UPTEC X 05 040
SEP 2005
ISSN 1401-2138

ERIC WIKTELIUS

Cloning, characterisation and folding of glutathione transferase from thermophilic microorganisms

Master’s degree project

UPPSALA UNIVERSITET

Molecular Biotechnology Programme

Uppsala University School of Engineering

UPTEC X 05 040	**Date of issue 2005-09**

Author

Eric Wiktelius

Title (English)

Cloning, characterisation and folding of glutathione transferase from thermophilic microorganisms

Title (Swedish)

Abstract

Glutathione transferases (GSTs) belong to an ancient enzyme family. Among their functions is an important role in the detoxification process, by conjugating glutathione to the toxin and thereby facilitating transport out of the cell. This study focuses on prokaryotic GSTs. Particularly interesting for folding and stability is the possible thermal stability that may be expected from a GST from a thermophilic organism. GSTs from thermophilic cyanobacteria *Thermosynechococcus elongatus* and its mesophilic relative *Synechococcus elongatus* were cloned, expressed and purified. Both proteins had activity with substrates used to study activity in mammalian GSTs, and a higher thermal stability could be observed with the protein from *T. elongatus* than with *S. elongatus*. Stability and refolding were studied through denaturation and renaturing. Even though the experiments were partially successful, method optimisation will be necessary in order to get clear results.

Keywords

Glutathione transferase, GST, thermophilic, folding, stability

Supervisor

Gun Stenberg

Department of Biochemistry, Uppsala University

Scientific reviewer

Per Jemth

Department of Medical Biochemistry and Microbiology, Uppsala University

Project name	Sponsors
Language **English**	Security
ISSN 1401-2138	Classification
Supplementary bibliographical information	Pages **28**

Biology Education Centre Biomedical Center Husargatan 3 Uppsala
Box 592 S-75124 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 555217

Cloning, characterisation and folding of glutathione transferase from thermophilic microorganisms

Eric Wiktelius

Sammanfattning

Glutationtransferaser (GST) är mycket viktiga enzymer som kan katalysera flera olika reaktioner i cellen. Speciellt viktiga är de för avgiftningssystemet. Vissa potentiellt farliga ämnen kan vara svåra att transportera ut från cellen p.g.a. sina kemiska egenskaper, och genom att GST konjugerar tripeptiden glutation till ämnet blir det lättare. Eftersom en stor del av avgiftningen sker i levern, är koncentrationen av glutation extra hög där. Det finns naturligtvis många naturligt förekommande substrat för GST, men för att studera aktiviteter hos GST brukar man använda syntetiska kemikalier, utan fysiologisk betydelse.

De flesta studierna gjorda med GST har hittills utförts med GST från däggdjur. I den här studien studeras GST från prokaryoter. Av speciellt intresse i denna studie är veckning och stabilitet av dessa proteiner. Veckning av ett protein innebär den slutgiltiga struktur som proteinet får p.g.a. sitt aminosyrainnehåll. För att kunna vara katalytiskt aktivt och stabilt måste en stabil veckningsstruktur kunna uppnås. I den här studien undersöks proteiner från organismer som lever i temperaturer mellan 45 och 65°C. Målet är att rena fram proteinerna, undersöka deras katalytiska förmågor och se om de varmlevande organismerna har GST med annorlunda strukturell stabilitet än de från icke-termofila.

GST från två arter cyanobakterier, en termofil och en icke-termofil, klonades och proteinerna uttrycktes i *E. coli*. Resultat från analyser av dessa proteiner visade att de prokaryota proteinerna hade aktivitet med en rad olika substrat och att termostabilitet kunde påvisas hos det ena proteinet. Veckningens stabilitet studerades genom att denaturera proteinet (veckla ut det) och sedan låta det vecka ihop sig igen. Dessa försök gav inte riktigt tydliga resultat, men var ändå lovande när det gäller möjligheten att sätta upp framtida försök för veckningsstudier av de nya proteinerna.

Examensarbete 20 p, Molekylär bioteknikprogrammet

Uppsala universitet september 2005

Table of contents

Abbreviations

A_{wl} – Absorbance at wavelength wl nm
Amp – Ampicillin
Bp – Base pair(s)
CDNB – 1-chloro-2,4-dinitrobenzene
CuOOH – Cumene hydroperoxide
Da – Dalton
E. coli – *Escherichia coli*
EDTA – Ethylenediaminetetraacetic acid
EPNP – 1,2-epoxy-3-(4-nitrophenoxy)-propane
E-flask – Erlenmeyer flask
EA – Ethacrynic acid
GSH – Glutathione
GST – Glutathione transferase
Gdn-HCl – Guanidium-HCl
IPTG – Isopropyl-1-thio-ß-D-galactopyranoside
LB – Luria-Bertani
M_w – Molecular weight
NADPH – Nicotinamide adenine dinucleotide phosphate
Nt – Nucleotide(s)
NBC – 4-nitrobenzylechloride
O/N – Over night
PCR – Polymerase chain reaction
RT – Room temperature
S. elongatus – *Synechococcus elongatus*
SDS-PAGE – Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
synGST – Glutathione transferase cloned from *Synechococcus elongatus*
T. elongatus – *Thermosynechococcus elongatus*
terGST – Glutathione transferase cloned from *Thermosynechococcus elongatus*
Tris-HCl – Trizma base (Sigma) with pH set using HCl.

1 Introduction

1.1 The role of GST

The ancient protein family of GSTs is a part of the detoxification system in the cell. There are multiple classes of GSTs with varying targets of activity. Cytosolic GSTs of many classes have been characterised to a large extent in mammals[1]. GSTs often play, due to their large variety of function, a role in disease and drug resistance development[2]. Common for the proteins is the use of GSH for catalysis of the reactions. GSH is the most abundant cysteine peptide in the cell, if the proteins are excluded[3], with concentrations ranging between 0.1 and as much as 5 mM. GSH is a tripeptide, consisting of a glutamic acid, a cysteine and a glycine. The molecule is unusual in the way, that it does not have the classic peptide bond between the α-carbons in cysteine and glutamic acid. Instead, the nitrogen in cysteine is bound to the γ-carbon of the glutamic acid (Figure 1).

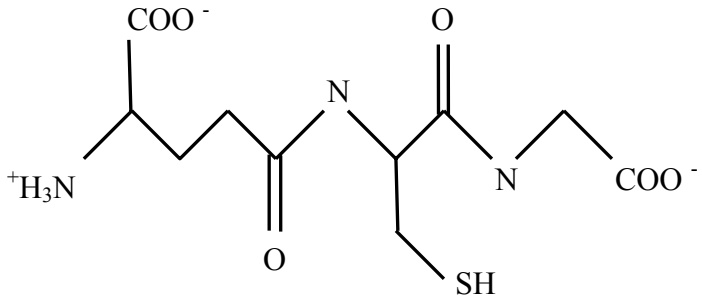


Figure 1. The GSH molecule. The sulphuric cysteine is the catalytic centre for the electrophilic substrates.

The most widely studied function is the conjugation of GSH to electrophilic substrates. Other examples of functions are steroid isomerisation and peroxidase activity[1]. Many organic and especially aromatic compounds are potentially toxic to the cell and can diffuse freely across the cellular membrane because of their hydrophobic nature. Due to their hydrophobicity, excretion of the compound can be troublesome. To enable excretion of the compound, GST can bind GSH to the aromatic compound through the sulphur in cysteine. For many molecules, for example an unsubstituted aromatic system, there are usually more enzymes involved in the complete handling of the toxic compound[1]. Such molecules have increased in number and variants over the years due to man-made chemicals coming from the industrial world. The enzymes are much older than human beings and have been shown to be able to detoxify toxic substances in nature, such as H_2O_2 and toxins from fungi and plants[1]. The general reaction catalysed by GST is[4]:

$$GSH + A\text{-}B \rightarrow GSA + HB$$

A is the main target molecule with a group B bound to it, for example a chloride, as is the situation with CDNB (figure 2).

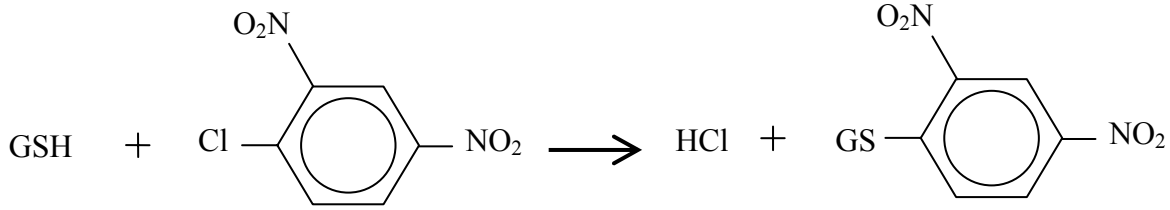


Figure 2. The conjugation of GSH with the synthetic substrate electrophile CDNB. Substitution of the chloride creates a charged group on the aromatic and a hypothetical secretion of the substrate out of the cell is simplified.

1.2 GST classification and characterisation

GSTs are divided into evolutionary classes. They have been classified according to their sequence, substrate specificity and antibody affinity[5]. Common for all known cytosolic GSTs is that they are dimers[6]. Both homo- and heterological protein variants exist. The subunits have not been shown to have any covalent contact. For human GST P1-1 (pi), the subunits are instead locked by a hydrophobic interaction between a residue in one subunit, and a pocket in the other[10]. Logically, between classes, the sequence varies largely at the electrophilic substrate binding part, whereas the GSH binding part is more conserved[5,11]. The catalytic site of GST is for most classes an N-terminal tyrosine, or in some cases a serine or a cysteine residue, which helps to stabilise the GS^- ion by interacting with the hydroxyl group in the tyrosine/serine case and forming a disulfide in the cysteine case[1]. When aligning GSTs of all organisms there are very few conserved amino acid residues to be found[7]. Only about 5 % can be expected to be conserved between different species of even closely related organisms. Of those repeatedly come across are a tyrosine in the N-terminal, a proline farther in, a glycine in the C-terminal part and an aspartic acid just 7 residues after the glycine (positions at 4, 54, 146 and 153 respectively in human GSTP1-1)[7]. GSTP1-1 has been studied for some of these residues' structural importance[5,10,11].

1.3 Prokaryotic GST

Although is has been known for quite some time that bacterial GSTs exist, they have been studied to a limited extent. There are only a few bacteria known to contain glutathione, but there are still many that remain to be tested. Early classifications placed all bacterial GSTs in the so called theta class which holds representatives from many higher organisms[6]. This led scientists to believe that this class of GSTs were the ancestors of all GSTs. Later classifications have placed bacterial GSTs in a separate group called beta[2]. These are the most widespread GSTs with examples in higher organisms. What is typical for the theta enzymes is that they have poor activity with xenobiotic CDNB, which is the most widely used substrate for characterising the activity of mammalian GSTs[2], and present low affinity for GSH matrices[6]. If and how there is such a high versatility in activity with prokaryotic GSTs as with for example mammalian GSTs have been far less characterised. Since studies of prokaryotic GSTs have attracted far less interest in the scientific world than eukaryotic have, many of the fundamental ideas about GSTs are presently unknown. Are prokaryotic GSTs dimers? What reactions do they catalyse? Are there as few conserved residues in these enzymes as there are between classes and particularly between organisms? One of these questions was actually partially answered in the very beginning of this project, when scanning through genomic databases for a prokaryotic GST to study. Cyanobacterium *T. elongatus* had examples of conserved areas in its annotated GST gene. One reason why bacterial GSTs are so poorly analysed, is that due to the highly varying sequences between classes of GST, it can be troublesome to find relevant open reading frames in newly sequenced genomes.

1.4 Cyanobacteria

The prokaryotic organism kingdom hosts a large amount of respirating bacteria, believed to be a predecessor of the eukaryotes and ultimately animals. But nonetheless the kingdom includes some 1500 species of bacteria[8], from which plants have evolved. These bacteria are summarised under the name cyanobacteria, being prokaryotes capable of photosynthesis. Cyanobacteria habit both fresh- and seawater ecosystems and are often mistaken for algae, because of their colour and habitat. As with other prokaryotes, there are cyanobacteria that

habit a warmer temperature than the average earthly organism above ground. *T. elongatus* flourishes in Japanese hot springs at 45-65°C. Its mesophilic relative *S. elongatus* has a more humane habitat. Cyanobacteria have a remarkably high concentration of GSH in their cytosol[6]. Both these species have been shown to have genes annotated as GSTs[9]. The cyanobacterial GST genes studied in this work had DNA sequences not too different from human GST P1-1 and even more in line with the sequence of human GST T1-1 (theta)[9].

1.5 Folding and stability of GST

Correct folding of a protein is necessary for a correct function. If a structural element of the protein is missing, an adequate folding of the protein is impossible. Whilst GSTs present only a few conserved amino acids in their primary structure, there is a similarity between the classes in terms of general folding of the polypeptide chain[5,10]. Folding studies have traditionally focused on investigating the role of a suspected folding determinant in a polypeptide[11]. By site-directed mutagenesis it is possible to change an amino acid of choice and see what the effect is on the folding of the protein. Recent folding experiments with GSTs, have in most cases handled human GST P1-1[11]. The way of studying folding has been through denaturation. The denaturing environment is changed in the direction of lower or higher denaturing potential to investigate the folding progression. The change in protein structure is in this study monitored by emitted fluorescence from an excited protein sample. Both the process of denaturation (unfolding) and renaturation by dilution (folding) can be studied this way.

Stability can be defined as how soluble the protein is, how a potential inactivation of the protein affects its activity and how resistant the protein is to a denaturant. It is also possible to calculate free energies for different unfolding states and the basic free energy of the fully native protein. Whilst calculating the free energy by analysing the bonds in detail would be far too labour-intensive, it is easier to just compare the native and denatured state. To make such a calculation one needs to know the relation between a denatured and a native state in a denaturing environment. Due to the changes in emission spectra between conformations, a protein can be studied at multiple concentrations of denaturant, and the conformation distribution is recorded. These values are compared with fully denatured and folded protein[12]. The unfolding equilibrium constant, K_u, for the reaction N(ative) ↔ D(enatured) is [D] / [N]. Using fluorescence values it is defined as:

$$K_u = (F_N - F) / (F - F_U)$$

F_N is the fluorescence signal for a native protein, F_U for a fully unfolded protein and F is the signal at the studied concentration of denaturant. The free energy, consisting of the total enthalpy gain of the molecule subtracted by the level of entropy embedded in the molecule, is a relation that can be calculated only with the aid of K_u and temperature by use of the famous Gibb's free energy formula:

$$\Delta G_u = -RT \ln K_u$$

By extrapolating the plot of free energies against concentrations, one can calculate the free energy change for the whole unfolding/folding procedure by subtracting the native protein's free energy with the unfolded's. A high difference in free energy, gives a higher folding stability of the protein.

By further studying spectral changes during the folding or unfolding process a stopped-flow apparatus is used. By quickly diluting either denatured protein with buffer for a refolding, or diluting a protein sample with denaturant for unfolding, the change in emission can be monitored in real-time. This provides information about how fast the conformational changes happen. The change is attempted to be adapted to a mathematical description. This mathematical description of the folding/unfolding process can as of now not be used to give ideas about the structure itself, only how fast an equilibrium is reached.

1.6 Why study GSTs?

The detoxification system has existed in the cell for 2.5 billion years[1]. Since GSTs are common in all types of organisms, even bacteria, their classification is a tool for documenting evolutionary events. GSTs are believed to be the result of an adaptation to the oxidative stress of the oxygenic world. They do have similarities with other stress related proteins, even with those that belong to organisms that do not have GSTs in their genome[1]. Protein structures can be analysed to a very high resolution by crystallisation, but what defines the structure remains largely unknown. GSTs from different organisms have a very similar structure, but the amino acid content vary extensively. GSTs are very suitable as model proteins, since they often can be easily purified by affinity chromatography with immobilised GSH[2]. The purification is very specific, but the affinity for GSH in the different GST classes is very varying. Assaying for activity can be varied substantially by use of a large number of substrates. Often synthetic compounds without physiological relevance are used as substrates for GST[6].

Human GSTs are, as expected, in their active form at 37°C. As a novel feature in the folding and stability experiments of GST, this study aims to show what characteristics a GST, from an organism living at a substantially warmer temperature than most mammals, has.

1.7 Cloning of recombinant DNA to expression systems

Why is it so common to use cloning to express a protein? To be able to purify a protein for characterisation studies, the purity does not only have to be high, so does the protein amount. The level of expression of a protein targeted for purification is usually much too low in the original cell. Cultures of cells would be in need of extreme volumes to achieve enough protein for more than a detection level. If the protein does not need post-translational modifications that cannot be handled by prokaryotes, expression in bacteria is fast and gives high yields. The most common variant is a genetically engineered and well characterised form of *E. coli*. The bacteria will not distinguish foreign proteins from its own, and as long as the environment inside the bacterium does not damage the protein, it will remain in the bacterium until the cell is lysed. The general procedure requires the gene encoding the protein to be cloned into a plasmid to be transformed into the bacteria, and the expression to be induced. An amplified fragment of DNA through PCR can be cloned into the pKK –D vector[13]. This is a circular vector with a promoter similar to the one in the *lac* operon upstream from where the gene to be expressed is ligated. Unlike lactose, being a part of the metabolic pathway in the bacterium, IPTG is a patented, very potent and non-degradable inducer that will induce transcription continuously. The vector also mediates ampicillin resistance, making it possible to select for bacteria having picked up the plasmid during the transformation. For a schematic layout of pKK –D, see Figure 3.

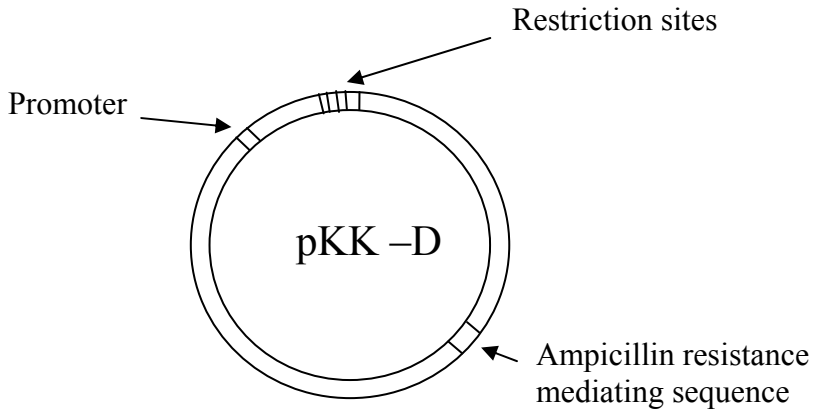


Figure 3. The pKK –D vector. The promoter sequence is the target site for IPTG, resulting in a continuous expression. There are multiple restriction sites upstream from the promoter, so that the insertion can be equipped with fitting ends to the vector. Restriction cleavage of the vector and insertion produces sticky ends, complementary to one another. These can be ligated to produce the final circular construct. In case there should be a sequence specific to a restriction enzyme in the insertion, many other restriction enzymes can be chosen from.

2 Materials and methods

For all large scale temperature controlled centrifugations a Beckman/Coulter model Avanti J-20XP was used. For all small scale centrifugations (microcentrifuge tubes) at RT a Beckman/Coulter model Microfuge 18 was used and at 4-8°C an Eppendorf model 5415 D was used. All PCR reactions were performed in a CR/960 G Gradient Thermal Cycler. For culture incubations a Gallenkamp Orbital Incubator was used for liquid cultures of *E. coli* and a Thermaks incubator for plates. For *T. elongatus* cultures a New Brunswick Scientific/G24 was used. Spectrophotometric measurements for activity and concentration were performed on a Bio-Tek unit model UVIKON XL in MERCK quartz cuvettes. The fluorometer unit was from Aminco model SPF-500. Stopped-flow apparatus was from Applied Photo Physics. Electrophoresis unit for SDS-PAGE was from LKB model 2050 Midget. Electrophoresis unit for agarose gels were manufactured in-house at Uppsala Biomedical Centre, Uppsala, Sweden. Chemicals used in these experiments were from Sigma Aldrich or MERCK except acrylamide which was from Bio-Rad. Brands of other materials are mentioned in the text.

2.1 Genome screening and gene isolation

Database searches. The protein BLAST[9] function was used to find GST in microorganisms of thermophilic nature with the genomic sequence of GSTP1-1 from *Homo sapiens* as query string. Results led to a pursuit of an annotated GST in the *T. elongatus* genomic sequence[8] and as comparison, the sequence of its fully sequenced mesophilic relative *S. elongatus PCC 6301*[9,14] was screened for annotated GST using the sequence from *T. elongatus* as query.

Growing culture for PCR. Slant growing *S. elongatus* (Pasteur Institute) culture were inoculated into 20 ml BG-11 medium (see Appendix for contents) and grown for five days without agitation in an E-flask at room temperature under a conventional 75 W light bulb until a dense green culture could be observed. *T. elongatus* (kindly supplied by Dr. Miwa Sugiura) cells were grown in the same fashion in 30 ml DTN medium (see Appendix for contents) at 55°C for 10 days with slow agitation.

PCR. To amplify the specific gene, PCR was used with primers constituting *Eco*RI and *Hind*III sites with an AA overhang to easier facilitate restriction cleavage. The primers used for the PCR had the following sequence:

S. elongatus 1: 5'-AAGAATTCATGCTGAAGCTGTACGGCAGCGCT-3'
S. elongatus 2: 5'-AAAAGCTTCTAGCGCCGACCAATCGACTGTTG-3'
T. elongatus 1: 5'-AAGAATTCATGCTCAAGCTCTATGGCGGCGCC-3'
T. elongatus 2: 5'-AAAAGCTTTCATGCCCTCGCTCCCATTAAACC-3'

The reactions were for *S. elongatus* 4 µl of culture as template, 0.2 mM dNTP (Amersham Biosciences), 0.8 µM of each primer (Thermo), 2.5 units of *Pfu* polymerase (Stratagene), buffer for the enzyme (Stratagene) and water to a volume of 50 µl. The PCR was started with a hot start at 95°C for 10 min followed by 35 cycles of 1 min 95°C denaturation, 1 min 58-62°C annealing and 1 min elongation at 72°C. The reaction was let to finalize at 72°C for 30 min after completion of the cycles. For *T. elongatus* the reaction was with 1 µl of culture, 10-fold diluted and undiluted, 0.2 mM dNTP, 0.8 µM of each primer, 2.5 units of *Pfu* polymerase, buffer for the enzyme and water to a volume of 50 µl. This had the same hot start as for *S. elongatus* and 35 cycles of 1.5 min 95°C denaturation, 1 min 67, 68 and 69°C annealing and 2 min elongation at 72°C. The PCR products were analysed on a 1 % agarose gel.

2.2 Preparation of competent cells

Heat-shock competent cells of *E. coli* strain XL-1 Blue (Stratagene) were prepared[15] by streaking cells from -80°C onto an LB (see Appendix for contents) + 10 µg/ml tetracycline agar plate and grown O/N at 37°C. A single colony was inoculated into 2 ml of LB medium + 10 µg/ml tetracycline in a 14 ml polystyrene tube with aerating cap and growing O/N with agitation at 37°C. The culture was transferred to an E-flask containing 200 ml LB + 10 µg/ml tetracycline and incubated with agitation at 37°C until the culture reached an OD of 0.46 at 600 nm. The cells were harvested and centrifuged at 7000 x g at 4°C for 5 min. The media was discarded and the cells were resuspended in 40 ml CM1 (10 mM NaOAc pH 5.6, 50 mM $MnCl_2$ and 5 mM NaCl). The cells were put on ice for 20 min, centrifuged as previously and resuspended in CM2 (10 mM NaOAc pH 5.6, 5 % (v/v) glycerol, 70 mM $CaCl_2$ and 5 mM $MnCl_2$). This suspension was placed in ~100 µl aliquots in sterile microcentrifuge tubes and stored at -80°C until used.

2.3 Cloning and transforming

Plasmid preparations. Competent cells *E. coli* strain XL-1 Blue were transformed[15] with commercially available pGEM–3Z (Promega) and constitutive expression vector pKK –D. 0.5 µl plasmid solution was added to an aliquot of cells and vortexed very gently and briefly. The mixture incubated for 1 h on ice. The reaction tubes were heat-shocked at 42°C for 2 min and moved back to ice for 1 min. 500 µl of LB was added to each tube and the cells were placed at 37°C for 1 h. The cells were dispersed on 2 agar plates (LB + 100 µg/ml amp + 15 g/l solidifying agent) per transformation and grown at 37°C O/N. Colonies were transferred to 2 ml LB + 100 µg/ml amp and grown O/N with agitation at 37°C. The plasmids were extracted from the culture using QIAprep Spin Miniprep Kit (QIAGEN). The plasmids were eluted in 50 µl water per culture.

Subcloning. For restriction digestion 10 units of *Hind*III (Fermentas), 10 units of *Eco*RI (Roche) and 3 µl of SuRE/Cut restriction buffer B (Roche) was added to 20 µl of the plasmid preparations and the volume was adjusted to 30 µl before incubation in 37°C for 3 h. The PCR products from section **2.1** were cut with a restriction enzyme as described for the plasmids. The restricted DNA was analysed on a 1% agarose gels. The bands were excised and the DNA was extracted using QIAquick Gel Extraction Kit (QIAGEN) and eluted into the minimal volume possible, 30 µl. The reaction consisted of 28 µl of each fragment preparation, 2 µl of extracted restriction digested pGEM–3Z preparation, 2 units of T4-DNA ligase (Roche), 4 µl of T4-DNA ligase buffer (Roche), 1 mM ATP (Sigma) and 2 µl water in a total volume of 40 µl. The ligation was let to incubate in RT for 24 h. The ligation reaction was then transformed into competent *E. coli*. 1-7 µl of the ligation was added to aliquots of competent cells and the transformation was carried out as mentioned before in this section.

Cloning to expression vector. To prepare pGEM–3Z containing the inserted fragment, the cultures were treated using QIAprep Spin Miniprep Kit as before. Of the eluted plasmid solution, 17 µl was incubated with 2 µl SuRE/Cut buffer B and 1 unit of *Hind*III and *Eco*RI each for 4 h at 37°C. The whole reaction was separated on a 1 % agarose gel. The bands containing the GST-fragments were excised from the gel and the DNA was extracted using QIAquick Gel Extraction Kit as before. To ligate the extracted fragments into pKK –D, the same ligation reaction as with pGEM–3Z was performed, with the exception that 4 µl pKK –D was used instead of 2 and the water was excluded. The ligation was transformed into *E. coli* using the same method as above. For a schematic drawing of the cloning procedure, see Figure 4. Plasmids were prepared from a 2 ml subculture as before, and the vector with insertion was analysed with Big-Dye Sequencing Kit (Applied Biosystems) (ABI Prism 310 protocol), both with primers used for PCR and complementary to the pKK –D vector.

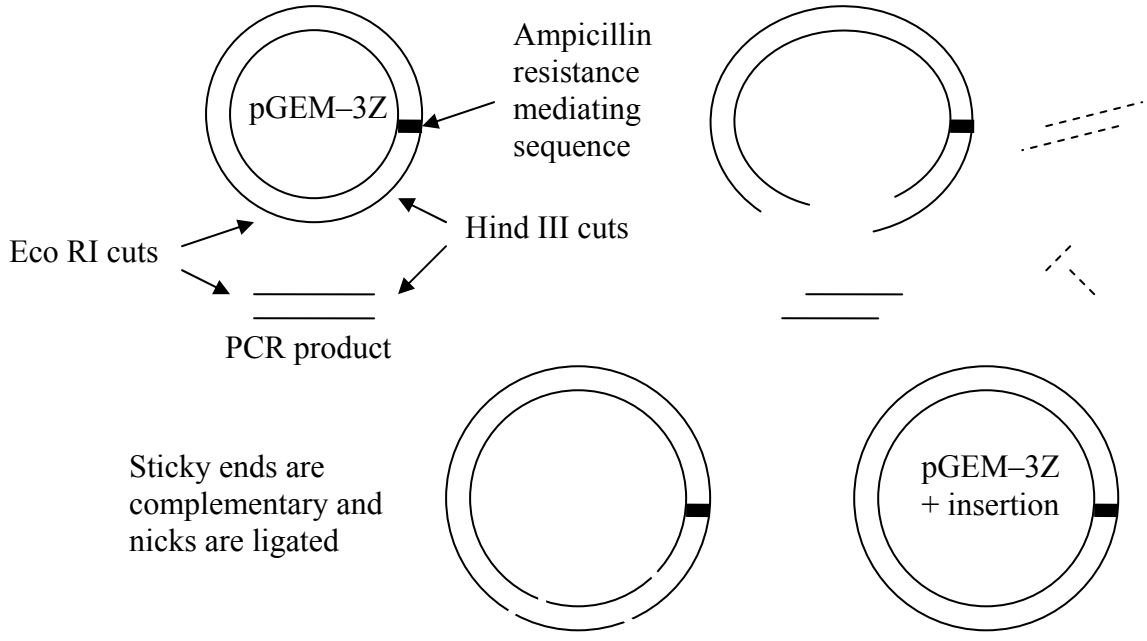


Figure 4. The subcloning into pGEM–3Z. After ligation the construct could be transformed and amplified in *E. coli*. These new preparations were cleaved and separated and ligated into restriction cleaved pKK –D. This last cloning step has been omitted but is performed in the same way as with pGEM–3Z.

2.4 Protein expression

Small scale expressions. For small scale expression tests, colonies of successfully transformed cells were inoculated into 5 ml LB + 100 µg/ml amp in 50 ml polypropylene tubes with conical bottom and grown for 3 h at 37°C with agitation when clearly visible growth could be observed. IPTG was added to a concentration of 1 mM to induce expression and the cultures were grown O/N. The cultures were transferred to microcentrifuge tubes and centrifuged for 2 min at 10000 x g. The supernatant was discarded and the cells were resuspended in 300 µl A-buffer (see Appendix for contents). The cells were disrupted by sonication (30 s with a Vibra Cell/CV 1F setting 8) and centrifuged at 10000 x g for 2 min. 50 µl of the supernatant was analysed on a small SDS-PAGE gel.

Large scale expressions. From clones where the expression was shown to be satisfactory in small scale, a small amount was used to inoculate 2 ml fresh LB + 100 µg/ml amp and grown O/N as before. 350 µl was transferred to 35 ml LB + 100 µg/ml amp in E-flasks. The culture was grown O/N with agitation at 37°C. For each of the two proteins the expression culture was grown in 6 x 0.5 l LB + 100 mg/l amp for terGST and in 6 x 0.5 l 2xYT (see Appendix for contents) + 100 mg/l amp for synGST. The containers were 2 l E-flasks with cotton stoppers. 5 ml from the O/N culture was inoculated into each E-flask and grown at 37°C with agitation. At an A_{600} of 0.45 the terGST expressing culture was induced with 1 mM IPTG. The synGST expressing culture was induced with 1 mM IPTG at an A_{600} of 0.54, and the culturing temperature was lowered to 30°C. Cultures were let to grow and express O/N.

2.5 Protein purification

Lysing cells. Cultures were centrifuged 15 min at 12000 x g. The media was discarded and the cells were resuspended in 70 ml A-buffer. Before sonication of the cells, three spatula tips of lysozyme was added to the cells and let to incubate on ice for 2 h. The cells were disrupted with a Vibra Cell/CV 1F sonicator 20 s with 1 min cooling pause on ice. The process was repeated 5 times. The lysed cells were centrifuged for 45 min at 27000 x g. synGST expressing cell lysate needed an additional hour of centrifugation.

Affinity gel synthesis. Epoxy-activated Sepharose® (Amersham Biosciences) was suspended and washed with 3 l water. The gel was transferred to an E-flask. 1.5 g of GSH was dissolved in 70 ml water at pH 13. The solution was mixed with the gel and the flask was slowly agitated at 30°C for 20 h. The gel was washed with 1.5 l water. The gel was transferred back to an E-flask and 70 ml 1 M ethanolamine pH 9 was added. The gel was agitated at 30°C for 4 h. The gel was ready to be used for affinity chromatography.

Binding, elution and dialysis. The gel was equilibrated with 1 l of A-buffer. The supernatant was mixed with the gel in an E-flask and the GST was bound to the gel at 4°C with slow agitation O/N. The gel was washed with 1 l of A-buffer and packed in a column and mounted to a GradiFrac® system (Pharmacia Biotech) with a pump, a spectrophotometer measuring A_{280}, a printer and a fraction collector. The whole procedure was carried out at 4°C. Elution of the bound protein was achieved by running 50 mM GSH in A-buffer through the column. The system eluted with 1.5 ml/min and fractions of 2 ml were collected. 50 µl of the fractions with high A_{280} was checked for activity with CDNB (see reaction scheme in **2.6** below). The fractions were pooled. To remove the glutathione, the pools were dialysed (Spektrapor

dialysis tubes 10 kDa cut-off) at 4°C, slowly stirred with 2 l A-buffer for 2 h, then moved to a 4.5 l of fresh A-buffer for two days and finally a new 4.5 l for 6 h. The protein samples were stored at 4°C or on ice at all times. Both synGST and terGST were purified using the same gel. Before purifying synGST (terGST was purified first) the gel was regenerated by washing the gel thoroughly with ~330 ml 0.1 M NaOAc and 1 M NaCl pH 4 followed by ~330 ml 0.1 M Tris-HCl and 0.5 M NaCl pH 8 three times. The gel was ready to be equilibrated from this step.

2.6 Characterisation experiments

Protein Concentration and Purity Determination. To determine protein concentration the Bradford Assay[16] (Bio-Rad reagents) was performed with immunoglobulin G as standard. As a rough estimate of an appropriate amount to analyse with Bradford, the samples were analysed measuring A_{280}. The absorption coefficient was estimated to be 1. Thus, Lambert-Beer's law $A = \varepsilon \times l \times c$, were ε is the absorption coefficient, l is the cuvette length and c is the concentration, could be used to immediately translate the absorption to mg/ml. The samples were also sent to be analysed for their contents in terms of amino acids. The results showed if the purified contents matched the theoretical amino acid composition of the protein compared to the genomic sequence. The protein concentrations in the samples were determined by using the concentrations of the amino acids, and comparing how many mg/ml of corresponding protein it comes to.

Formula using amino acid AA: amount AA (mol) / number of AA in GST / sample volume analysed * protein M_w = Protein concentration.

The concentrations could be used to calculate the absorption coefficient to be used when determining unknown concentrations with A_{280}.

Enzymatic Activity. As an experiment of the enzymatic activity of these new GSTs, activity was studied with a variety of substrates commonly used when studying GST cloned from other organisms. For details regarding structural details and reaction type for each substrate, see the Appendix. Table 1 shows the substrates and under what conditions the activity was measured[17]. The reaction took place at 30°C. Measuring activity was performed by first adding tempered sodium phosphate assay buffer to a quartz cuvette, followed by GSH and an appropriate amount of protein sample. Last the substrate (all dissolved in 95 % ethanol) is added and the cuvette is inverted quickly a few times before put in a spectrophotometer. The change in absorbance at the studied wavelength is picked up and activity is calculated with respect to the protein's concentration. An experiment was performed to see if there was an increase or decrease in activity with a lower and higher reaction temperature. The stability of the proteins was tested through subjecting the samples to a potentially inactivating high temperature for 10 min before adding them to the reaction.

Substrate	Sodium phosphate assay buffer properties	Concentration of substrate in reaction (mM)	Concentration of GSH in reaction (mM)	Additional substrates/comments	Assay wavelength (nm)	Extinction coefficient (1/mM/cm)
CDNB	0.1 M pH 6.5	1	1	-	340	9.6
EPNP	0.1 M pH 6.5	0.5	10	10 min incubation before adding enzyme	360	0.5
EA	0.1 M pH 6.5	0.2	0.25	-	370	5
NBC	0.1 M pH 6.5	0.25	5	20 min incubation before adding enzyme	310	1.9
CuOOH	0.1 M pH 7.0	1.5	1	0.3 U/ml glutathione reductase + 0.1 mM NADPH	340	-6.2

Table 1. Activity assay properties.

As a further study of the protein stability, the protein samples were subjected to chemical denaturation. Volumes corresponding to 4.4 µg of protein were mixed with different concentrations of Gdn-HCl. The mixture is left to react for 1 h. The mixture was put in a fluorometer cuvette (MERCK) and the sample was excited at 295 nm wavelength. The emission spectra for the samples were determined. The emission maxima for both denatured and native proteins were measured from 220-500 nm. Also, the samples were diluted 10 x and left 1 h to see if the original spectrum could be recovered. That is, if refolding of a denatured protein is possible.

Refolding kinetics. The refolding kinetics reaction was studied with the stopped-flow apparatus. Protein samples of 0.9 mg/ml for terGST and 0.09 mg/ml for synGST were as before denatured using different concentrations of Gdn-HCl for 1 h. The refolding procedure diluted the samples 11 x in A-buffer and measured for 2 s for terGST and 1 s for synGST with photomultiplier voltage at 550 V. The samples were excited at 295 nm and the overall emission over 320 nm is measured. The change in emission was monitored and attempted to be fit to a mathematical description. The curves were analysed for their rate constants, i.e. how quickly the refolding proceeds.

3 Results

3.1 Genomic screening

The results from the BLAST analyses performed at the NCBI website[9] are shown below. Alignment 1 was to find thermophilic GST which led to the settlement for *T. elongatus BP-1* GST, called terGST in this thesis. Alignment 2 shows the similarity between terGST and synGST in its relative *S. elongatus PCC 6301*. There were multiple other genes in *S. elongatus PCC 6301* annotated as GST or GST-like, but none had such striking similarity with terGST.

```
Homo sapiens GST P1-1
Thermosynechococcus elongatus BP-1 GST (terGST)

MPPYTV|VYFPVRGRCAALRMLLADQGQSWKEEVVTVETWQEGS---LKASCLYGQLPKF
      |+Y   + R + +R  L + G  ++  ++ ++  ++     LK + + G++P
---MLK|LYGGAKSRASIVRWYLEELGIPYEFVLIDLQAGEQHQPEFLKLNPM-GKVPVI

QDGDLTLYQSNTILRHLGRTLGLYGKDQQEAALVDMVNDGVEDLRCKYISLIYTNYEAGK
 DGD+ L++S  IL +L +  G   KD   AA V       + +     +L    + A
VDGDVVLWESGAILLYLAQVHGELPKDAAAAAQV------YQWVLFANSTLTQAMFPAET

DDYVKALPGQLKPFETLLSQNQGGKTFIVGDQISFADYNLLDLLLIHEVLAPGCLDAFPL
 D  + LP  LK  ET L     G+++I+G   S AD  L  +L   ++L    L  +P
RD--RQLPPLLKGIETALM----GQSYILGKDFSVADVALGSMLAYLQMLFQVDLSPYPA

LSAYVGRLSARP|KLKAFLASPEYVNLPINGNGKQ
++ YV RL  RP|
VADYVARLQQRP|AFQKGLMGARA----------
```

Alignment 1. Upper sequence in the alignment is the query sequence from human GSTP1-1 and the lower is the one for terGST. A bold character indicates exact amino acid match and a + indicates a similarity in amino acid characteristics. A – means the alignment has inserted a gap to optimise alignment further down the sequence. Note that the whole sequences are shown. The vertical bold lines indicate where the alignment was started and ended. The gene in *T. elongatus* is situated at nt 187886-188446.

```
Thermosynechococcus elongatus BP-1 GST (terGST)
Synechococcus elongatus 6301 GST (synGST)

|MLKLYGGAKSRASIVRWYLEELGIPYEFVLIDLQAGEQHQPEFLKLNPMGKVPVIVDGD
|MLKLYG A++RAS+V WYLEELG PYE V +D+ AGE  QP FL LNP GKVP +VD
|MLKLYGSARTRASLVAWYLEELGQPYESVAVDMAAGEHRQPTFLSLNPFGKVPALVDDA

VVLWESGAILLYLAQVHGEL--PKDAAAAAQVYQWVLFANSTLTQAMFPAETRDRQLPPL
  LWESGAILLYLA+  GEL  P+D A AA   QWVLFAN+TL   +F    R ++ P L
FTLWESGAILLYLAEKFGELQTPQDRAIAA---QWVLFANATLGPGLFVEANRAKEAPRL

LKGIETALMGQSYILGKDFSVADVALGSMLAYLQMLF-QVDLSPYPAVADYVARLQQRPA
L  + T L  + Y++    SV+DVA+GS LAY+ M F   DLSPYPAV DY+ RL QRPA
LGSLNTLLSQREYLVSDRLSVSDVAVGSYLAYVPMFFPDFDLSPYPAVQDYIRRLTQRPA

FQKGL|MGARA
+Q+ +|
YQQSI|GRR--
```

Alignment 2. Upper sequence in the alignment is the query sequence, this time being terGST. The lower is the one for synGST. The gene in *S. elongatus PCC 6301* is situated at nt 341171-341725 and is transcribed complementarily.

3.2 Cloning and purification

Template growth, PCR and cloning. Cultures of *T. elongatus* were attempted to be grown in other media at first, including *S. elongatus* medium BG-11, but only with DTN could a steady and dense growth be obtained. The PCR reactions rendered a DNA band in the desired range and could be performed without visible traces of contaminations. See Figure 5 for a view of a PCR reaction and Figure 6 for a view of a digestion of the pKK –D with insertion. The digested fragments from *S. elongatus* and *T. elongatus* could not immediately be ligated into the expression vector, pKK –D. Instead, the fragments were ligated into the pGEM–3Z vector

first. DNA sequencing of the insertion was successful and resulted in a sequence without gaps or mutations as compared with the sequence as given in GenBank gi:47118315.

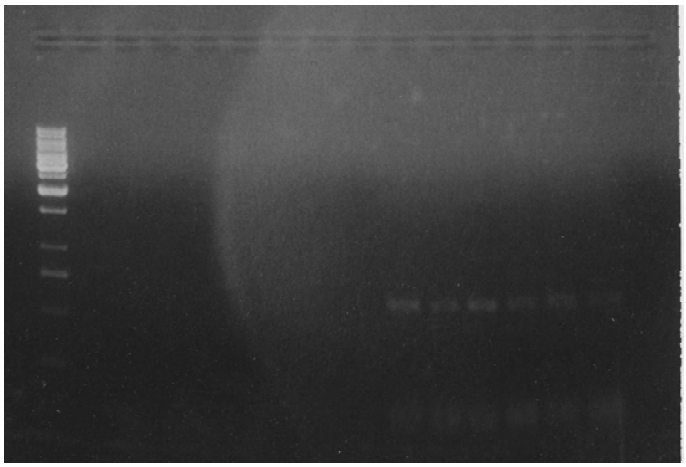


Figure 5. Reactions of the PCR of *T. elongatus*. The PCRs have bands just above the 500 bp marker (Fermentas) band (hardly visible). The closest one in length above is the 750 bp band. *T. elongatus* GST has an expected size of 561 nt. *S. elongatus* PCRs gave a little more product than *T. elongatus* did. The very weak bands further down at lower nt are primers for the PCR.

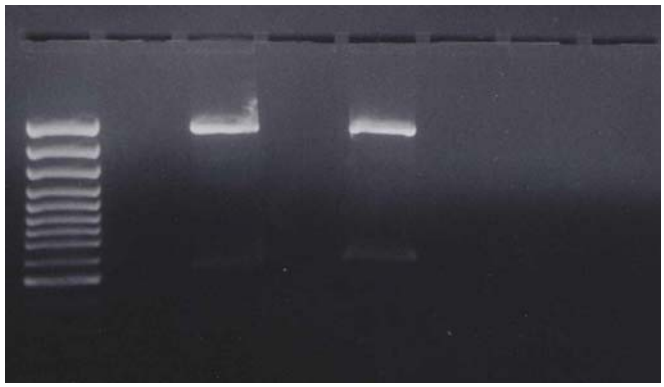


Figure 6. Restriction digestions of preparations of DNA from successful transformations of terGST ligated into pKK –D. These preparations were chosen as candidates for sequencing. The bands corresponding to the insert (weak) are 500-600bp in size as estimated by comparisons with the size marker (Fermentas).

Expressions analysis. Overexpression could be seen with small scale expressions in both cases, but to a larger extent with terGST than from synGST. Figure 7 shows SDS-PAGE gel from the final results of the purification compared to the lysates and flow-through (saved A-buffer after gel has bound O/N and prior to washing the gel before elution). Note the clearly visible contaminations in the *S. elongatus* sample. It is also hard to distinguish which of the bands in the interesting region is synGST.

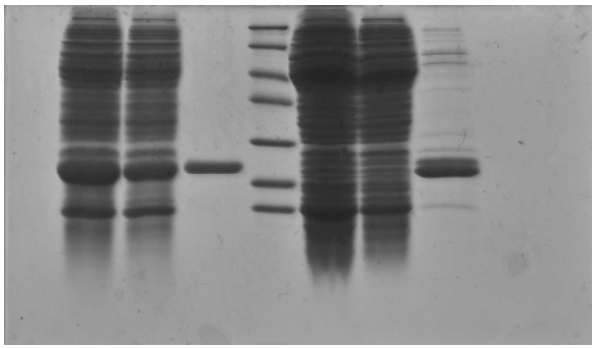


Figure 7. Leftmost is 5 µl of terGST lysate. After that 5 µl of terGST flow-through and then 15 µl of 20 x diluted purified terGST sample. The protein M_w marker (Fermentas) has bands 116.0, 66.2, 45.0, 35.0, 25.0, 18.4 and 14.4 kDa from top to bottom. Finally 5 µl of synGST lysate, 5 µl synGST flow-through and 25 µl purified synGST sample. The expression is very much higher with *T. elongatus* and there is also a lot of unbound terGST in the flow-through that was not purified.

3.3 Protein quantity and purity determination

Elution of bound protein. The absorbance peaks were detected almost immediately after elution of the GSH-gel started. Both protein elutions presented two distinct peaks. The first one with a very high absorbance in 4 fractions and a second one in 6 fractions 1-2 fractions after the first one.

Protein Concentration and Purity Determination. The amino acid analysis is presented in Tables 2 and 3. Amino acids glycine, cysteine and glutamic acid/glutamine could not be accurately determined in the analysis. This is probably because there was a significant amount of glutathione left in the protein sample, despite the dialysis. The concentration of glutamic acid and glutamine cannot be determined separately. The same is true for aspartic acid and asparagine. The Bradford Assay did not give an accurate concentration, and the standard curve had an unsatisfactory fit to its linear optimum. A_{280} for undiluted samples was 6.8 for terGST pool 1, 0.9 for terGST pool 2, 1.2 for synGST pool 1 and 0.6 for synGST pool 2. The absorbance units can be directly translated into mg/ml.

For terGST 133 µl was estimated to contain amino acids corresponding to 28 nmol of monomers. For synGST 250 µl was estimated to contain 5.9 nmol of monomers. The M_w was estimated to 20542 Da for terGST and 20550 for synGST[18].

Concentration of terGST = 4.4 mg/ml.
Concentration of synGST = 0.48 mg/ml.

Amino acid	Content in terGST as published	Theoretical %	Analysis % 1st pool	Analysis % 2nd pool
Glycine	14	-	Not determined	Not determined
Alanine	23	15.8%	16.0%	15.7%
Valine	15	10.3%	9.3%	8.9%
Leucine	27	18.5%	19.0%	18.1%
Isoleucine	7	4.8%	4.3%	4.3%
Serine	8	5.5%	5.7%	6.0%
Cysteine	0	-	Not determined	Not determined
Threonine	4	2.7%	2.8%	3.3%
Methionine	7	4.8%	4.7%	4.5%
Proline	11	7.5%	7.9%	8.6%
Phenylalanine	7	4.8%	4.6%	4.5%
Tyrosine	9	6.2%	6.1%	5.9%
Tryptophane	3	-	Not determined	Not determined
Histidine	2	1.4%	1.3%	1.4%
Lysine	8	5.5%	5.5%	5.7%
Arginine	7	4.8%	4.9%	5.1%
Aspartic acid	9	6.2%	7.7%	8.2%
Glutamic acid	9	-	Not determined	Not determined
Asparagine	2	1.4%	See aspartic acid	See aspartic acid
Glutamine	14	-	Not determined	Not determined

Table 2. The amino acid content of samples of terGST. Both pools seemed to have the same content but in lower protein concentration in pool 2. The following experiments handled only pool 1.

Amino acid	Content in synGST as published	Theoretical %	Analysis % 1st pool	Analysis % 2nd pool
Glycine	11	-	Not determined	Not determined
Alanine	22	14.7%	13.8%	13.1%
Valine	13	8.7%	8.9%	8.6%
Leucine	26	17.3%	15.8%	13.0%
Isoleucine	4	2.7%	3.5%	4.5%
Serine	13	8.7%	8.5%	7.4%
Cysteine	0	-	Not determined	Not determined
Threonine	7	4.7%	5.2%	5.9%
Methionine	3	2.0%	1.7%	1.8%
Proline	12	8.0%	8.8%	11.7%
Phenylalanine	9	6.0%	5.2%	5.1%
Tyrosine	10	6.7%	5.9%	5.3%
Tryptophane	3	-	Not determined	Not determined
Histidine	1	0.7%	1.1%	1.4%
Lysine	4	2.7%	3.8%	4.8%
Arginine	13	8.7%	8.3%	7.0%
Aspartic acid	9	6.0%	9.5%	10.1%
Glutamic acid	10	-	Not determined	Not determined
Asparagine	4	2.7%	See aspartic acid	See aspartic acid
Glutamine	10	-	Not determined	Not determined

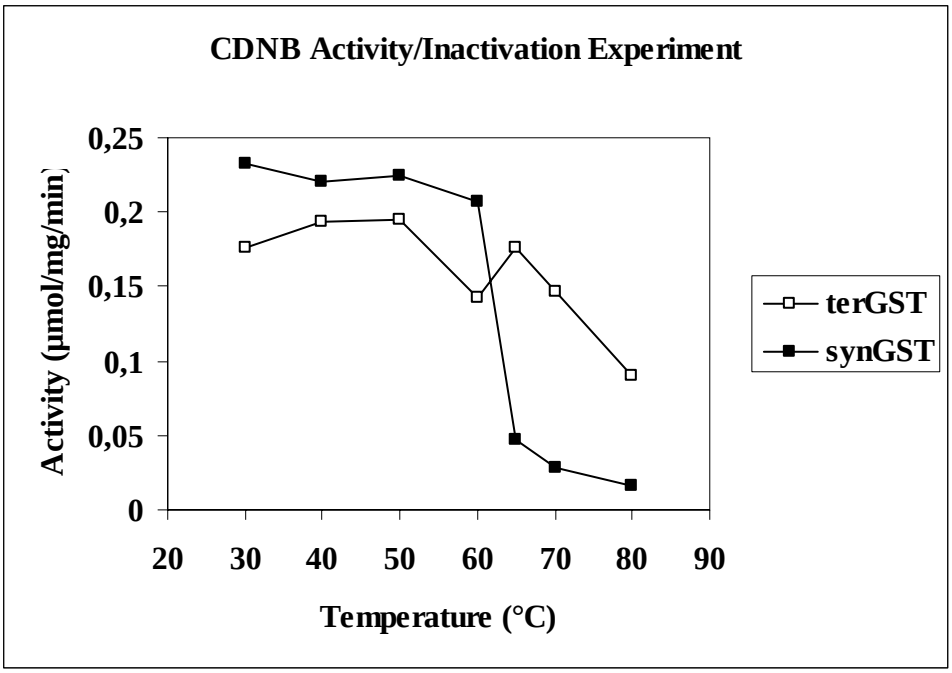
Table 3. The amino acid content of samples of synGST. As for terGST, both pools seemed to have the same content but in lower protein concentration in pool 2. Notable here when comparing theoretical % to the analysis results, is that pool 1, which is what the SDS-PAGE shows, has a lot of bands in non-specific regions. The following experiments handled only pool 1.

3.4 Enzymatic activity

The activities are measured according to Table 1 in **2.6** and activities are presented in Table 4. The results of the activity measurements are calculated with respect to the proteins' concentration mg/ml. The unit (µmol/mg/min) for activity is µmol substrate processed by one mg of protein per minute. Graph 1 shows the result of the thermal inactivation at different temperatures.

Substrate	terGST activity (µmol/mg/min)	synGST activity (µmol/mg/min)
CDNB 30°C	0.14	0.18
CDNB 25°C	0.14	0.15
CDNB 40°C	0.13	0.13
EPNP	0.1	Not detected
EA	0.021	0.059
NBC	0.049	Not detected
CuOOH	0.84	1.33

Table 4. Activities with tested substrates including experiment with an assay buffer with varying temperature.



Graph 1. The experiment with temperature inactivation showing an apparent drop in activity for synGST.

3.5 Folding equilibrium studies

Both proteins had their emission peaks at 335 nm for native structures and 358 nm for unfolded states (a concentration of Gdn-HCl of 6.25 M). Therefore, the emission for the series was measured at these wavelengths. Examples of spectra for a native and a denatured sample of the same protein are shown in Figure 8 and 9 respectively. Graphs 2 and 3 show the results for terGST and synGST respectively.

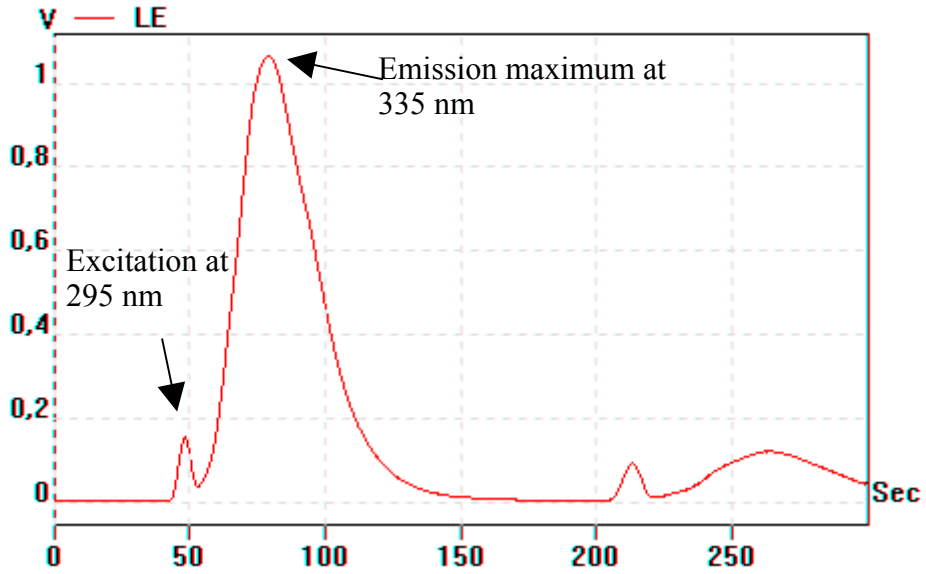


Figure 8. The emission spectrum from a terGST sample. The fluorometer used could not provide spectrums and the signal had to be transferred to a computer recording samples with 4 samples per second. Therefore, the x-axis shows an emission spectrum picked up from 220 nm to 750 nm as time samples. The small peak at about 50 s is the excitation wavelength of 295 nm and the other at 220-230 s is the double excitation wavelength.

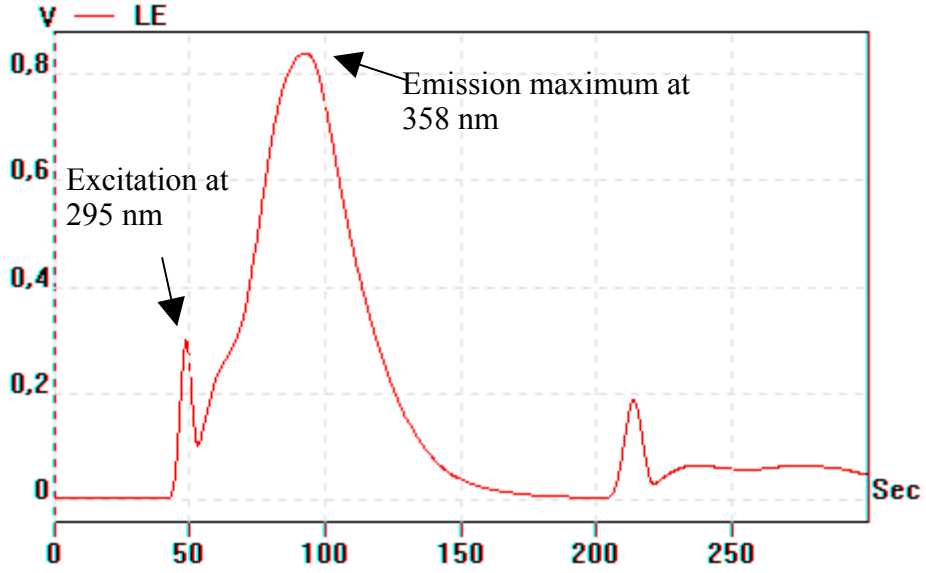
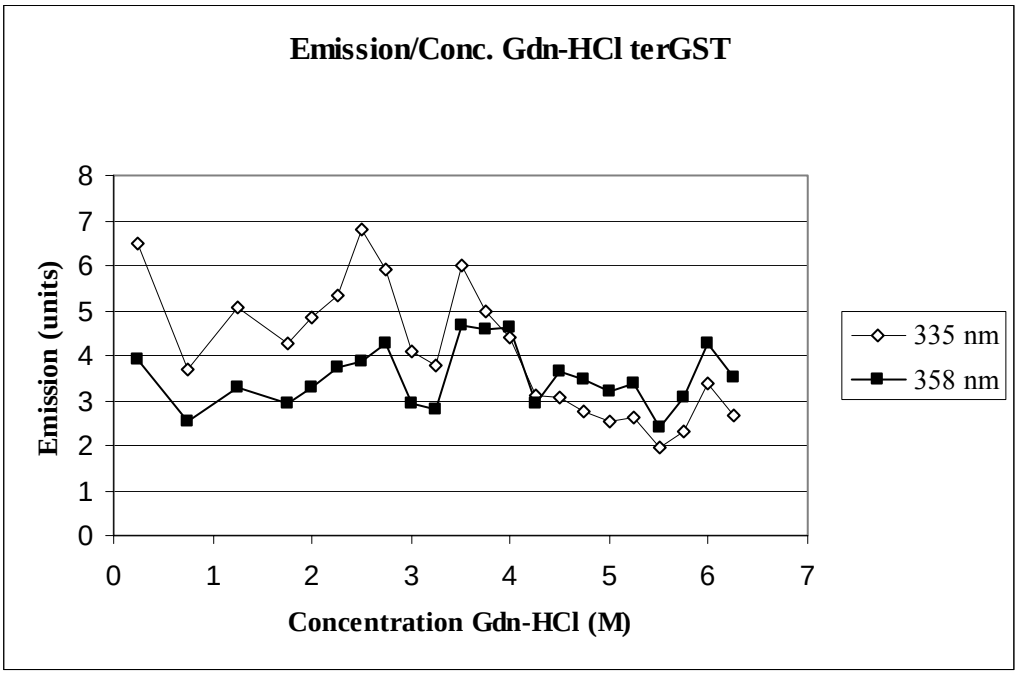
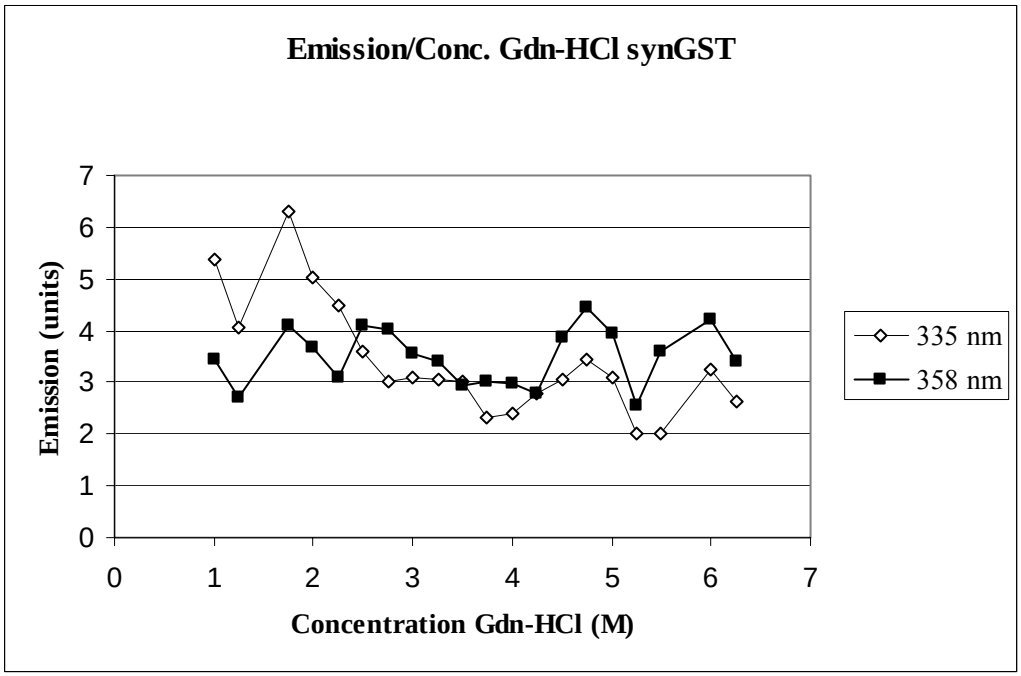


Figure 9. The emission spectrum from the same protein sample but a denatured one. Note the clear shift in emission resulting in different emission peaks.



Graph 2. Emissions of terGST at the different concentrations of Gdn-HCl. No general trend in the data could be found and therefore it was not possible to calculate the unfolding equilibrium constant K_u, since a rising concentration of Gdn-HCl should affect the fluorescence in a consistent way . An estimation of the K_u from this data would most probably be based on too much noise. The values for the emission seemed to be changing decreasingly whilst measuring.



Graph 3. Emissions of synGST at the different concentrations of Gdn-HCl. As for the case with terGST, the data was not usable for K_u calculations.

3.6 Refolding kinetics

The change in fluorescence during the refoldings of the proteins could only be fit into equations at very few concentrations. At low concentrations there was almost no change in fluorescence during the dilution of the protein-Gdn-HCl solution. The values obtained are presented in Table 5.

Protein – Gdn-HCl conc.	Amp (u)	Rate (1/s)	Protein – Gdn-HCl conc.	Amp (u)	Rate (1/s)
terGST – 1M	-	-	synGST – 1M	-	-
terGST – 2M	-	-	synGST – 2M	-0.0454	22.61
terGST – 3M	-	-	synGST – 3M	-0.107	13.1
terGST – 4M	-	-	synGST – 4M	-0.106	11.94
terGST – 5M	0.6169	2.267	synGST – 5M	-0.119	10.45
terGST – 6M	0.6424	2.412	synGST – 6M	-0.106	7.699

Table 5. The values for the amplitudes and rates of the refolding reactions. Both proteins had fluorescence emission changing according to a single exponential function. The simplest equation showing the relation between emission signal is: $E = amp*e^{-rate*t}$, with E being the emission at time t.

4 Discussion

The first interesting observation in this work has to be the surprising result of the alignment made with the BLAST tool at NCBI. The sequence similarity of terGST with human GSTP1-1 was the fifth most significant hit among microbes. It was quite interesting that a thermophilic organism would have such a high accordance with a human sequence, when experience told us that GSTs have very few conserved amino acids in general. From studying the sequence in detail there is a conserved sequence in the C-terminal part that shows the traditionally conserved sequence in terGST, but not in synGST. The first is a highly conserved glycine, number 146 in GST P1-1, and the so called N-capping box of GST P1-1 with a serine at position 150 and an aspartic acid at 153. Position 150 may vary somewhat between enzymes, but the amino acid usually does not differ much from serine in characteristics. Whilst terGST shows the exact same pattern as most mammalian GSTs in this area, synGST lacks the highly conserved glycine. The importance of these amino acids for folding has been proven to be critical for GST P1-1[5,10,11].

Interestingly, neither the terGST nor synGST could be cloned into the expression vector immediately, despite repeated efforts. It seemed a subcloning of the DNA into the carriage vector pGEM–3Z was necessary in all cases. The restriction cleavage and ligation necessary was performed in vitro, under the same circumstances and with the same restriction enzymes for both plasmids. Therefore it appears like cleaving the fragments out of pGEM–3Z, gives the fragments a higher accessibility for ligation into pKK –D. This still does not explain why it was possible to cleave and insert into pGEM–3Z immediately from PCR product, but it is a smaller plasmid, more easily processed and amplified vector with high cellular copy number. Worth mentioning was also the result of a very pure PCR product, which indicates a very specific PCR reaction, even though unpurified cultures were used as templates.

The expression level of synGST was remarkably lower that for terGST. The proteins are both prokaryotic and the genes ought to present codons easily transcribed and translated in *E. coli*. In mammalian genomes it can often be effective to change codons rarely used by *E. coli* to more common ones. Considering that the optimal growth temperature of *S. elongatus* is more

close to *E. coli* used as the expression system, it could be hypothesised that this protein would be more effectively expressed as compared with the GST from the thermophilic *T. elongatus*. Expression rates of wild type GST P1-1 have usually been quite high, and it could be because of the similarity with this protein, that terGST is very much more efficiently expressed. The expression rates for terGST were repeatedly high also for the small scale experiments. The results for synGST in terms of purity are very promising looking at the amino acid content, but the SDS-PAGE shows a different picture. Since the concentration of contaminants is much lower than that of synGST the amino acid content is adequate but adding more purifying steps will is necessary. Although a gel filtration cannot separate with infinite resolution, it would take care of at least some of the contaminants in the synGST sample. More importantly, it would also remove the excess GSH that was left in the sample after the dialysis.

Looking at the activities for these new GSTs, it is interesting to see that activity with such a wide range of reaction mechanisms and variants of molecules could be observed. The first observation was that the mesophilic synGST has a higher overall activity with the studied substrates than terGST. It can be argued if the level is significantly higher, but the result of a higher activity was repeated for all substrates. Due to their difference in natural habitat, it can be fortunate to investigate other substrates to see if they react more easily with terGST. The studied substrates were chosen because the human GST T1-1, which is the closest homologue of the bacterial GSTs, was very active with these substrates[19]. EPNP and NBC are substrates that GST T1-1 has high activity with. Although the cyanobacterial GSTs were indeed shown to be active on these substrates, the activities are in the order of 1/10 of the activity of GST T1-1. Interestingly, the coupled reaction to reduce CuOOH showed 2-3 times higher activity in these new enzymes. GST T1-1 does not readily catalyse substitution of the chloride in CDNB, but CDNB showed about 3 times higher activity with the enzymes in this study. EA is a substrate more easily processed by enzymes like GST P1-1, and the activities observed in this study are in fact almost 1/100 of that of GST P1-1. Actually this goes for CDNB also, but it should be emphasised that activities vary very much between GST classes even in the same organism.

A proposed hypothesis was that terGST would be more unaffected or maybe even helped by an increased assay temperature, but no such significant conclusion could be drawn. However, the thermal stability of terGST appears to be undoubted. Even though the data is a little shaky, there is no apparent decrease in activity until 70°C. The very sudden and steep drop for synGST is more in line with how thermal stability is described for other GSTs.

Unfortunately, the equilibrium studies for the denaturation of both enzymes failed. The conclusion to be drawn is that the denaturation reactions of these enzymes have not been allowed to reach equilibrium. This would further explain the fact that the emission values at the chosen wavelengths changed as the measurements went on. However, this experiment can most probably be redone with a more successful result, since it was possible to see a clear change in spectrum when denaturing the proteins. The refolding of the proteins using the stopped-flow analysis shows that refolding was possible, and showed rate constants of reasonable values, but that optimal circumstances could give more results on the kinetics behind it. For example, it was not possible to get refolding rates for terGST at low concentrations of Gdn-HCl, which would indicate denaturation method should be changed. This is also a result pointing to a high stability of terGST compared to that of synGST, not only in terms of thermal resistance, but also structure stability.

To further investigate these enzymes and their characteristics it is necessary to prepare a high purity synGST and investigate method optimisations. Even though the final goal of studying protein folding was rather unsuccessful, this method proves that continuing the work would probably be fruitful. The study does show that that a chemical denaturation of these proteins is reversible by fluorescence spectroscopy and regaining activity (data not shown). Furthermore, since prokaryotic GSTs have been such a small area of research in the past, this study shows that the traditional purification problems hindering this work are not always present. It was also possible to get adequate expressions; the proteins did bind to the affinity matrix and could easily be assayed for activity. With only minor protocol enhancements, there is great hope in obtaining conclusive results on both folding stability and kinetics.

5 Acknowledgements

Thanks to Dr. Miwa Sugiura at The Department of Applied Biological Chemistry, Faculty of Agriculture, Osaka Prefecture University, Osaka, Japan for providing *T. elongatus* culture and recipe for DTN medium.

Also, useful help was provided by Peter Lindblad at The Department of Physiological Botany, Uppsala University on how to grow, store and isolate genes from cyanobacteria.

The biggest thanks go to my supervisor, Dr. Gun Stenberg. With enormous patience and pedagogical skills, she has provided me with basically all my knowledge of biochemical research and laboratory skills.

6 Appendix

Contents of A-buffer

Tris-HCl – 10 mM
EDTA – 1 Mm

HCl added to set pH at 7.8.

Contents of 2xYT

NaCl – 5 g/l
Yeast extract – 10 g/l
Tryptone – 16 g/l

Contents of LB

NaCl – 10 g/l
Yeast extract (OXOID) – 5 g/l
Tryptone – 10 g/l

Contents of BG-11

Base Medium

$NaNO_3$ – 1.5 g/l
K_2HPO_4 x $3H_2O$ – 0.04 g/l
$MgSO_4$ x $7H_2O$ – 0.075 g/l
$CaCl_2$ x H_2O – 0.032 g/l
Citric acid – 0.006 g/l
Ferric ammonium citrate – 0.006 g/l
EDTA – 0.001 g/l
Na_2CO_3 – 0.02 g/l
Trace metal mix – 1 ml/l

Trace Metal Mix

H_3BO_4 – 2.86 g/l
$MnCl_2$ x $4H_2O$ – 1.81 g/l
$ZnSO_4$ x $7H_2O$ – 0.222 g/l
Na_2MoO_4 x $2H_2O$ – 0.39 g/l
$CuSO_4$ x 5H2O – 0.079 g/l
$Co(NO_3)_2$ x $6H_2O$ – 0.049 g/l

Contents of DTN

DTN Medium (1 l)

DTN is prepared from stocks of the 7 solutions below. Note: $NaHCO_3$ was added after autoclaving.

40 x DTN Medium Base – 25 ml
$FeCl_3$ – 0.26 ml
Sulphuric Mixture – 5 ml
Tricine – 16 ml
NH_4Cl – 0.2 ml
Micronutrients – 0.5 ml

10 M NaOH was added to adjust pH to 7.6. The volume was adjusted to ~990 ml and the solution was sterilised by autoclaving.

$NaHCO_3$ – 10 ml (0.2 μm filter sterilised)

Volume adjusted to 1 l with sterile H_2O.

40 x DTN medium base

Bisodium EDTA – 7.6 g/l
$MgSO_4$ x $7H_2O$ – 4 g/l
KNO_3 – 4 g/l
$NaNO_3$ – 28 g/l
$NaHPO_4$ x $2H_2O$ - 7 g/l
$CaCl_2$ x $4H_2O$ – 2.8 g/l
NaCl – 0.32 g/l

10 M NaOH added to adjust to pH 7.6.

Sulphuric mixture

$Na_2S_2O_3$ x $5H_2O$ – 40 g/l
Na_2SO_3 – 20 g/l
$NaHSO_3$ – 33.6 g/l

Micronutrients

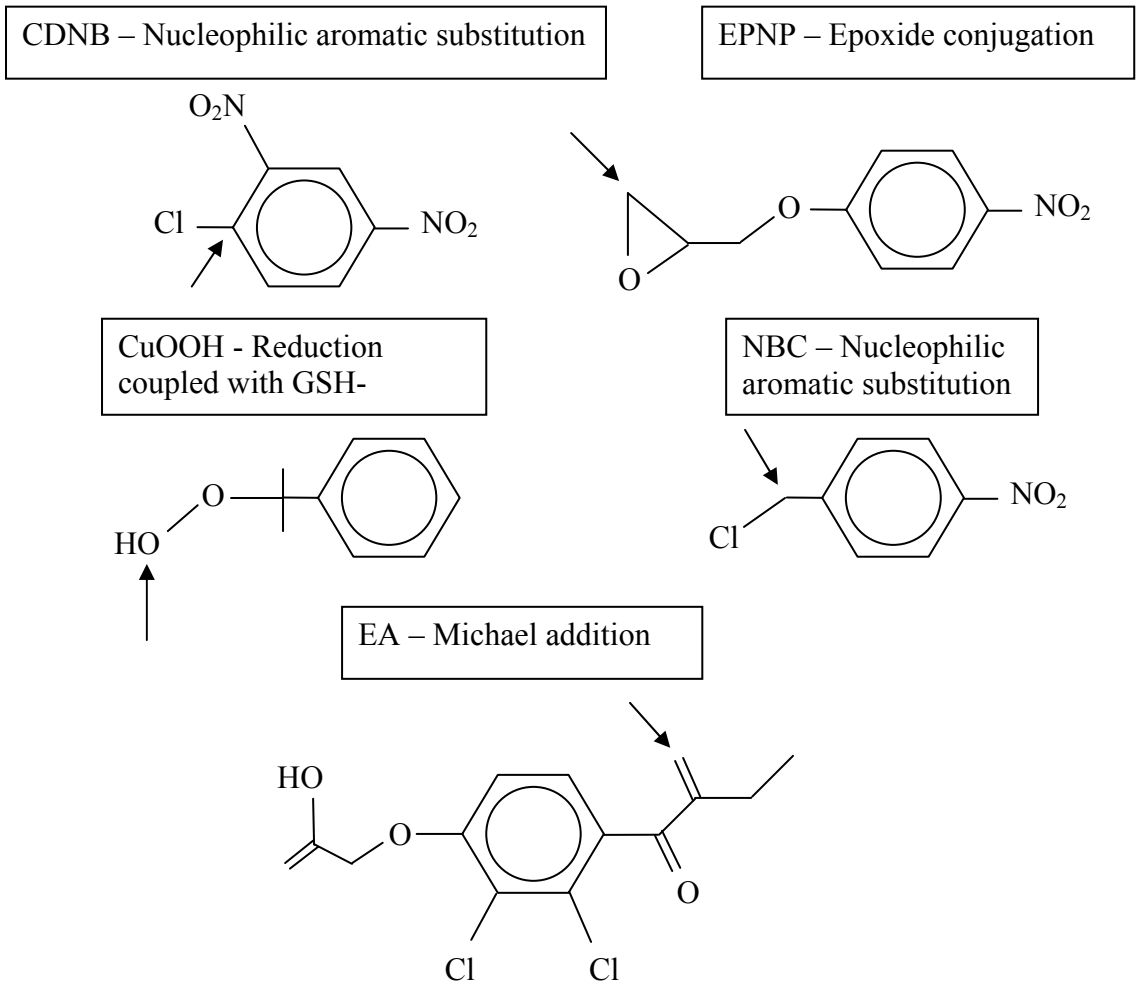
H_2SO_4 (concentrated) – 0.5 ml/l
H_3BO_4 – 0.5 g/l
$MnSO_4$ x H_2O – 2.7 g/l
$ZnSO_4$ x $7H_2O$ – 0.5 g/l
$CuSO_4$ x $5H_2O$ – 0.025 g/l
Na_2MoO_4 x $2H_2O$ – 0.025 g/l
$CoCl_2$ x $6H_2O$ – 0.045 g/l
$(NH_4)_2Ni(SO_4)_2$ x $6H_2O$ – 0.019 g/l
Na_2SeO_4 – 0.004 g/l

Other solutions

$FeCl_3$ x $6H_2O$ – 3.9 g/l
Tricine – 179.2 g/l (+ 10 M NaOH to pH 8.)
NH_4Cl – 53.45 g/l
$NaHCO_3$ – 84 g/l

These solutions were prepared separately.

Substrate structures and reaction mechanisms



The arrows indicate the atom to which GSH is added.

7 References

[1] Sheehan D, Meade G, Foley V M and Dowd C A. *Structure, Function and Evolution of Glutathione Transferases:Implications for Classification of Non-Mammalian Members of an Ancient Superfamily.* Biochemical Journal, 2001. 360, 1-16.

[2] Vuilleumier S and Pagni M. *The Elusive Role of Bacterial Glutathione S-Transferases: New Lessons from Genomes*. Applied Microbiology and Biotechnology, 2002. 58, 138-146.

[3] Chasseaud L F. *The Role of Glutathione and Glutathione S-transferases in the Metabolism of Chemical Carcinogens and Other Electrophilic Agents*. Advances in Cancer Research, 1979. 29, 175-274.

[4] Armstrong R N. *Structure, Catalytic Mechanism and Evolution of the Glutathione Transferases*. Chemical Research in Toxicology, 1996. 10, 2-18.

[5] Dragani B, Stenberg G, Meloni S, Petruzzelli R and Mannervik B. *The Conserved N-capping Box in the Hydophobic Core of Glutathione S-Transferase P1-1 Essential for Refolding.* The Journal of Biological Chemistry, 1997. 272 (41), 25518-25523.

[6] Vuilleumier S. *Bacterial Glutathione S-Transferases: What Are They Good For?* Journal of Bacteriology, 1997. 179 (5), 1431-1441.

[7] Snyder M J and Maddison D R. *Molecular Phylogeny of Glutathione S-Transferases.* DNA and Cell Biology, 1997. 16 (11), 1373-1384.

[8] Nakamura Y, Kaneko T, Sato S, Ikeuchi M, Katoh H, Sasamoto S, Watanabe A, Iriguchi M, Kawashima K, Kimura T, Kishida Y, Kiyokawa C, Kohara M, Matsumoto M, Matsuno A, Nakazaki N, Shimpo S, Sugimoto M, Takeuchi C, Yamada M and Tabata S. *Complete Genome Structure of the Thermophilic Cyanobacterium Thermosynechococcus elongatus BP-1.* DNA Research, 2002. 9 (4), 123-130 135-148.

[9] NCBI, National Centre for Biotechnology Information: http://www.ncbi.nlm.nih.gov/BLAST/ (31 Jan 2005)

[10] Hegazy U M, Mannervik B and Stenberg G. *Functional Role of the Lock and Key Motif at the Subunit Interface of Glutathione Transferase P1-1.* The Journal of Biological Chemistry, 2004. 279 (10), 9586-9596.

[11] Kong G K-W, Polekhina G, McKinstry W J, Parker M W, Dragani B, Aceto A, Paludi D, Principe D R, Mannervik B and Stenberg G. *Contribution of Glycine 146 to a Conserved Folding Module Affecting Stability and Refolding of Human Glutathione Transferase P1-1.* The Journal of Biological Chemistry, 2003. 278 (2), 1291-1301.

[12] Kellis T J, Nyberg K and Fersht A R. *Energetics of Complementary Side-Chain Packing in Protein Hydrophobic Core*. Biochemistry, 1989. 28 (11), 4916-4921.

[13] Björnestedt R, Widersten M, Board P G and Mannervik B. *Design of Two Chimaeric Human-Rat Class Alpha Glutathione Transferases for Probing the Contribution of C-Terminal Segments of Protein Structure to the Catalytic Properties*. The Biochemical Journal, 1992. 282 (Part 2) 505-510.

[14] Sugita M and Sugita C. *Unpublished*. Submission to NCBI 21 Dec 2004.

[15] Sambrook J and Russell D W. *Molecular Cloning: A Laboratory Manual 3rd Editionl.* Cold Spring Harbor Laboratory Press, 2001. Chapter 1, Protocol 25 (Modified using NaOAc/$MnCl_2$).

[16] Bradford, M M. *A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding*. Analytical Biochemistry, 1976. 72, 248-254.

[17] Jemth P and Mannervik B. *Measurement of Glutathione Transferases.* Current Protocols in Toxicology, 1999. 6.4.1-6.4.10.

[18] ExPASy, Expert Protein Analysis System: http://www.expasy.org/tools/pi_tool.html (15 Apr 2005)

[19] Shokeer A, Larsson A K and Mannervik B. *Residue 234 in Glutathione Transferase T1-1 Plays a Pivotal Role in the Catalytic Activity and the Selectivity Against Alternative Substrates.* The Biochemical Journal, 2005. 388 (Part 1), 387-392.