Streptoverticillium mobaraense*: application of a two-stage agitation spedd control strategy. *Process Biochemistry* **40:** 963-968.

Yanagita, T., and Nomachi, Y. 1967. Kinetic analysis of the region or protease formation in the hypha of *Aspergillus niger*. *Journal of General and Applied Microbiology* **13:** 227-235.

Yasueda, H., Kumazawa, Y., and Motoki, M. 1994. Purification and characterization of a tissue-type transglutaminase from red sea bream (*Pagrus major*). *Bioscience, Biotechnology and Biochemistry* **58:** 2041-2045.

Yee, V., Pedersen, L., Le Trong, I., Bishop, P., Stenkamp, R. and Teller, D. 1994. Three-dimensional structure of a transglutaminase: human blood coagulation factor XIII. *Proceedings of the National Academy of Science USA* **91:** 7296-7300.

Yokoyama, K., Nakamura, K., Seguro, K. and Kubota, K. 2000. Overproduction of microbial transglutaminase in *Escherichia coli*, in vitro refolding, and characterisation of the refolded form. *Bioscience, Biotechnology and Biochemistry* **64:** 1263-1270.

Yurimoto, H., Yamane, M., Kikuchi, Y., Matsui, H., Kato, N. and Sakai, Y. 2004. The pro-peptide of *Streptomyces mobaraensis* transglutaminase functions in cis and trans to mediate efficient secretion of active enzyme from methylotrophic yeasts. *Bioscience, Biotechnology and Biochemistry* **68:** 2058-2069.

Zavalova, L., Lukyanov, S., Baskova, I., Snezhkov, E., Akopov, S., Berezhnoy, S., Bogdanova, E., Barsova, E., and Sverdlov, E. 1996. Genes from the medicinal leech (*Hirudo medicinalis*) coding for unusual enzymes that specifically cleave endo-epsilon (gamma-Glu)-Lys isopeptide bonds and help to dissolve blood clots. *Molecular and General Genetics* **253:** 20-25.

Zheng, L., Halberg, R., Roels, S., Ichikawa, H., Kroos, L., and Losick, R. 1992. Sporulation regulatory protein GerE from *Bacillus subtilis* binds to and can activate or repress transcription from promoters for mother-cell-specific genes. *Journal of Molecular Biology* **226:** 1037-1050.

Zhu, Y., Rinzema.A, Tramper.J and Bol.J. 1995. Microbial transglutaminase- a review of its production and application in food processing. *Applied Microbiology and Biotechnology* **44:** 277-282.

Zhu, Y., Rinzema, A. and Tramper, J. 1996. Medium design based on stoichiometric analysis of microbial transglutaminase production by *Streptoverticillium mobaraense,. Biotechnology and Bioengineering* **50:** 291-298.

Zotzel, J., Pasternack, R., Pelzer, C., Ziegert, D., Mainusch, M. and Fuchsbauer, H. 2003b. Activated transglutaminase from *Streptomyces mobaraensis* is processed by a tripeptidyl aminopeptidase in the final step. *European Journal of Biochemistry* **270:** 4149-4155.

Zotzel, J., Keller, P. and Fuchsbauer, H. 2003a. Transglutaminase from *Streptomyces mobaraensis* is activated by an endogenous metalloprotease. *European Journal of Biochemistry* **270:** 3214-3222.

Appendix

Appendix I: Commonly Used Buffers

5x SDS-PAGE Sample buffer: 40% (v/v) glycerol,
10% (w/v) SDS,
50 mM DTT,
0.2 M Tris-HCl pH 6.8

50x TAE: 2 M Tris,
1 M glacial acetic acid,
1 mM EDTA pH 8

Agarose gel Loading Buffer: 0.25% (w/v) bromophenol blue,
40% (w/v) sucrose

Bacterial Lysis Solution: 50 mMTris HCl pH7.4,
1.5 mg/ml lysozyme,
1 mM DTT,
2 mM EDTA,
1 mM PMSF

Coomassie Blue stain: 0.5% w/v Coomassie brilliant blue,
10% glacial acetic acid,
50% methanol

PBS: 7.5 mM Na_2HPO_4,
2.5 mM $NaH_2PO_4.2H_2O$,
145 mM NaCl

SDS-PAGE Running Buffer: 1.5% (w/v) Tris Base,
7% (w/v) glycine,
0.5% (w/v) SDS

TBS pH 7.4:   20 mM Tris base,
137 mM NaCl,
pH adjusted with HCl

TBS Tween 20:   0.1% Tween 20 in TBS

Transfer Buffer (Towbin):   25 mM Tris,
192 mM glycine,
20% (v/v) methanol

Appendix II: Vectors

pET21d(+) (Novagen, UK)
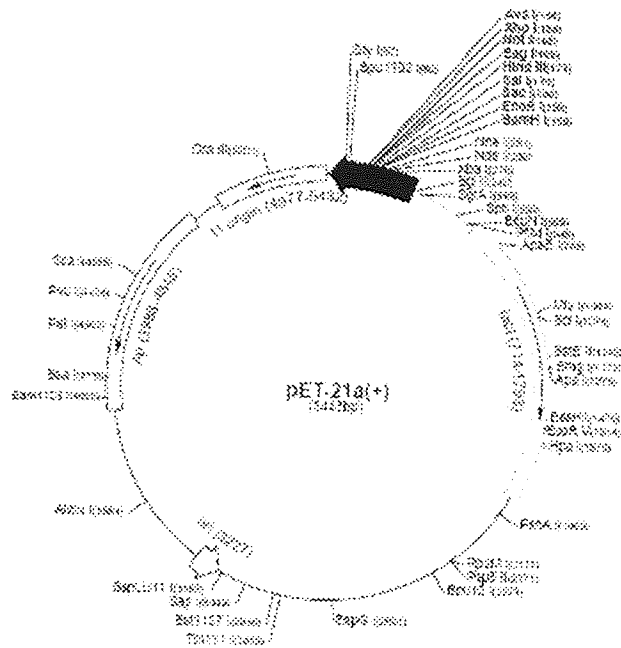
pET22b(+) (Novagen, UK)
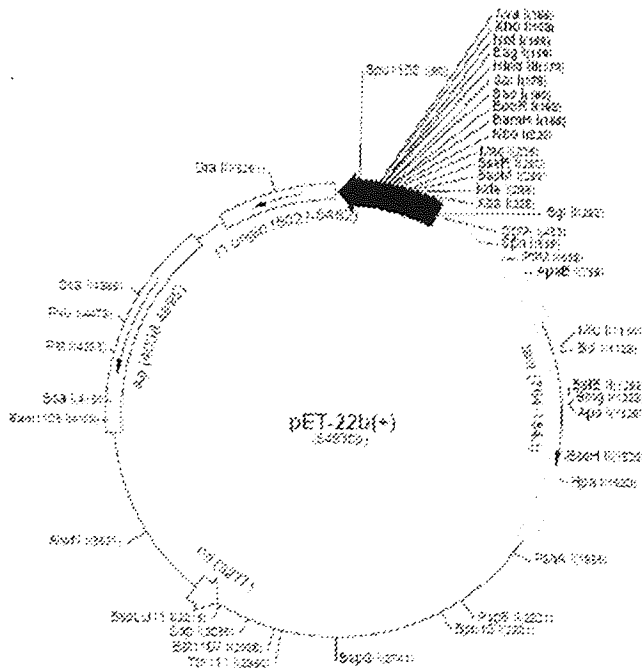
pWH1520 (MoBiTec, Germany)
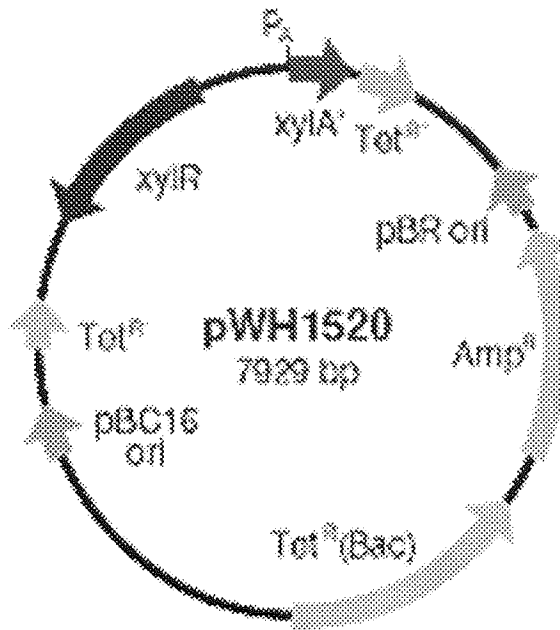
pGEMT (Promega, UK)
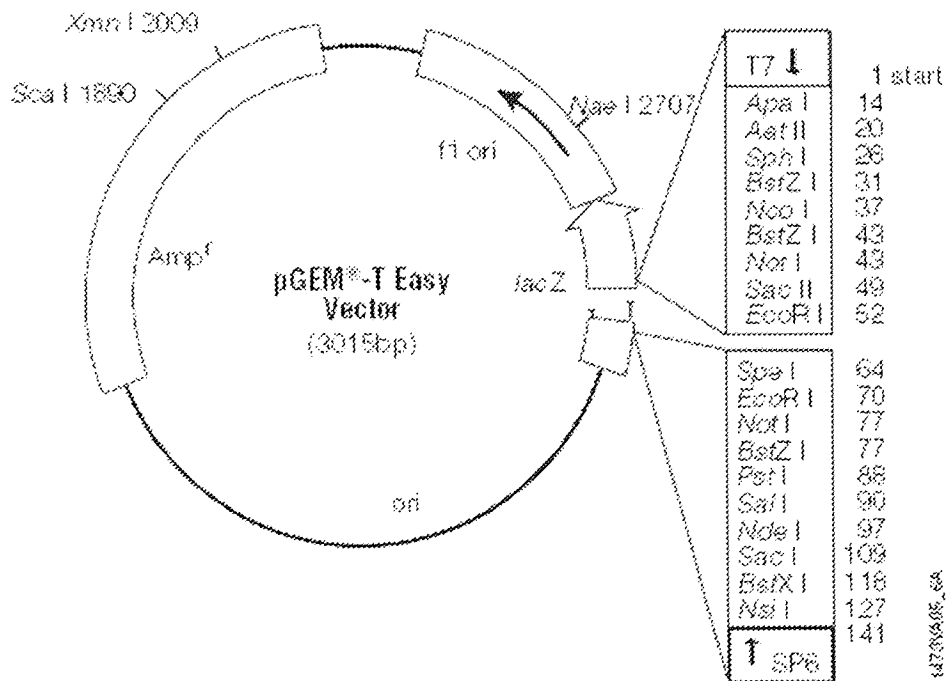
pGEX (Amersham Bioscience, UK)
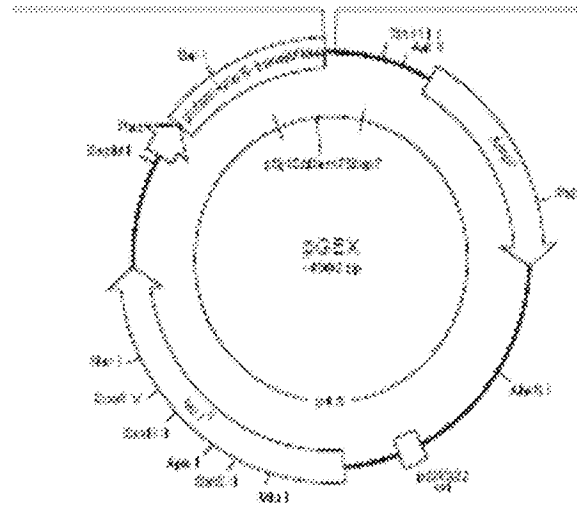
Appendix III: Micro organisms

Growth media used are indicated in parentheses corresponding to those listed in Appendix IV.

A: E.coli *Strain Genotypes*

Grown in medium no. 4.

E.coli DH5α (Stratagene); genotype: F' Phi80dlacZ δM15 δ(lacZYA-argF)U169 deoR recA1 endA1hsdR17(rK-mK+)phoA supE44 λ- thi-1

E. coli BL21 (DE3)(Novagen); genotype: F-, mpT, hsdSβ(rβ-mβ-), dcm, gal, (DE3) tonA

E. coli BL21 (DE3) pLysS (Novagen); genotype: F-, OmpT, hsdSβ(rβ-mβ-), dcm, gal, (DE3), pLysS(CmR) tonA

E.coli Tuner(DE3) (Novagen); genotype: F− OmpT hsdSB(rB− mB−) gal dcm lacY1(DE3)

E. coli NovaBlue (Novagen); genotype endA1 hsdR17(rK12mK12+) supE44 thi-1recA1 gyrA96 relA1 lacF'[proA+B+ lacIqZ .M15 ::Tn10(TcR)]

B: Fungi and Yeast

Aspergillus fumigatus	NRRL 2869	(2,7,9)
Aspergillus nidulans	ATCC 10074	(2,7,9)
Aspergillus niger	ATCC 1004	(2,7,9)
Aspergillus oryzae	ATCC 12892	(2,7,9)
Aspergillus parasiticus	lab strain	(2,7,9)
Aspergillus sojae	ATCC 46250	(2,7,9)
Aspergillus terreus	TUB F-177	(2,7,9)
Aspergillus terreus	IFO 089356	(2,7,9)
Aspergillus terreus	IFO 36364	(2,7,9)
Aspergillus terreus	IFO 20542	(2,7,9)
Aspergillus terreus	IFO 7079	(2,7,9)
Aspergillus versicolor	ATCC 26939	(2,7,9)
Circinella nigra	lab strain	(13)
Curvularia fakata	lab strain	(10)
Eremascus albus	lab strain	(5)
Fusarium monoliforme	lab strain	(2,7,9)
Neurospora tetrasperma	lab strain	(2,7,9)
Octosporomyces octosporus	IMI 140427	(11)
Penicillium chrysogenum	ATCC 10002	(2,7,9)
Penicillium claviforme	lab strain	(2,7,9)
Penicillium roquefortii	lab strain	(2,7,9)
Saccharomyces cerevisiae	ATCC 9763	(11)

C: Bacteria

Acinetobacter calcoaceticus	ATCC 15308	(1,4,6)
Alcaligenes eutrophus	lab strain	(1,4,6)
Arthrobacter globiformis	lab strain	(1,4,6)
Arthrobacter simplex	lab strain	(1,4,6)
Bacillus subtilis	NCIMB 10106	(8,4)
Bacillus licheniformis	NCIMB 8874	(8,4)
Bacillus megaterium	NTUCC 323	(8,4)
Bacillus polymyxa	Lab Strain	(8,4)
Bacillus brevis	NCIMB 9372	(8,4)

Brevibacterium helvolum	NCTC 11083	(1,4,6)
Brevibacterium vitarumenis	lab strain	(1,4,6)
Lactobacillus plantarum	NCIMB 6376	(1,4,6)
Lactobacillus casei	NTUCC 676	(1,4,6)
Micrococcus luteus	NCIMB 196	(1,4,6)
Micrococcus roseus	NTUCC 394	(1,4,6)
Neisseria pharynges	NCTC 4590	(4,12)
Nocardia agarocarbon-oxydans	Lab strain	(2,3,9)
Pseudomonas mendocina	ATCC 25411	(4,6,1)
Pseudomonas putida	NCIMB 9782	(4,6,1)
Pseudomonas putida	NCIMB 9876	(4,6,1)
Pseudomonas putida	ATCC 29607	(4,6,1)
Pseudomonas vesicularis	ATCC 11426	(4,6,1)
Proteus mirabilis	ATCC 14153	(4,6)
Saccharomonospora viridis	ATCC 15386	(2,3,9)
Saccharopolyspora hirsuta sp. Hirsuta	ATCC 27875	(2,3,9)
Staphylococcus aureus	ATCC 6538	(4,6)
Staphylococcus epidermidis	NCTC 7944	(4,6)
Streptococcus durans	ATCC 11576	(4,6)
Streptomyces albidoflavus	NRRL B-1271	(2,3,9)
Streptomyces aureofaciens	NTUCC 215	(2,3,9)
Streptomyces badius	252 ATCC 39117	(2,3,9)
Streptomyces baldaccii	NRRL B-3500	(2,3,9)
Streptomyces chartreusis	TUB B-122 (ATCC 43360)	(2,3,9)
Streptomyces griseus	IFO 13427	(2,3,9)
Streptomyces mobaraensis	NRRL B-3729	(2,3,9)
Streptomyces omiyaensis	TUB B-125 (ATCC 43363)	(2,3,9)
Streptomyces paucisporogenes	ATCC 12596	(2,3,9)
Streptomyces phaecochromogenes	TUB B-118 (ATCC 43363)	(2,3,9)
Streptomyces platensis	ATCC 13865	(2,3,9)
Streptoverticillium waksmanii	NTUCC 538	(2,3,9)
TUB B-549 (unidentified Streptomycete, thermotolerant)		(2,3,9)
TUB B-718 (unidentified streptomycete, thermotolerant)		(2,3,9)
Xanthobacter autotrophicus	DSMZ 853	(1,4,6)
Xanthobacter flavus	lab strain	(1,4,6)

Appendix IV: Growth Media

1. Corynebacterium medium: 10 g/l casein peptone,

5 g/l yeast extract,

5 g/l glucose,

5 g/l sodium chloride

2. D-1:   15 g/l soybean meal,

5 g/l corn steep liquor,

15 g/l glucose,

5 g/l NaCl,

2g/l $CaCO_3$, pH 7.2

3. GYM: 4 g/l glucose,

4 g/l yeast extract,

10 g/l malt extract, pH 7.2

4. LB: 10 g/l tryptone,

5 g/l yeast extract,

10 g/l NaCl, pH 7.4

5. M 40 Y: 400 g/l sucrose,

20 g/l malt extract,

5 g/l yeast extract

6. NB: 5 g/l peptone,

3 g/l meat extract

7. PDB: 4 g/l potato infusion

(from 200g potatoes),

20 g/l dextrose, pH 5.1

8. Schaeffer's medium:

0.8% nutrient broth,

27 mM KCl,

2 mM $MgSO_4 \cdot 7H_2O$,

1 mM $Ca(NO_3)_2 \cdot 4H_2O$,

0.1 mM $MnCl_2 \cdot 4H_2O$,

0.2 1 µM $FeSO_4 \cdot 7H_2O$

9. TG-7:

15 g/l soybean meal,

5 g/l corn meal,

15 g/l sucrose,

5 g/l NaCl,

2 g/l $CaCO_3$, pH 6.5

10. V-8 Juice medium:

20% (v/v) V-8 juice,

3 g/l $CaCO_3$ pH7.3

11. YM:

3 g/l yeast extract,

3 g/l malt extract,

5 g/l soybean peptone,

10 g/l glucose

12. YpSs medium:

4 g/l yeast extract,

15 g/l soluble starch,

1 g/l K_2HPO_4,

0.5 g/l $MgSO_4.7H_2O$

Appendix V: Pseudomonas PCR

Differences to the published sequence are highlighted in red. Translated amino acid sequence shown. Bases in green represent those originating from the vector. * denotes stop codon.

```
  1   gccaaactgtccattcgccacgacacgacctatagctacgccagc
          K   L   S   I   R   H   D   T   T   Y   S   Y   A   S

 46   gacgtgtgcaacagtatccagttcctgcgcctgacgccgcgcagc
          D   V   C   N   S   I   Q   F   L   R   L   T   P   R   S

191   agcgaacgccagcgtatcaatcaatggcaactggacctgccctgc
          S   E   R   Q   R   I   N   Q   W   Q   L   D   L   P   C

136   aaggtcaagggccagatcgacccttacggcaacatcctgcatgtg
          K   V   K   G   Q   I   D   P   Y   G   N   I   L   H   V

181   ctgacactggacaaaccccatgcccacctggcgcttaccgcacac
          L   T   L   D   K   P   H   A   H   L   A   L   T   A   H

226   ggccaggtcgagatcgacccggattgcgaacacgaaaccgagagc
          G   Q   V   E   I   D   P   D   C   E   H   E   T   E   S

271   cagtcgcccctgccgttcctgcgcggcagccacctgacgcaggcc
          Q   S   P   L   P   F   L   R   G   S   H   L   T   Q   A

316   gacgacaccctgactgctttcgccgcccggcaatgtggcagtcac
          D   D   T   L   T   A   F   A   A   R   Q   C   G   S   H

361   cgcgaccgggcagcactgatcgggctgatgcaaggcctggccgag
          R   D   R   A   A   L   I   G   L   M   Q   G   L   A   E

406   cacatgccctacagccctggggccacgtcagtgggcaccacggcc
          H   M   P   Y   S   P   G   A   T   S   V   G   T   T   A

451   atcgaagccttcaacggtggcgcaggtgtgtgtcaggaccacacg
          I   E   A   F   N   G   G   A   G   V   C   Q   D   H   T

496   cacgcgttcctggcctgcgcacgcagcctgggggtaccggcacgc
          H   A   F   L   A   C   A   R   S   L   G   V   P   A   R

441   tatgtgtccggctacctgtgcaccgaggacgaacaacacctggcc
          Y   V   S   G   Y   L   C   T   E   D   E   Q   H   L   A

486   agccacgcctgggccgaagcctgggtcgacgatgcctggtacagc
          S   H   A   W   A   E   A   W   V   D   D   A   W   Y   S

531   ttcgacattaccaaccgcctgacccggccagagcgtcacctgaag
          F   D   I   T   N   R   L   T   R   P   E   R   H   L   K
```

```
576 ctggcggtgggcctggactaccttgatgcctgcccggtcaggggc
     L  A  V  G  L  D  Y  L  D  A  C  P  V  R  G
621 gtacgccgcggcggcggcgccgaatcgatgcaggccagcgttcac
     V  R  R  G  G  G  A  E  S  M  Q  A  S  V  H
666 gtgcatcgccagtga 681
     V  H  R  Q  *
```

```
576 ctggcggtgggcctggactaccttgatgcctgcccggtcaggggc
     L  A  V  G  L  D  Y  L  D  A  C  P  V  R  G
```

Appendix VI: Pseudomonas putida Degenerate PCR Sequence

Degenerate PCR carried out using *Ps. putida* genomic DNA as template. PCR was carried out using primers F1 and R1 with an annealing temperature of 50 °C for 30 s as described in Section 2.11.6. The product was gel purified and subcloned into pGEM-T (Promega, UK) for sequencing. Sequence in blue denotes that of the vector.

```
1    GAATTCGATTTTGCATCGACGCGTGCGACAAGGCGGCTATGAGGGACAAA
51   GAGGAATTGCAGGATTGCTTCATGAACAGCTCTTGTAGGAATCGTGGTGA
101  CGAATGGGGTCGCAGGCCCGGCCCTGCCTAGGCCTGCGGCTCCATGCGCA
151  TGGGAATGGTGACGGGCCCGTCATTGACCAGGTGCACCTGCATGTCGGCG
201  GCGAACTCGCCGGTCTGCACCTGAGCGTGGGCAAGCCTGGCCTGGTCCAC
251  AAAGTATTCGTAGAGGCGTCGGCCTTCGTCGGGCGCGGCGGCGGCGGTAA
301  AGCTGGGGCGATTGCCGCCTCTGGTATCGGCCGCCAGCGTGAACTGGCTG
351  ACCACCAGCAGCCCTCCGCCAATGTCCTGCAGGCTTTTGTTCATCTTTCC
401  GGCCTCGTCGCTGAAGATGCGCAGCTTCAGCATCTTGGCCAGCAACTTGT
451  CGGCCTCAGCCATTGTGGTACTTAACCAATCACTAGT 488
```

Appendix VII: *Streptomyces* Degenerate PCR Nucleotide Alignment

Genomic DNA was isolated from Streptomyces species and used as starting material for PCR using primers F2 and R2. After PCR, described in Section 2.11.6, bands were gel purified and sub-cloned in pGEM-T (Promega, UK) for sequencing.
The corresponding sequence from *S. mobaraensis* has been shown for reference. Differences to the sequence of *S. mobaraensis* TGase are highlighted in red.

```
S. mobaraensis      CAGCAGATGACCGAGGAGCAGCGGGAGTGGCTGTCCTACG
S.viridis           CAGCAGATGACCGAGGAGCAGCGGGAGTGGCTGTCCTACG
S.paucisporogenes   CAGCAGATGACNGAGGAGCAGCGAGAACAGCTGTCCTACG
S.platensis         CAGCAGATGACNGAGGAGCAGCGAGAACAGCTGTCCTACG
S.baldaccii         CAGCAGATGACGGAGGAGTAGCGAGAAAAGCTGTCCTACG


S. mobaraensis      GCTGCGTCGGTGTCACCTGGGTCAATTCGGGTCAGTACCC
S.viridis           GCTGCGTCGGTGTCACCTGGGTCAATTCGGGTCAGTACCC
S.paucisporogenes   GCTGCGTCGGCGTCACCTGGGTCAATACAGGCCCCTACCC
S.platensis         GCTGCGTCGGCNTCACCTGGGTCAATACAGGCCCCTACCC
S.baldaccii         GTTGCGTCGGCGTCACCTGGGTCAATTCGGGTCCCTACCC


S. mobaraensis      GACGAACAGACTGGCCTTCGCGTCCTTCGACGAGGACAGG
S.viridis           GACGAACAGACTGGCCTTCGCGTCCTTCGACGAGGACAGG
S.paucisporogenes   GACGAACAGACTGGCGTTCGCGTTCTTCGACGAGAACAAG
S.platensis         GACGAACAGACTGGCGTTCGCGTTCTTCGACGAGAACAAG
S.baldaccii         GACGAACAAATTGGCGTTTGCGTCCTTCGACGAGAACGAG


S. mobaraensis      TTCAAGAACGAGCTGAAGAACGGCAGGCCCCGGTCCGGCG
S.viridis           TTCAAGAACGAGCTGAAGAACGGCAGGCCCCGGTCCGGCG
S.paucisporogenes   TACAAGAACGATCTGGAAAAACAGCCGACCGCAACCCAACG
S.platensis         TACAAGAACGATCTGGAAAAACAGCCGACCGCAACCCAACG
S.baldaccii         TACAAGAACGACCTCAAAAACATCAGCCCCCGACCCGGGG


S. mobaraensis      AGACGCGGGCGGAGTTCGAGGGCCGCGTCGCGAAGGAGAG
S.viridis           AGACGCGGGCGGAGTTCGAGGGCCGCGTCGCGAAGGAGAG
S.paucisporogenes   AGACGCAGGCGGAGTTCGAGGGGCGCATCGCCAAGGGCAG
S.platensis         AGACGCAGGCGGAGTTCGAGGGGCGCATCGCCAAGGGCAG
S.baldaccii         AAACGCGGGCAGAGTTCGAGGGCCGCATCGCCAAGGACAG


S. mobaraensis      CTTCGACGAGGAGAAGGGCTTCCAGCGGGCGCGTGAGGTG
S.viridis           CTTCGACGAGGAGAAGGGCTTCCAGCGGGCGCGTGAGGTG
S.paucisporogenes   TTTCGATGAGGGGAAGGGTTTCAAGCGGGCGCGTGAGGTG
S.platensis         TTTCGATGAGGGGAAGGGTTTCAAGCGGGCGCGTGAGGTG
S.baldaccii         TTTCGACGAGGAGAAGGGCTTCAAGCGGGCGCGTGATGTG
```

```
S. mobaraensis    GCGTCCGTCATGAACAGGGCCCTGGAGAACGCCCACGACG
S.viridis         GCGTCCGTCATGAACAGGGCCCTGGAGAACGCCCACGACG
S.paucisporogenes GCGTCCGTCATGAACAAGGCCCTGGACAGCGCGCACGACG
S.platensis       GCGTCCGTCATGAACAAGGCCCTGGACAGCGCGCACGACG
S.baldaccii       GCGTCCATCATGAACAAGGCCCTGGAAAGCGCCCACGACG

S. mobaraensis    AGAGCGCTTACCTCGACAACCTCAAGAAGGAACTGGCGAA
S.viridis         AGAGCGCTTACCTCGACAACCTCAAGAAGGAACTGGCGAA
S.paucisporogenes AGGGGACGTACATCGACCCACCTCAAGACAGAGCTCGCGAA
S.platensis       AGGGGACGTACATCGACCCACCTCAAGACAGAGCTCGCGAA
S.baldaccii       AGGGGGCGTACATCGACAACCTCAAGACGGGGCTCACGAA

S. mobaraensis    CGGCAACGACGCCCTGCGCAACGAGGACGCCCGTTCCCCG
S.viridis         CGGCAACGACGCCCTGCGCAACGAGGACGCCCGTTCCCCG
S.paucisporogenes CAAAAACGACGCTCTGCTCTACGAGGACAGCCGCTCAAGC
S.platensis       CAAAAACGACGCTCTGCTCTACGAGGACAGCCGCTCAAGC
S.baldaccii       CAACAATGACGCTCTGCTCTACGAGGACAGCCGATCGAAC

S. mobaraensis    TTCTACTCGGCGCTGCGGAACACGCCGTCCTTCAAGGAGC
S.viridis         TTCTACTCGGCGCTGCGGAACACGCCGTCCTTCAAGGAGC
S.paucisporogenes TTTTACTCGGCGCTGAGGAATACGCCGTCCTTCAAGGAAA
S.platensis       TTTTACTCGGCGCTGAGGAATACGCCGTCCTTCAAGGAAA
S.baldaccii       TTCTACTCGGCACTGAGGAACACACCGTCTTTCAAGGAAA

S. mobaraensis    GGAACGGAGGCAATCACGACCCGTCCAGGATGAAGGCCGT
S.viridis         GGAACGGAGGCAATCACGACCCGTCCAGGATGAAGGCCGT
S.paucisporogenes GGGATGGAGGCAACTACGACCCATCCAAGATGAAGGCGGT
S.platensis       GGGATGGAGGCAACTACGACCCATCCAAGATGAAGGCGGT
S.baldaccii       GAGACGGAGGCAACTACGACCCGTCCAAGATGAAGGCGGT

S. mobaraensis    CATCTACTCGAAGCACTTCTGGAGCGGCCAGGACCGGTCG
S.viridis         CATCTACTCGAAGCACTTCTGGAGCGGCCAGGACCGGTCG
S.paucisporogenes GGTCTACTCGAAGCACTTCTGGAGCGGGCAGGACCCGCGG
S.platensis       GGTCTACTCGAAGCACTTCTGGAGCGGGCAGGACCCGCGG
S.baldaccii       GATCTACTCGAAGCACTTCTGGAGCGGGCAGGACCAGCGG

S. mobaraensis    AGTTCGGCCGACAAGAGGAAGTACGGCGACCCGGACGCCT
S.viridis         AGTTCGGCCGACAAGAGGAAGTACGGCGACCCGGACGCCT
S.paucisporogenes GGCTCCTCCGACAAGAGGAAGTACGGCGATCCGGATGCCT
S.platensis       GGCTCCTCCGACAAGAGGAAGTACGGCGATCCGGATGCCT
S.baldaccii       GGCCCTTCTGACAGGAGGAAGTACGGCGATCCGGAAGCCT

S. mobaraensis    TCCGCCCCGCCCCGGGCACCGGCCTGGTCGACATGTCGAG
S.viridis         TCCGCCCCGCCCCGGGCACCGGCCTGGTCGACATGTCGAG
S.paucisporogenes TCCGCCCCGACCAGGGCACCGGCCTGGTAGACATGTCGAG
S.platensis       TCCGCCCCGACCAGGGCACCGGCCTGGTAGACATGTCGAG
S.baldaccii       TCCGCCCCGCCCAGGGTACCGGCCTGGTCGACATGTCGAA
```

```
S. mobaraensis     GGACAGGAACATTCCGCGCAGCCCCACCAGCCCCGGTGAG
S.viridis          GGACAGGAACATTCCGCGCAGCCCCACCAGCCCCGGTGAG
S.paucisporogenes  GGACAGGAACATTCCGCGCAGTCCCGCCAAAACCTGGTGAA
S.platensis        GGACAGGAACATTCCGCGCAGTCCCGCCAAAACCTGGTGAA
S. baldaccii       GGACAGAAGCATCCCGCGCAGTCCCGCCAACGCCGGCGAA

S. mobaraensis     GGATTCGTCAATTTCGA
S.viridis          GGATTCGTCAATTTCGA
S.paucisporogenes  AGTTGGGTAAACTTCGA
S.platensis        AGTTGGGTAAACTTCGA
S.baldaccii        AGCTGGGTCAACTTCGA
```

Appendix VIII: Comparison of Nucleotide Sequence of Streptomyces TGases

Sequences aligned using Clustal W. Dashes (-) incorporated to give best alignment. Identical bases are indicated by an *. Where S_bal, S_cin, S_fra, S_pla and S_mob represent the genetic sequence of TGase from *S. baldaccii*, *S. cinnamoneus*, *S. fradiae*, *S. platensis* and *S. mobaraensis* respectively.

```
S_b  ATGTACAAACGTCGGAGAATTCTCACCTTCGCCACTGCGGGTGCAGTCATATGCACCGCC  60
S_c  ------------------------------------------------------------
S_f  ------------------------------------------------------------
S_p  ------------------------------------------------------------
S_m  ----------------------------------------------------------CG  2


S_b  GGACTCATCCCGTCGGTCAGCCAGGCCGCCGGCAGCGGCGACAGGGGAGAGAAAGGGTCC  120
S_c  ----------CTTCGGTCAGCCAGGCCGCCAGCAGTGGCGATGGGGAAGAGAAGGGGTCC  50
S_f  ----------CGTCGCTCAGCCAGGCCGCCGACAGCGGCGACGGGGCGAAGAAGGGGTCC  50
S_p  ----------CGTCGGTCAGCCATGCCGCCAGCGGCGGCGACGGGGAATGGGAGGGGTCC  50
S_m  TCCGCAGGCGCGGCCACCGGCAGTGGCAGTGGCAGCGGCACCGGGGAAGAGAAGAGGTCC  62
               *    *   * **    *  *     * *  ***    ***    *  *   ***** 


S_b  TACGCCGAAACGCACGGCCTGACGGCGGATGACGTCAAGAACATCAACGCACTCAACGAG  180
S_c  TACGCCGAAACGCACGGCCTGACGGCGGATGACGTCGAGAGCATCAACGCACTGAACGAA  110
S_f  TACGCCGAAACGCACGGCCTGACGGCGCATGACGTCAAGAACATCAACGCGCTCAACGAA  110
S_p  TACGCCGAAACGCACGACCTGACGGCGGAGGACGTCAAGAACATCAACGCGCTGAACAAA  110
S_m  TACGCCGAAACGCACCGCCTGACGGCGGATGACGTCGACGACATCAACGCGCTGAACGAA  122
     ***************    *********  * ******  *    ********  *** *** *


S_b  AGCGCCCTGACTCCGGAGCAACCCGGCAGGTCACCAGGGGAATTGCCTCCGAGTGCCAGC  240
S_c  AGAGCTCTGACTCTGGGCCAACCTGGCAAGCCTCCGAAGGAATTACCTCCGAGCGCCAGC  170
S_f  AGGGCTCCGGCCCTCGGTCAGCCGGGCAAGCCCCCGG-----------CGGGCGC--GC  156
S_p  AGAGCTCTGACTGCGGGTCAACCCGGGAATTCGCCGGCGGAATTGTCGCCGAGCGCCAGT  170
S_m  AGCGCTCCGGCCGCT-----------------------------------TCGAGCGCCGGT  149
     ** ** * *  *                                   ** * **   *


S_b  CCGTCCTCCCGGGCCCCCGGCTCTACCGACGACAGGGAAACCCCTCCGGCCGAGCCGCTC  300
S_c  GCGCCCTCCCGGGCCCCC------TCCGATGACCGGGAAACTCCTCCCGCCGAGCCGCTC  224
S_f  CC-CCCTTCCGAACCCCCGCCCTGGTCGACGACAGGGAAACCCCTCCCGCCGAGCCGCTC  215
S_p  GCGCTCTTCCGGGCCCCCGACGCCGTCGATGACAGGGTGACCCCTCCCGCCGAGCCGCTC  230
S_m  CCGTCCTTCCGTGCTCCTGAT---TCAGATGAGCGGGTGACTCCTCCCGCCGAGCCGCTC  206
       *    ** ***   * **         **  **   ** **  ***** ***********


S_b  GAAAGGATGCCTGATGCGTACCGGGCCTACGGAGGCAGGGCCACTACGGTCATCAACAAC  360
S_c  GACAGGATGCCTGAGGCGTACCGGGCCTACGGAGGCAGGGCCACTACGGTCGTCAACAAC  284
S_f  GACAGGATGCCCGACGCGTACCGGGCCTACGGAGGCAGAGCCACTACGGTCGTCAACAAC  275
S_p  AACCGGATGCCTGACGCGTACCGGGCCTACGGAGGCAGGGCCACTACGGTCGTCAACAAC  290
S_m  GACCGGATGCCCGACCCGTACCGGCCCTCGTACGGCAGGGCCGAGACGATCGTCAACAAC  266
        * *******  **  ******** *** *** *** *** ** ** * * *******


S_b  TACATACGCAAGTGGCAGCAGGTCTACAGTCACCGCGACGGCAAGAAGCAGCAAATGACC  420
S_c  TACATACGCAAGTGGCAGCAGGTCTACAGTCACCGCGACGGAAAGAAACAGCAAATGACC  344
S_f  TACATACGCAAGTGGCAGCAGGTCTACAGTCAGCGCGACGGCAAGAAGCAGCAAATGACC  335
S_p  TACATACGCAAGTGGCAGCAGGTCTACAGTCAACGCGGCGGCAACCCACAGCAAATGACC  350
S_m  TACATACGCAAGTGGCAGCAGGTCTACAGCCACCGCGACGGCAGGAAACAGCAGATGACC  326
     *****************************  ** **** *** *     ****  ******
```

```
S_b GAAGAGCAGCGAGAAAAGCTGTCCTACGGTTGCGTCGGCGTCACCTGGGTCAATTCGGGT 480
S_c GAAGAGCAGCGAGAAAAGCTGTCCTACGGTTGCGTTGGCGTCACCTGGGTCAACTCGGGC 404
S_f GAAGAGCAGCGAGAGAACCTGTCCTACGGTTGCGTCGGCGTCACCTGGATCAATTCAGGC 395
S_p GAAGAGCAGCGAGAACAACTGTCCTACGGCTGCGTCGGCGTCACCTGGGTCAATACAGGC 410
S_m GAGGAACAGCGGGAGTGGCTGTCCTACGGTTGCGTCGGTGTCACCTGGGTCAACTCGGGC 386
    ** ** ***** **   **********  **** ** ******** ****  * **

S_b CCCTACCCGACGAACAAATTGGCGTTTGCGTCCTTCGACGAGAACAAGTACAAGAACGAC 540
S_c CCCTACCCGACGAACAGATTGGCGTTCGCGTCCTTCGACGAGAACAAGTACAAGAACGAC 464
S_f TTCTACCCGACGAACAAATTGGCGTTCGCGTTCTTCGACGAGAACAAGTACAAGAACGAT 455
S_p CCCTACCCGACGAACAAACTCGCGTTCGCGTTCTTCGACGAGAACAAGTACAAGAACGAC 470
S_m CAGTATCCGACGAACAGGCTGGCTTTCGCGTTCTTCGACGAGGACAAGTACAAGAACGAG 446
      ** *********    * ** **  ****  ********* ****************

S_b CTCAAAAACATCAGCCCCCGACCCGGGGAAACGCGGGCAGAGTTCGAGGGCCGCATCGCC 600
S_c CTGAAGAACACCAGCCCCCGACCCGATGAAACGCGGGCGGAGTTCGAGGGTCGCATCGCC 524
S_f CTGAAGAATACCAGCCCCCGACCCAATGAAACGCGGGCGGAGTTCGAGGGCCGCATCGCC 515
S_p CTGGAAAACAGCAGACCGCGCGACCCAACGAGACGCAGGCGGAGTTCGAGGGGCGCATCGCC 530
S_m CTGAAGAACGGCAGGCCCCGGTCCGGCGAAACGCGGGCGGAGTTCGAGGGGCGCGTCGCC 506
    **  * **   *** ** **  **    ** **** ***  *********** *** *****

S_b AAGGACAGTTTCGACGAGGAGAAGGGCTTCAAGCGGGCGCGTGATGTGGCGTCCATCATG 660
S_c AAGGGCAGTTTCGACGAGGGGAAGGGTTTCAAGCGGGCGCGTGATGTGGCGTCCGTCATG 584
S_f AAGGCCAGCTTCGACGAGGGGAAGGGCTTCAAGCGGGCGCGTGATGTGGCCTCCATCATG 575
S_p AAGGACAGTTTCGATGAGGGAAAGGGTTTCAAGCGGGCGCGTGAGGTGGCATCCGTCATG 590
S_m AAGGACAGCTTCGACGAGGCGAAGGGGTTCCAGCGGGCGCGTGACGTGGCGTCCGTCATG 566
    ****  *** ***** ****    *****  *** ************* ***** *** *****

S_b AACAAGGCCCTGGAAAGCGCCCACGACGAGGGGGCGTACATCGACAACCTCAAGACGGGG 720
S_c AACAAGGCCCTGGAAAATGCCCACGACGAGGGGACTTACATCAACAACCTCAAGACGGAG 644
S_f AACAAGGCCCTGGAGAACGCCCCCGACGAGGGAACGTACCTCAAGAACCTCAAGACGGAC 635
S_p AACAAGGCCCTGGATAACGCGCACGACGAGGAGACTTACATCGGCCACCTCAAGACAGAG 650
S_m AACAAGGCCCTGGAGAACGCCCACGACGAGGGGGCGTACCTCGACAACCTCAAGAAGGAG 626
    **************  *   ** * ********    * *** **         ********   *

S_b CTCACGAACAACAATGACGCTCTGCTCTACGAGGACAGCCGATCGAACTTCTACTCGGCA 780
S_c CTCACGAACAACAATGACGCTCTGCTCCGCGAGGACAGCCGCTCGAACTTCTACTCGGCG 704
S_f CTCACGAACAAAAATGATGCCCTGCTCCACGAGGACAGCCGCTCGAACTTCTACTCGGCA 695
S_p CTCGCGAACAAAAACGACGCTCTGCTCTACGAGGACAGCCGCTCGAGCTTTTACTCGGCG 710
S_m CTGGCGAACGGCAACGACGCCCTGCGGAACGAGGATGCCCGCTCGCCCTTCTACTCGGCG 686
    **    *****    ** ** ** **  ****    ****** *** ***   *** *******

S_b CTGAGGAACACACCGTCTTTCAAGGAAAGAGACGGAGGCAACTACGACCCGTCCAAGATG 840
S_c CTGAGGAACACACCGTCCTTCAAGGAAAGGGACGGCGGCAACTACGACCCGTCCAAGATG 764
S_f CTGAGGAATACACCGTCCTTCAGGGAAAGAGACGGAGGCAACTACGACCCGTCCAAGATG 755
S_p CTGAGGAATACGCCGTCCTTCAAGGAAAGGGATGGAGGCAACTACGACCCGTCCAAGATG 770
S_m CTGCGGAACACGCCGTCCTTCAAGGACCGCAACGGCGGCAATCACGACCCGTCCAAGATG 746
    ***  **** ** ***** **** ***   *   * ** *****   ******************

S_b AAGGCGGTGATCTACTCGAAGCACTTCTGGAGCGGGCAGGACCAGCGGGGCCCTTCTGAC 900
S_c AAGGCGGTGATCTACTCGAAGCACTTCTGGAGCGGGCAGGACCAGCGGGGCTCCTCCGAC 824
S_f AAGGCGGTGATCTACTCCAAGCACTTCTGGAGCGGGCAGGACCAGCGGGGCTCCGCCGAC 815
S_p AAGGCGGTGGTCTACTCGAAGCACTTCTGGAGCGGGCAGGACCAGCGGGGCTCCTCCGAG 830
S_m AAGGCCGTCATCTACTCGAAGCACTTCTGGAGCGGCCAGGACCGGTCGGGCTCCTCCGAC 806
    *****  **  ******* ******************* ******  * ****  *  * **

S_b AGGAGGAAGTACGGCGATCCGGAAGCCTTCCGCCCCGCCCAGGGTACCGGCCTGGTCGAC 960
S_c AAGAGGAAGTACGGCGACCCGGAAGCCTTCCGCCCCGACCAGGGTACCGGCCTGGTCGAC 884
S_f AAGAGAAAGTACGGCGACGCGGAAGCTTTCCGCCCCGACCAGGGCACCGGCCTGGTCGAC 875
S_p AAGAGGAAGTACGGTGACCCGGACGCCTTCCGCCCCGGCCAGGGCACAGGTCTGGTAGAC 890
S_m AAGAGGAAGTACGGCGACCCGGAGGCCTTCCGCCCCGACCGCGGCACCGGCCTGGTCGAC 866
    * *** ********* **  **** ** ** ********** **  ** ** ** ***** ***
```

```
S_b ATGTCGAAGGACAGAAGCATCCCGCGCAGTCCCGCCAACGCCGGCGAAAGCTGGGTCAAT1020
S_c ATGTCGAAGGACAGAAGCATTCCGCGCAGTCCGGCCAAGCCCGGCGAAGGTTGGGTCAAT 944
S_f ATGTCGAAGGACAGAAACATTCCGCGCAGTCCGGCCCGTCCTGGCGAAGGTTGGGTCAAT 935
S_p ATGTCGAGGGACAGGAACATTCCGCGTAGTCCCGCAAAACCTGGCGAAAGTTGGGTCAAT 950
S_m ATGTCGAGGGACAGGAACATTCCGCGCAGCCCCACCAGCCCCGGCGAGAGTTTCGTCAAT 926
    ******* ****** * *** ***** ** ** *    * *****  * *  *****
```

```
S_b TTCGACTACGGCTGGTTCGGGGGCTCAAACGGAAGCGGATGCCGACAAGACGGTATGGACC1080
S_c TTCGACTACGGTTGGTTCGGGGGCTCAAACAGAAGCGGATGCCGACAAAACCACATGGACC1004
S_f TTCGACTACGGGTGGTTCGGGGGCTCAAACGGCAGCGGACGCCGACGAAACAACATGGACC 995
S_p TTCGACTACGGCTGGTTCGGGGGCTCAGGCAGAAGCGGATGCCGACAAAACCGTATGGACC1010
S_m TTCGACTACGGCTGGTTCGGGAGCGCAGACGGAAGCGGACGCCGACAAGACCGTATGGACC 986
    ********** ******* ** **  * * ****** ****** * **   ******
```

```
S_b CACGGCGACCACTACCACGCGCCCAATAGTGGCCTGGGCCCCATGCACGTATACGAGAGC1140
S_c CACGGCGACCACTACCACGCGCCCAATAGCGACCTGGGCCCCATGCACGTACACGAGAGC1064
S_f CACGGCGACCACTATCACGCACCCAATAGCGGCCTGGGCCCCATGCATGTCCACGAGAGC1055
S_p CACGCCAACCACTATCATGCGCCCAATGGCGGCATGGGCCCCATGAACGTATACGAGAGC1070
S_m CACGGCAACCACTACCACGCGCCCAATGGCAGCCTGGGTGCCATGCACGTGTACGAGAGC1046
    ****  * ******* ** ** ****** *   * ****  ***** * **  *******
```

```
S_b AAGTTCCGGAACTGGTCTGCCGGGTACGCGGACTTCGACCGCGGAACCTACGTGATCACG1200
S_c AAGTTCCGGAAGTGGTCTGCCGGGTACGCGGACTTCGACCGCGGAGCCTACGTGATCACG1124
S_f AAGTTCCGGAAGTGGTCCGCCGGGTACGCGGACTTCGACCGCGGAACCTACGTGATCACG1115
S_p AAGTTCCGGAACTGGTCTGCGGGGTACGCGGACTTCGACCGCGGAGCCTACGTCATCACG1130
S_m AAGTTCCGCAACTGGTCCGACGGTTACTCGGACTTCGACCGCGGAGCCTACGTGGTCACG1106
    ******** ** ***** *  ** *** ***************** ******  *****
```

```
S_b TTCATACCCAAGAGCTGGAACACCGCCCCCGCCAAGGTGCAGCAGGGCTGGCCGTAA 1257
S_c TTCATACCCAAGAGCTGGAACACCGCCCCCGCCAAGGTGGAGCAAGGCTGGCCGTGA 1181
S_f TTTATACCCAAGAGCTGGAACACCGCCCCCGACAAGGTGGAGCAAGGCTGGCCGTAG 1172
S_p TTCATACCCAAGAGCTGGAACACCGCCCCCGCCGAGGTGAAGCAGGGCTGGCCGTAA 1187
S_m TTCGTCCCCAAGAGCTGGAACACCGCCCCCGACAAGGTCAAACAAGGTTGGCCATAA 1163
    **  *  ********************** * ****   * ** ** ***** *
```