Purification and determination of oxidative modifications of the thermostable alcohol dehydrogenase from *Thermococcus guaymasensis*

by

Navin Asokumar

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Biology

Waterloo, Ontario, Canada, 2019

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Hyperthermophiles are microorganisms that grow optimally at 80°C and above. They have been found to be resilient to extremes of pH, redox potential, pressure and salinity. Moreover, they can utilize a wide range of carbohydrates as carbon source and produce ethanol as an end product. Alcohol dehydrogenase (ADH) is a key enzyme responsible for alcohol production, catalyzing interconversions between alcohols and corresponding ketones or aldehydes. ADHs from hyperthermophiles are of great interest due to their thermostability, high activity and enantioselectivity. One such zinc-containing homotetrameric ADH from hyperthermophilic archaeon *Thermococcus guaymasensis* was found to be 39,395.72 Da per subunit in size, with activity of 1,049 U/mg. Regardless of being a zinc-containing enzyme, this ADH was found to lose its activity when exposed to air. Since zinc is a divalent cation which cannot be further oxidized, a plausible explanation would be that certain amino acids, especially cysteines, succumb to oxidative change, leading to the inactivation of the enzyme under aerobic conditions. TgADH has four Cys residues per subunit; Cys39 (which coordinates with zinc), Cys56, Cys212, and Cys305. The recombinant wild-type (45,260.13 Da per subunit) and Cys mutant TgADH were partially purified from recombinant *E. coli* host using a heat treatment step, with a specific activity of over 300 U/mg. MALDI-TOF mass spectrometry was used for investigating the modifications due to oxidation in the partially purified recombinant wild-type TgADH, when exposed to air. External calibration resulted in an accuracy range of 0.0013 % - 0.19 %. The higher end of the accuracy range was ≈ 145 times more than the lower end, invalidating the method for the required purpose. However, since 0.0013 % of accuracy was seen for Protein A $[M+H]^+$, which has a mass almost equal to recombinant TgADH, and due to the impure nature of the sample, the method was used for analysis of modifications in the recombinant TgADH, without prior digestion of the

enzyme. The measured masses of the active and air-oxidized recombinant TgADH were 47 and 12 Da less than reported mass, respectively. The H_2O_2-oxidized (1:10, 1:40, 1:160, and 1:400 mol/mol protein subunit:oxidant) TgADH were less than reported masses, whereas they should have been more than the reported mass due to the modification(s). Only the 1:1600 mol/mol H_2O_2-oxidized sample showed an increase of 360 Da suggesting that Cys residues, along with the Met residues were oxidized. The difference between the measured masses of the active and air-oxidized recombinant TgADH was ≈ 35 Da, which is almost equal to two oxygens. However, due to the large discrepancy in the accuracy range it cannot be conclusively said that two oxygens are being added when the enzyme is exposed to air. Enzyme activity analysis of partially purified C212S and C305S TgADH mutants showed no increase in oxygen tolerance compared to the wild-type TgADH. Since previous studies have eliminated the oxidation of Cys56 being the cause for the loss of enzyme activity, this study puts forward the hypothesis that the oxidation of the zinc coordinating Cys residue, Cys39, could be the reason behind the loss of activity under aerobic conditions.

Acknowledgements

I would like to thank my supervisor, Dr. Kesen Ma, for giving me this opportunity and guiding me from the beginning to end of my degree. I thank my committee members, Dr. Todd Holyoak and Dr. Trevor Charles, for their invaluable advice regarding my research. I also would like to thank Dr. Cheryl Duxbury for her kindness and support throughout my degree.

I extend my sincere gratitude to my parents, Asokumar Nadesapillai and Sakunthala Asokumar, uncle, Sivakumar Thangaraj, and grandmother, Kamalam Thangaraj, who have supported me immensely, been there for me whenever I needed, shared my joys and sorrows, picked me up when I was down and made sure I was focused on my goals.

I also like to thank my former and present lab mates and colleagues that I have worked with during my time: Sarah Kim, Faisal Alharbi, Mobolaji Adegboye, Olayinka Aiyegoro, Junrui Zhan, Sadique Mostafa bin Amin, and Can Wang. I thank Sura Ali for her support and advice for my career. I thank Lian Hickey for making sure I was feeling well.

I sincerely thank Valerie Goodfellow from the Mass Spectrometry Facility of University of Waterloo for helping me with my mass spectrometry analyses.

Finally, I thank God for taking me by my hand and helping me cross each and every obstacle.

Dedication

This thesis is dedicated to my Appa and Amma

Table of Contents

List of Figures ..... **ix**
List of Tables ..... **xi**
List of Abbreviations ..... **xii**
1. Introduction **1**
1.1 Hyperthermophiles 2
1.2 Microbial production of ethanol 2
1.3 Alcohol dehydrogenase in hyperthermophiles 7
1.3.1 Zinc-containing ADHs 8
1.3.2 Iron-dependent ADHs 10
1.3.3 Native and recombinant ADH of *Thermococcus guaymasensis* 11
1.4 Inactivation of enzymes containing reactive Cys residue(s) 15
1.5 MALDI-TOF mass spectrometry 20
1.6 Aims of this study 23
2. Materials and Methods **25**
2.1 Chemicals used 26
2.2 Instruments used 27
2.3 Microorganism(s) 28
2.4 Growth and storage of microorganism(s) 28
2.5 Enzyme purification 29
2.5.1 Preparation of cell-free extract 29
2.5.2 Purification of recombinant TgADH 30
2.5.3 Heat treatment conditions 31
2.6 Determination of protein concentration 31
2.7 Determination of enzyme activity 32
2.7.1 Oxygen sensitivity 32
2.7.2 Inactivation by H_2O_2 32
2.8 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) 33
2.9 MALDI-TOF mass spectrometry 33
2.10 Enzyme structural modelling 34
3. Results **35**

3.1 Over-expression of *T. guaymasensis* ADH in *E. coli* 36
3.2 Selection of media for growth and overexpression of recombinant *E. coli* cells.....38
3.3 Heat treatment conditions.....41
3.3.1 Incubation temperature.....41
3.3.2 Incubation period.....45
3.3.3 Concentration of CFE for incubation 48
3.4 Purification of the recombinant wild-type *T. guaymasensis* ADH 50
3.5 Oxygen sensitivity of C212S and C305S TgADH mutants 56
3.6 Mass spectrometry analysis.....62
3.6.1 Calibration 62
3.6.2 Optimal matrix.....66
3.6.3 Analysis of oxidized recombinant TgADH.....68
3.7 Oxidative degradation of recombinant TgADH.....72
3.8 Structural modelling and analysis of TgADH.....81
4. Discussion & Conclusion 83
4.1 Growth and expression of recombinant wild-type TgADH 84
4.2 Purification of recombinant wild-type TgADH 85
4.3 Oxygen sensitivity, oxidative modification(s), and degradation of TgADH 86
4.3.1 Oxygen sensitivity of Cys mutant TgADH 86
4.3.2 Oxidative modification(s) in recombinant wild-type TgADH 88
4.3.3 Degradation of recombinant TgADH.....90
4.4 Structure modelling of TgADH and analyses 91
4.5 Conclusions 92
4.6 Recommendations for future studies.....93
5. References.....95
6. Appendix.....104
List of characterised hyperthermophilic ADHs, to date, and their biophysical and catalytical properties.....104

List of Figures

Figure 1-1: Glycolysis/gluconeogenesis pathway, as reported in KEGG database 6

Figure 1-2: Maps of pGEM-T Easy and pET-30a vectors showing the insertion site of TgADH encoding gene 13

Figure 1-3: Amino acid sequence of recombinant wild-type TgADH..... 14

Figure 1-4: Sequential oxidation states of cysteine 16

Figure 1-5: Brief strategy of this study 24

Figure 3-1: Analysis of over-expression level of *Tg*ADH in *E. coli* induced by IPTG, using SDS-PAGE (12.5 %) 37

Figure 3-2: Analysis of over-expression of *Tg*ADH in *E. coli*, in 2YT, LB and TSB media, induced by IPTG 39

Figure 3-3: Analysis of over-expression of *Tg*ADH in *E. coli*, in 2YT, LB and TSB media, induced by IPTG 40

Figure 3-4: Heat treatment of CFE at 60°C for 1 hour 42

Figure 3-5: Heat treatment of CFE at five different temperatures..... 43

Figure 3-6: Heat treatment of CFE at five different temperatures..... 44

Figure 3-7: Heat treatment of CFE for different durations at 100°C 46

Figure 3-8: Heat treatment of CFE for different durations at 100°C 47

Figure 3-9: Heat treatment of different concentrations of CFE at 100°C..... 49

Figure 3-10: Purification of recombinant TgADH using a phenyl-sepharose column, with segmented gradient..... 51

Figure 3-11: Fractions collected from phenyl-sepharose column when segmented gradient was used 52

Figure 3-12: Purification of TgADH using a phenyl-sepharose column, with a linear gradient.. 53

Figure 3-13: Fractions collected from phenyl-sepharose column when linear gradient was used ..... .54

Figure 3-14: Partial purification of TgADH using heat treatment at 100°C for 10 minutes..... 55

Figure 3-15: Oxygen sensitivity of purified recombinant wild-type TgADH 57

Figure 3-16: Oxygen sensitivity of partially purified TgADH Cys mutant C305S 59

Figure 3-17: Oxygen sensitivity of partially purified TgADH Cys mutant C212S 60

Figure 3-18: Inactivation of C212S, C305S and wild-type TgADH 61

Figure 3-19: External calibration of mass spectrometer using a standard mixture 63

Figure 3-20: The difference between reported and measured masses of each component in the standard mixture 65

Figure 3-21: Selecting a matrix for MALDI-TOF analysis..... 67

Figure 3-22: Correlation between enzyme activity and change in mass due to oxidation by H_2O_2 71

Figure 3-23: Degradation of recombinant TgADH in phenyl-sepharose fractions 73

Figure 3-24: Enzyme activity of TgADH under different storage conditions 74

Figure 3-25: Degradation of TgADH under different storage conditions 76

Figure 3-26: Streak plate of heat treated CFE stored under anaerobic conditions at room temperature after 144 hours 77

Figure 3-27: Degradation of TgADH after inoculating with colony A, B or C..... 78

Figure 3-28: Degradation of TgADH by colonies A, B and C under aerobic and anaerobic conditions 80

Figure 3-29: Modelled structure of TgADH..... 82

List of Tables

Table 1-1: Selection of enzymes that that are known to be inactivated due to Cys thiol oxidation 19

Table 2-1: Chemicals used 26

Table 2-2: Instruments used 27

Table 3-1: Purification of recombinant TgADH using a phenyl-sepharose column, with segmented gradient 51

Table 3-2: Purification of recombinant TgADH using a phenyl-sepharose column, with linear gradient 53

Table 3-3: Partial purification of recombinant TgADH using heat treatment 55

Table 3-4: Partial purification of recombinant Cys mutants 58

Table 3-5: External calibration of mass spectrometer using a standard mixture 64

Table 3-6: Comparison of protein masses measured by MALDI-TOF/MS vs reported values ...69

Table 3-7: Growth of colonies A, B and C under aerobic and anaerobic conditions and their ability to degrade TgADH 80

List of Abbreviations

ADH	Alcohol dehydrogenase
AH	After heat treatment
BH	Before heat treatment
CFE	Cell-free extract
CV	Column volume
DTT	1,4-Dithiothreitol
Fe	Iron
FPLC	Fast Protein Liquid Chromatography
HlADH	Horse Liver alcohol dehydrogenase
MALDI	Matrix assisted laser desorption/ionization
NAD	Nicotinamide adenine dinucleotide
NAD^+	Nicotinamide adenine dinucleotide (oxidized)
NADP	Nicotinamide adenine dinucleotide phosphate
$NADP^+$	Nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
SDT	Sodium dithionite
TbADH	*Thermoanaerobacter brockii* alcohol dehydrogenase
TgADH	*Thermococcus guaymasensis* alcohol dehydrogenase
TOF	Time of flight
YADH I	Yeast alcohol dehydrogenase I
Zn	Zinc

1. Introduction

1.1 Hyperthermophiles

Hyperthermophilic microorganisms are those which grow optimally at 80°C and above (Stetter, 2006; Ma and Tse, 2015). Many hyperthermophilic archaea and bacteria have been isolated from aquatic environments that are heated geothermally or volcanically, such as hydrothermal vents, hot springs and solfataric fields (Stetter, 2006; Ma and Tse, 2015). Hyperthermophilic microorganisms are not only capable of thriving at high temperatures, but are also able to grow at extremes of pH, redox potential, pressure, and salinity. Moreover, they have been found to utilize a broad range of substrates due to flexible metabolism (Chang and Yao, 2011; Tse and Ma, 2016). This makes them prime research candidates in order to understand thermostability, and the functioning and properties of thermostable enzymes for biotechnological applications.

One of the most interesting enzyme groups from hyperthermophiles are the alcohol dehydrogenases (ADHs) (Radianingtyas and Wright, 2003). They are studied for their industrial importance and to understand the characteristics of evolutionary relationships of other enzymes from different sources (Radianingtyas and Wright, 2003; Ma and Tse, 2015). ADHs are industrially valued for their use in production of potable and commodity alcohol, solvents, acetic acid, and enantiomerical products. (Radianingtyas and Wright, 2003; Yan, 2010; Ma and Tse, 2015).

1.2 Microbial production of ethanol

Concerns such as energy security, economic stability, environmental impact, and climate change have increased the demand for biofuels as substitutes for fossil fuels. The Energy Information Administration of USA predicts that the production of fuel ethanol will increase to ≈ 25 billion

gallons by 2035 (National Research Council, 2011). Shell LTD hopes to utilize plant biomass on a scale exceeding that of oil, in 2060 (Lynd et al., 1999). One of the most popular liquid biofuels is ethanol. When mixed with gasoline at a certain percentage, bio-ethanol can be used as a drop-in substitute for liquid fossil fuel (Regalbuto, 2009). Moreover, it is also a precursor for many other commodity chemicals such as acetaldehyde, acetic acid and their derivatives.

The most common ethanologenic microorganisms intensively studied or used in industry are *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Escherichia coli*, and *Klebsiella oxytoca*. However, these organisms cannot be termed as 'perfect' ethanogens. A suitable microorganism that can efficiently convert raw biomass to ethanol should have some other features additional to being able to ferment a wide variety of sugars. These features include but are not limited to having high ethanol yield, high tolerance to fermentation products and simple growth requirements (Zaldivar et al., 2001).

There is a widespread interest in using hyper/thermophilic microorganisms for production of ethanol. Hyper/thermophiles are a suitable choice because they naturally produce highly stable enzymes which can withstand fluctuations in temperature, pH, and solvent concentration (Ying and Ma, 2011; Ma and Tse, 2015). Moreover, compared to common ethanologenic mesophiles, many hyper/thermophiles can ferment hexose and/or pentose sugars as well as more complex substrates such as cellulose, hemicellulose, and xylan (Chang and Yao, 2011). The requirements of high temperature and anaerobic environment (in most cases) of these organisms reduce or eliminate the possibility of oxidative degradation, cooling of the fermenter, gas solubility, and external contamination. High temperature also facilitates product solubility and makes it much easier to recover ethanol due to its high volatility (Zaldivar et al., 2001; Taylor et al., 2009; Olson et al., 2015).

The use of hyper/thermophilic organisms and/or enzymes also has its disadvantages, the main one being their low substrate tolerance and inhibition by the synthesized products (Chang and Yao, 2011). Also, some of the thermophiles are mixed fermenters, producing many types of products (including ethanol) during growth (Zaldivar et al., 2001). However, recent past has seen the application of metabolic engineering to circumvent many of the issues and have developed strains that give ethanol yields as much as the yeast-based systems (Olson et al., 2015). Thermophilic species such as *Clostridium thermocellum*, many members of the genus *Geobacillus*, extremely thermophilic species such as *Thermoanaerobacter ethonolicus*, *Thermoanaerobacter tengcongensis*, and *Thermoanaerobacter pentosaceus* have been intensively studied for high ethanol production capabilities. However, compared to thermophilic ethanol producers, very little is known about ethanol production in extremely thermophilic and hyperthermophilic microorganisms. A few archaeal hyperthermophilic species such as *Pyrococcus furiosus*, *Thermococcus* sp. strain ES1, *Thermococcus guaymasensis* and bacterial hyperthermophiles such as those belonging to Thermotogales have been reported to produce small amounts of ethanol, but more studies are required to dissect the metabolic pathways involved and possibly increase the ethanol produced.

Considerable interest has been shown in microbial ADHs (EC 1.1.1.1, Figure 1-1) due to their importance in process and production roles, the main one being generation of bio-alcohol, mainly ethanol. Alcohol fermentation in microbes, which results in ethanol as primary or co-metabolite, can follow two pathways (Fig 1-1). Pyruvate and acetaldehyde are the common intermediates for both the pathways, however the route taken for the formation of acetaldehyde from pyruvate can either be 1) A two-step pathway where pyruvate is non-oxidatively decarboxylated to acetaldehyde and carbon dioxide, which is catalyzed by pyruvate decarboxylase. Acetaldehyde is then converted

to ethanol that is catalyzed by ADH (Eram and Ma, 2013) or 2) A three-step pathway where pyruvate is oxidatively decarboxylated to acetyl-coenzyme A (acetyl-CoA) by the metalloenzyme pyruvate ferredoxin oxidoreductase and/or pyruvate formate lyase. Acetyl-CoA is then converted to acetaldehyde by a bifunctional CoA-dependent-acetylating acetaldehyde dehydrogenase. Acetaldehyde is then reduced to ethanol by ADH (Eram and Ma, 2013).

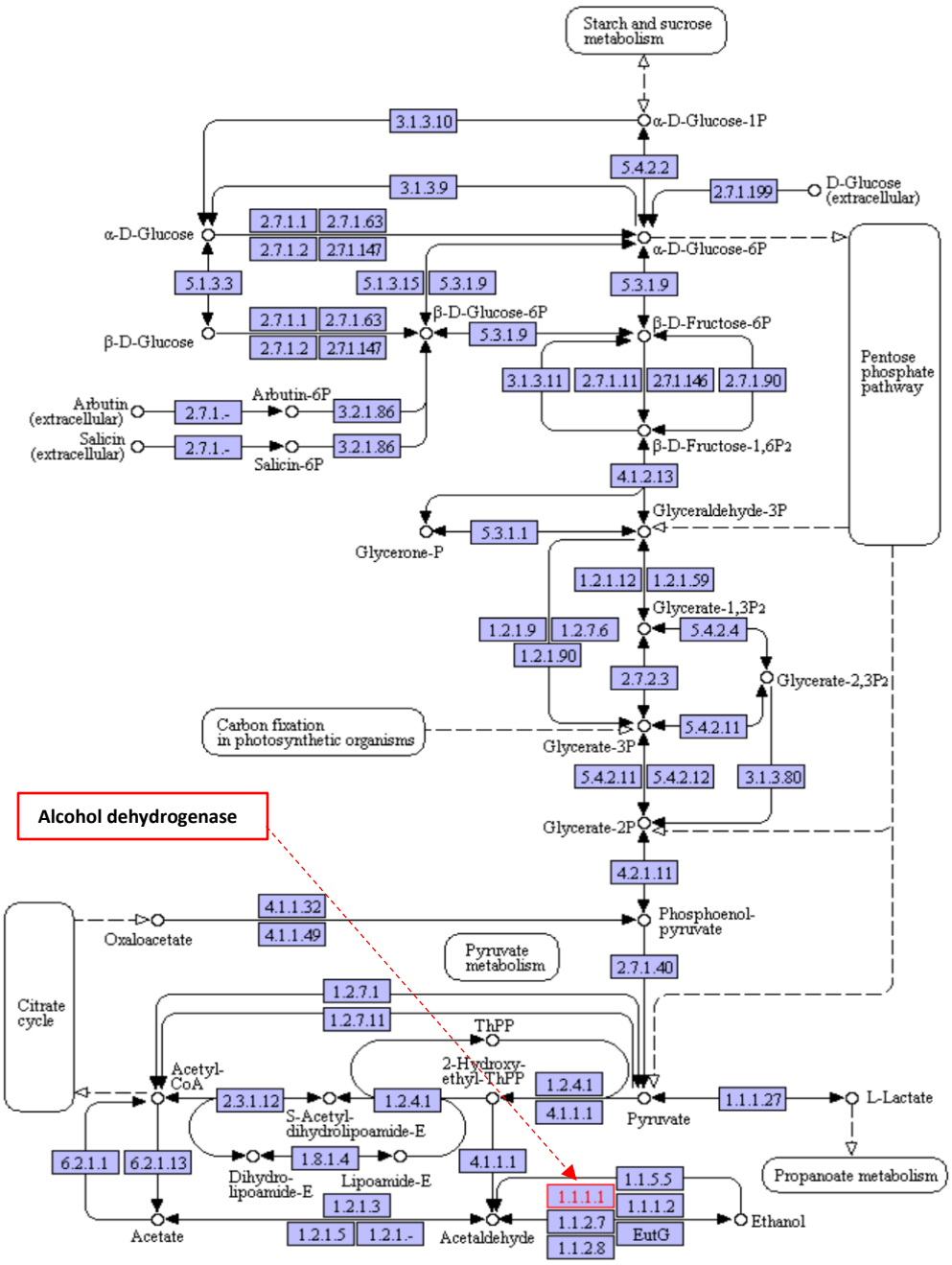


Figure 1-1: Glycolysis/gluconeogenesis pathway, as reported in KEGG database

The pathway shows the enzymes involved (purple numbered boxes) and the products formed during starch and sucrose metabolism. One of the products is ethanol which is produced by reducing acetaldehyde. The enzyme involved in this reduction (1.1.1.1, denoted by the broken red arrow) is alcohol dehydrogenase (Kanehisa and Goto, 2000).

1.3 Alcohol dehydrogenase in hyperthermophiles

One of the most interesting groups of enzymes from hyperthermophiles are the ADHs (Radianingtyas and Wright, 2003). ADHs belongs to the oxidoreductase super family and ubiquitous to all three domains of life (Bacteria, Archaea, and Eucarya) (Radianingtyas and Wright, 2003; Ma and Tse, 2015; Asokumar et al., 2018). It is a key enzyme in the alcohol fermentation pathway, producing alcohol such as ethanol, one of the most sort-after renewable energy sources. In general, ADHs catalyse the interconversion between alcohols and the corresponding aldehydes and ketones (Ma and Tse, 2015; Asokumar et al., 2018). They are considered an interesting and intriguing set of enzymes because depending on the source their physical and enzymatic properties and substrate specificities can vary widely (Radianingtyas and Wright, 2003; Ma and Tse, 2015). They can react with primary and secondary, linear and branched, aliphatic and aromatic alcohols and with their corresponding aldehydes and ketones.

Studying and understanding ADHs from hyperthermophilic microorganisms have been in focus recently. Characterizing thermostable ADHs would aid in better understanding of activity under high temperature and mechanisms of thermostability.

ADHs can be broadly categorized into three groups, depending on electron carriers required; (i) NAD or NADP (ii) pyrollo-quinoline quinone, heme or cofactor F_{420} and (iii) flavin adenine dinucleotide (FAD) (Reid and Fewson, 1994; Littlechild et al., 2003; Asokumar et al., 2018). The NAD or NADP dependent ADHs can be further classified based on their molecular size and metal requirements; (a) Type I, also known as medium chain ADHs, contain approximately 370 amino acid residues per chain, and contain zinc. (b) Type II contain short chain ADHs that do not contain metal ions and have approximate lengths of 250 amino acid residues per chain. (c) Type III

contains iron-dependent long chain ADHs which range from 380-900 amino acid residues per chain (Reid and Fewson, 1994; Ma and Tse, 2015; Asokumar et al., 2018).

1.3.1 Zinc-containing ADHs

Zinc-containing ADHs are classified as type I ADHs and mostly exist as tetramers or dimers. The dimeric forms are usually found in higher eukaryotes while the tetrameric forms are found in prokaryotes. Some exceptions do exist. For example, ADHs characterized from *Rhodococcus ruber* DSM 44541 (Kosjek et al., 2004), *Thermomicrobium roseum* (Yoon et al., 2002), *Thermococcus sibiricus*, and *P. furiosus* (Ma et al., 1999) were found to be dimer while an ADH belonging to *Picrophilus torridus* was found to be a dodecamer (Hess et al., 2008).

Zinc-containing ADHs usually, but not always, have two tetrahedrally coordinated zinc ions per subunit; one for catalytic function, found in the active site, and the other for structural stabilization (Auld and Bergman, 2008; Baker et al., 2009). The catalytic zinc ion interacts with four ligands which can be a combination of Cys, His, Asp, and Glu in microbial ADHs (Korkhin et al., 1998). For example, in the ADH of *Sulfolobus solfataricus* the catalytic zinc interacts with two Cys, a His, and a Glu residue (Esposito et al., 2002) while The ADH of *Aeropyrum pernix*, the zinc ion interacts with a Cys, His, Asp, and a Glu residue (Guy et al., 2003).

The structural zinc ion also interacts with four amino acid residues in a tetrahedral fashion. In higher eukaryotes and bacteria these four ligands would be Cys residues. In Horse liver ADH structural zinc ion coordinates with Cys97, Cys100, Cys103, and Cys111.

The structural zinc ion in *Pyrobaculum aerophilum* coordinates with four Cys residues while in *S. solfataricus* three Cys residues and a Glu coordinates with zinc (Esposito et al., 2002; Vitale et al., 2013). In *A. pernix* ADH the zinc ion coordinates with three Cys and an Asp residue (Guy et al.,

2003). Some hyperthermophilic zinc-containing ADHs only have one zinc that either plays catalytic or structural role. ADHs from *T. brockii*, *T. ethanolicus*, and *T. guaymasensis* have been found to have only the catalytic zinc ion while some others such as the one from *P. aerophilum* has only the structural zinc (Ying and Ma, 2011; Vitale et al., 2013).

Zinc-containing ADHs are nicotinamide cofactor-binding protein, which use NAD^+ or $NADP^+$ as cofactors. The specificity for using nicotinamide cofactors is brought by certain highly conserved domains. One such structure found in all zinc-containing ADHs is the Rossmann fold domain which is composed of alternating beta strands and alpha helical segments. This motif, annotated as $GXGX_2G$ (X denotes undetermined amino acid), can be seen in all type I ADHs (Tse et al., 2017). The specificity of using NAD or NADP is conferred by other amino acid residues present in the chain. In type I ADHs the presence of an Asp residue at a certain position equivalent to Asp223 in HlADH makes the enzyme NAD(H) specific. The Asp residue binds to the adenosine ribose of NAD^+ via two hydrogen bonds with the hydroxyl groups (Ma and Tse, 2016). In NADP-dependent type I ADHs a Gly, Ser, Arg, and sometimes a Tyr residue can be found approximately at the position of the Asp residue as in an NAD specific enzyme (Ma and Tse, 2016). This can be seen in the ADH of *T. guaymasensis* (Ying and Ma, 2011).

Usually zinc-containing ADHs are active under aerobic conditions since the Zn^{2+} ion cannot be further oxidized. However, there are exceptions. The ADH of hyperthermophilic archaea *T. guaymasensis* is oxygen sensitive even though it's a zinc containing enzyme (Ying and Ma, 2011). The only other well-known example of oxygen inactivation is the zinc-containing ADH from mesophilic *S. cerevisiae*, whose inactivation was due to oxidation of free -SH groups (Buhnur and Sund, 1969).

1.3.2 Iron-dependent ADHs

The iron-dependent ADHs are classified as type III ADHs. The well studied example of type III ADHs is the mesophilic *Z. mobilis* ADH 2 (ZmADH2) (Kinoshita et al., 1985). Structural studies of ZmADH2 showed that the iron ion coordinates with three highly conserved His residues and one acidic residue such as Asp (Moon et al., 2011). Sequence and structural studies of all type III hyperthermophilic ADHs have showed the presence of the three highly conserved His residues coordinating with the metal ion (Moon et al., 2011; Larson at el., 2019). The iron ion has catalytic and/or structural functions. Type III ADHs are always oxygen sensitive since the Fe^{2+} ion oxidizes to Fe^{3+} under oxidative conditions (Ma and Tse, 2016). It should be noted that mesophilic ADHs of this group do not contain iron when purified, but require it externally for catalytic activity. In contrast, the hyper/thermophilic ADHs of this group contain iron and exhibit activity after purification (Ma and Adams, 1999).

In type III ADHs, recognition of cofactor such as NAD^+ or $NADP^+$ is brought by a highly conserved glycine-rich motif equivalent to Gly96-Ser99 in ZmADH2 (Moon et al., 2011). The specificity between the cofactors is confirmed by many residues, but one specific residue, an Asp residue at the 39^{th} position in ZmADH2, has been identified to be crucial in spatial determination of accepting NAD^+ over $NADP^+$ and vice versa (Moon et al., 2011). All type III hyperthermophilic ADHs characterized so far are $NADP^+$ dependent and have a Gly instead of the Asp residue. The Gly residue provides more space and flexibility for the accommodation of the phosphate group in $NADP^+$ (Montella et al., 2005). So far only 6 hyperthermophilic iron-containing ADHs have been characterized and they all belong to the genus *Thermococcus* (See Appendix).

1.3.3 Native and recombinant ADH of *Thermococcus guaymasensis*

The euryarchaeal order *Thermococcales* can be divided into three genera; *Thermococcus*, *Pyrococcus* and *Palaeococcus*. Many species from these genera, such as *P. furiosus*, *P. horikoshii*, *Pyrococcus abys*ii, *T. kodakaraensis*, and *Palaeococcus ferrophilus* are well studied, while new species are being identified under extreme conditions. According to German Collection of Microorganisms and Cell Cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen, DSMZ), the genus *Thermococcus* contains 35 members. Most of the *Thermococcus* species are hyperthermophiles and are usually spherical and obligately anaerobic.

T. guaymasensis is a hyperthermophilic archaeon that was first isolated from the Guaymas Basin hydrothermal vent site (Canganella et al., 1998). It is a strictly anaerobic chemoorganotroph that prefers proteinaceous substrates such as yeast extract and trypticase but also utilizes casein, dextrose, maltose and starch. It produces acetate, propionate, isobutyrate, isovalerate, CO_2 and H_2S (in the presence of elemental sulfur) as metabolites (Canganella et al., 1998). In 2011, Ying and Ma showed that *T. guaymasensis* produces ethanol and acetoin via glucose fermentation and characterized the ADH (TgADH) which catalyzes the ethanol production. This ADH was found to be a zinc containing homotetrameric enzyme of 364 amino acids per subunit (39,395.72 Da) and specific for NADP as coenzyme. TgADH was found to be active over a broad temperature range (30°C – 95°C) while the optimal temperature was over 95°C with a half-life ($t_{1/2}$) of approximately 24 hours. This makes it the most thermostable zinc-containing ADH known to date (Hao, unpublished; Ying and Ma, 2011), with a high specific activity of 1,149 U/mg, compared to other zinc-containing ADHs. The enzyme also has broad substrate specificity. It oxidizes various primary and secondary alcohols, polyols, and diols while reducing various aldehydes and ketones. Moreover, it's unique stereoselectiveness catalyzes the reduction of racemic (R/S)-acetoin to (2R,

3R)-2,3-butanedial and meso-2,3-butanediol. Furthermore, TgADH is also highly solvent tolerant and remains active in the presence of methanol up to 40 % (v/v) (Ying and Ma, 2011).

The TgADH encoding gene was cloned and overexpressed in *E. coli* (Hao, unpublished). The entire encoding gene, with *EcoRI* and *XhoI* sites at the N- and C-terminal, respectively, was first inserted to the overhanging thymidine region of a pGEM-T Easy vector and transformed to *E. coli* DH5α cells (Hao, unpublished). After confirmation by sequencing, the insert from pGEM-T Easy vector was released by *EcoRI* digestion (Figure 1-2). The entire TgADH encoding gene, with overhanging primer including *EcoRI* and *XhoI* restriction sites, was then inserted to a *EcoRI* and *XhoI* double digested pET-30a expression vector, driven by the T7 lac promotor (Figure 1-2) (Hao, unpublished) and cloned and overexpressed in *E. coli* BL21 cells. The recombinant TgADH has 417 amino acids with a molecular mass of 45,260.13 Da, per subunit, (Figure 1-3) and very similar catalytic properties as the native enzyme (Hao, unpublished).

Regardless of being a zinc-containing ADH, both native and recombinant TgADH were found to be oxygen sensitive and lose activity when exposed to aerobic conditions (Ying and Ma, 2011). The half-life ($t_{1/2}$) against oxygen inactivation was about 4 hours, and loss of activity was slightly protected in the presence of reducing agents (Hao, unpublished). TgADH has four Cys residues per subunit (Cys39, Cys56, Cys212, and Cys305) and is hypothesized that the oxidation of -SH groups of one or more of these Cys residues may be the reason for oxygen sensitivity (Hao, unpublished; Ying and Ma, 2011).

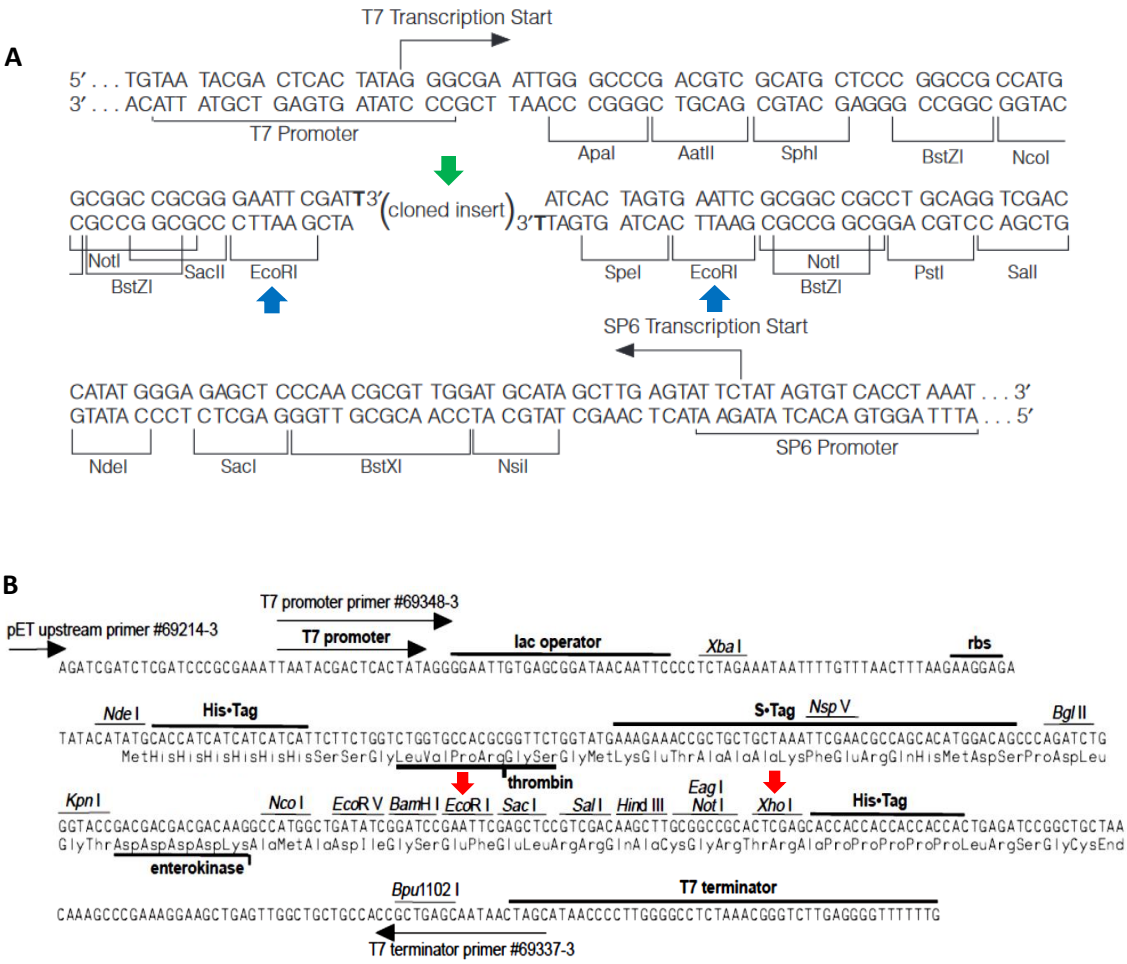


Figure 1-2: Maps of pGEM-T Easy and pET-30a vectors showing the insertion site of TgADH encoding gene

The TgADH encoding gene was (A) first inserted to the overhanging thymidine region of pGEM-T Easy vector (denoted by green arrow) and released by *EcoRI* digestion (restriction sites denoted by blue arrows). The insert was then (B) cloned to a pET-30a vector that was double digested using *EcoRI* and *XhoI* (restriction sites denoted by red arrows) (Adapted from EMD Millipore, 1998; Promega, 2018)

```
        10         20         30         40         50         60
MHHHHHHSSG LVPRGSGMKE TAAAKFERQH MDSPDLGTDD DDKAMADIGS EFMSKMRGFA

        70         80         90        100        110        120
MVDFGKAEWI EKERPKPGPY DAIVKPIAVA PCTSDIHTVF EAAFPREMCE FPRILGHEAV

       130        140        150        160        170        180
GEVVEVGSHV KDFKPGDRVV VPAITPDWRT LDVQRGYHQH SGGMLAGWKF SNPLKEGGKD

       190        200        210        220        230        240
GVFAEYFHVN DADMNLAHLP DEIKPEVAVM ATDMMTTGFH GAELADIPLG GTVAVIGIGP

       250        260        270        280        290        300
VGLMAVAGAR LLGAGRIIAV GSRPVCVEAA KYYGATDIVN RREHPDIAGR ILELTGGEGV

       310        320        330        340        350        360
DSVIIAGGNV DVMKTAVKIV KPGGTVANIN YFGSGDYLPI PRIEWGQGMA HKTIKGGLCP

       370        380        390        400        410
GGRLRMERLL DLIKYGRVDP SRLITHKFKG FDKIPEALYL MKDKPKDLIK PVVIIEE*
```

Figure 1-3: Amino acid sequence of recombinant wild-type TgADH

Recombinant wild-type TgADH has 417 amino acid residues (including the N-terminal His tag) and is 45,260.13 Da per subunit. The residues highlighted in yellow belong to the vector pET-30a. The residue underlined is the first residue of the native TgADH sequence. Asterisk denotes the end of the sequence.

1.4 Inactivation of enzymes containing reactive Cys residue(s)

Oxidation is a common modification of proteins. No universal mechanism exists for enzyme inactivation due to oxidizing factors. Some enzymes can still retain activity even after significant changes in the native confirmation of the molecule, while others get inactivated without visible conformational changes (Jin et al., 2004; Rodacka et al., 2016). The mechanism of inactivation depends mainly on the amino acids involved in the catalytic reaction. Among all the amino acids tryptophan, cysteine, tyrosine, histidine, and methionine residues are highly susceptible to oxidative attack by reactive oxygen species (Rodacka et al., 2016). Although these residues usually represent minor components in proteins, they often play a key role in the catalytic sites or are located near catalytic sites of enzymes and in stabilization of protein structures (Rodacka et al., 2016).

Enzymes with cysteine residues that are directly involved in the catalytic reaction are the most susceptible to oxidative inactivation because of sulfur atoms on the surface (Rodacka et al., 2016). Cysteine is one of the least abundant residues among all other amino acids in proteins. It makes about 2-2.6% in mammalian proteins and 0.4-0.5 % in Archaea (Miseta and Csutora, 2000; Gupta and Carroll, 2014). Nevertheless, up to 80 % of Cys play various functional roles, starting from catalytic activity to structural stabilization (disulfide bridges and metal binding) and regulatory activity (Marino and Gladyshev, 2010; Go et al., 2015). Depending on the pKa values in the physiological pH, Cys residues may appear as negatively charged thiolate anions, which are highly nucleophilic than the protonated forms and can be easily oxidized (Rodacka et al., 2016). Under oxidative conditions the free thiol group of a Cys residue can go through different stages of oxidation, depending on the type of oxidant, amount of oxidant, distance between other free thiols, and the environment surrounding the Cys residue (Figure 1-4).

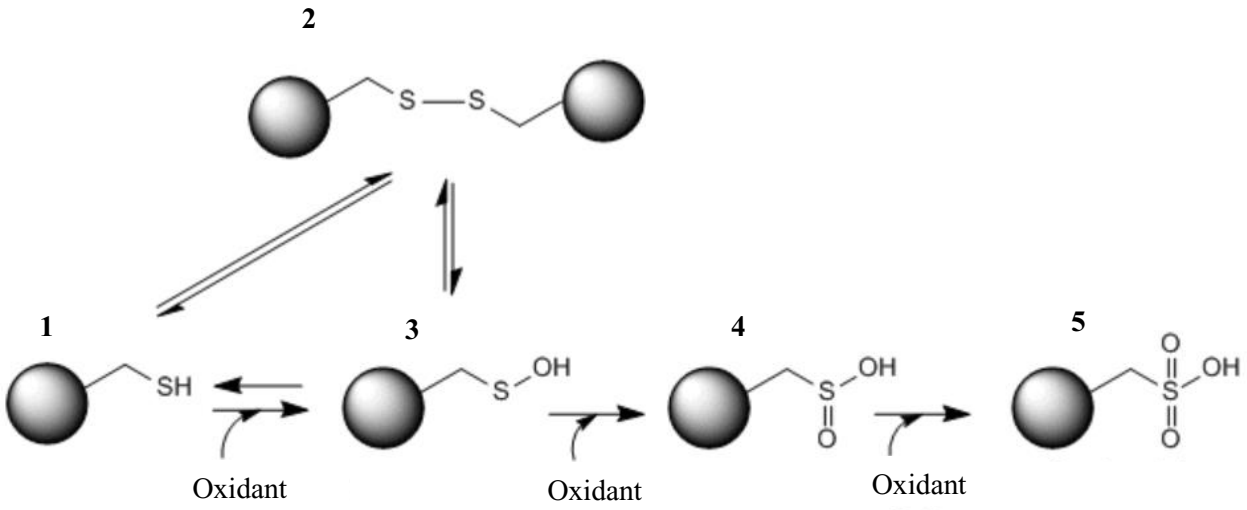


Figure 1-4: Sequential oxidation states of cysteine

1, Cysteine; 2, Disulfide formation (reversible under reducing conditions); 3, Sulfenic acid (Addition of one oxygen, reversible under reducing conditions); 4, Sulfinc acid (Addition of two oxygens); 5, Sulfonic acid (Addition of three oxygens). (Adapted and modified from Jakubowski, 2016)

Many prokaryotic as well as eukaryotic enzymes have been shown to be inactivated due to oxidation of Cys thiolates at the catalytic sites or that take part in the catalytic reaction (Table 1-1). For example, human peroxiredoxin I, a peroxidase that reduces hydrogen peroxide and alkyl hydroperoxides to water and alcohol, was found to be inactivated due to selective oxidation of the catalytic site Cys_{51}-SH to Cys_{51}-SO_2H (Yang et al., 2002). Glyceraldehyde-3-phosphate dehydrogenase plays a vital role in the glycolytic pathway. Catalytic activity of this enzyme depends on three amino acid residues, cysteine residue at position 152, histidine residue at position 179 and arginine residue at position 234. Studies showed that Cys152 has low pKa=6.03 and is in a highly reactive nucleophilic thiolate state at physiological pH (Martyniuk et al., 2011). The oxidation of Cys152 rapidly inactivates this enzyme. The sulfur in the thiol group of Cys152 gets subsequently oxidized, starting from sulfenic acid (R-SOH), to sulfinic acid (R-SO_2H) and sulfonic acid (R-SO_3H) (Morgan et al., 2002; Kowalczyk et al., 2008).

Zinc containing proteins have been found to be highly susceptible to oxidative stress because of the zinc ion being coordinated with Cys residue(s). For example, Liver betaine-homocysteine S-methyl transferase undergoes inactivation when not exposed to reducing agents (Castro et al., 2008). Further analysis showed that two of the three Cys residues that coordinate with the catalytic zinc form a disulfide bond under reducing agent-free conditions, causing complete loss of activity (Castro et al., 2008). Other zinc containing enzymes such as protein kinase C, histone acetyltransferases and DNA/RNA binding polymerases have also been shown to be inactivated under oxidative conditions (Rodacka, 2016).

Type I ADHs are zinc containing proteins as well, but not many of them have been characterized as becoming inactivated due to oxidative stress. One well studied example is YADH I. YADH I is a homotetrameric protein with a subunit molecular mass of 36 kDa (Buhnur and Sund, 1969; Men

and Wang, 2006; Rodacka, 2016). Each subunit contains two zinc ions. The catalytic zinc ion is coordinated by two cysteines and one histidine residue while the fourth coordination position is occupied by a hydroxyl group of ethanol, which is replaced with water in the absence of substrate. The structural zinc ion is coordinated by four cysteine residues. Buhnur and Sund (1969) observed that when YADH I crystals were stored in ammonium sulfate at 4°C, the activity declined steadily over time. The loss of activity was found to be due to the oxidation of free-SH groups (Buhnur and Sund, 1969). The -SH groups get oxidized to disulfide groups and a linear relationship was observed with decreasing enzymatic activity and the numbers of free thiol groups. *In vitro* oxidative studies using hydrogen peroxide showed that Cys43 and Cys153 at the active site get oxidized to cysteine sulfinic and sulfonic acid, respectively leading to the loss of zinc (Men and Wang, 2006). Furthermore, three disulfide bond formation were seen between Cys43-Cys153 (in the catalytic center), Cys103-Cys111 and Cys276-Cys277 (Men and Wang, 2006).

Table 1-1: Selection of enzymes that that are known to be inactivated due to Cys-thiol oxidation

Enzyme	Oxidant	Half-life ($t_{1/2}$, minutes)	Affected Cys residue(s)	Oxidative modification(s) of affected Cys residue(s)	Reference
Arabidopsis thaliana ADH	H_2O_2	80	Cys_{47} and $Cys2_{43}$	Cys_{47}-S=S-Cys_{243}	Dumont et al., 2018
Saccharomyces cerevisiae ADH	H_2O_2	nd	Cys_{43} and Cys_{153}	Cys_{43}-SO_2H and Cys_{153}-SO_3H	Men and Wang, 2006
Saccharomyces cerevisiae ADH	Air	nd	nd	nd	Buhnur and Sund, 1969
Trigonopsis variabilis D-amino acid oxidase	HOCl	60	Cys_{108}	Cys_{108}-SO_2H	Mueller et al., 2010
Glyderaldehyde-3-phospahte dehydrogenase	H_2O_2	>30	Cys_{152}	Cys_{152}-S=S-Cys_{156} Cys_{152}-SOH → Cys_{152}-SO_2H → Cys_{152}-SO_3H	Morgan et al., 2002
Human peroxiredoxin I	H_2O_2	0.5	Cys_{51}	Cys_{51}-SOH → Cys_{51}-SO_2H	Yang et al., 2002
Liver betaine-homocysteine S-methyltransferase	na	na	Cys_{217} and Cys_{299}	Cys_{217}-S=S-Cys_{299}	Castro et al., 2008

nd=not determined; na=not applicable

1.5 MALDI-TOF Mass Spectrometry

Mass spectrometry is an analytical method used to measure the ratio of mass to charge (m/z) by ionizing chemical compounds into charged molecules (Singhal et al., 2015). Even though mass spectrometry was invented for analysis of chemical compounds, development of electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) made it possible to apply the technique to analyze large biological molecules like proteins (Signor and Erba, 2013; Singhal et al., 2015). Both the techniques have advanced by leaps and bounds, enabling accurate mass determination of proteins to evaluate protein primary sequence, presence of mutations and/or post-translational modifications and protein degradation (Signor and Erba, 2013). Both ESI and MALDI convert peptides into ions by either addition or loss of one or more protons. MALDI-TOF MS has certain advantages over ESI-MS; (i) MALDI-TOF MS produces singly charged ions, thus interpretation of data is easy compared to ESI-MS, (ii) prior purification/separation of protein is required for ESI-MS analysis, but not mandatory for MALDI-TOF MS analysis, (iii) efficiency of ESI-MS is affected by the presence of salt, detergents, and contaminants, whereas MALDI-TOF MS has better salt tolerance (Signor and Erba, 2013; Singhal et al., 2015). Moreover, the simplicity and speed of data acquisition and interpretation of MALDI-TOF MS makes it the obvious choice for large scale proteomics work.

The sample for analysis by MALDI MS is prepared by mixing or coating with matrix, a solution of an energy-absorbent organic compound (Karas and Hillenkamp, 1988; Signor and Erba, 2013; Singhal et al., 2015). For MS examination of intact protein, a single organic matrix such as sinapinic acid (3,5-Dimethoxy-4-hydroxycinnamic acid), α-cyano-4-hydroxycinnamic acid (α-CHCA) or dihydroxybenzoic acid is used. The choice of matrix and technique used for matrix deposition (e.g. dried droplet and thin layer) influence the quality of the MALDI spectra. When

the matrix dries and crystalizes, the sample within the matrix also co-crystallizes. The sample within the matrix is then ionized with a laser beam (Signor and Erba, 2013; Singhal et al., 2015). Desorption and ionization generate singly protonated ions from the sample. The protonated ions are then accelerated at a fixed potential, where they separate from each other based on their mass-to-change ratio (m/z) (Signor and Erba, 2013; Singhal et al., 2015). The charged analytes are then detected and measured using different types mass analyzers such as quaderupole mass analyzers, ion trap analyzers, time of flight (TOF) etc (Signor and Erba, 2013; Singhal et al., 2015). During MALDI-TOF analysis, the m/z of an ion is measured by determining the time it takes to travel the length of the flight tube. Some TOF analysers use a mirror at the rear end which reflects back the ions, thereby increasing the length of the flight tube and correcting small differences in energy among ions (Signor and Erba, 2013; Singhal et al., 2015). Based on the TOF information, a spectrum called peptide mass fingerprint is generated for the analytes in the sample.

Accuracy and resolution of a mass spectrometry method should be considered for the desired purpose. MALDI-TOF is usually chosen for the analysis of large proteins since it has an accuracy of 0.05-0.2 %, resolution of 2000 and a m/z range of >300,000 (Lewis et al., 2000). The accuracy can be further increased if the protein is digested into peptides before analysis. The accuracy and resolution of MALDI-TOF depends on the sample and sample preparation (Lewis et al., 2000). Acquiring optimum data depends on choosing suitable matrices and solvents, the functional and structural properties of the analyte, sample purity, and how the sample is applied on the MALDI sample plate. Signor and Erba (2013) showed that using a mixture of DAB and α-CHCA increased the resolution compared to using sinapinic acid, when analyzing 20 pmoles of intact beta-galactosidase. However, only sinapinic acid was used for the analysis of large plasma proteins, by

Horton and Remaley (2006). This shows that the choice of matrix depends on the characteristics of the protein and is not transferable.

MALDI-TOF has been used successfully for analyzing large intact proteins to peptide fragments. For example, Signor and Erba (2013), showed that they were able to analyse proteins over 100 kDa using MALDI-TOF. Horton and Remaley (2006) reported that they were able to analyze major proteins in blood plasma, of which some were over 150,000 Da along with post translational modifications. Post-translational modifications such as phosphorylation, glycosylation, cysteine redox modifications, and nitrosylation have also been studied using MALDI-TOF mass spectrometry. Almost all the studies on analysis and characterisation of protein post-translational modifications are done using a 'bottom-up' approach, where the intact protein is first enzymatically digested and the peptide fragments are analysed using MALDI-TOF. This approach increases the accuracy of MALDI-TOF up to 0.1 Da – 0.01 Da (Mann and Kelleher, 2008). However, this approach has its drawbacks. When subjecting peptide fragments to mass spectrometry, only a small fraction of the tryptic peptides are normally detected, and only a fraction of these yield useful data that could help build a whole picture. Moreover, enzymatic digestion of a protein can sometimes introduce or abolish certain changes, which could cause unexpected increase or decrease in measurements (Liu et al., 2007).

1.6 Aims of this study

Enzymes from hyperthermophilic microorganisms serve as model systems in understanding enzyme evolution, molecular mechanisms for protein thermostability, and adaptation of enzymes to their unusual environments. ADHs from hyperthermophiles are of great interest due to their thermostability, high activity, being able to use complex carbohydrates as substrates, and enantioselectivity, making them potential industrial biocatalysts. One such zinc-containing homotetrameric ADH from hyperthermophilic archaeon *T. guaymasensis* was found to lose its activity when exposed to air. It was hypothesized that the loss of activity could be due to the oxidative damage of Cys residue(s) in each subunit. The main focus of this study was to investigate the oxidative modifications in recombinant wild-type TgADH and deduce the oxygen sensitivity of two different Cys single mutants of TgADH. Analyzing the modifications due to oxidation will help in estimating the number of oxygen atoms added under oxidative conditions and the oxidation state of the free thiol groups of the Cys residues.

Oxygen sensitivity of two different Cys single mutants of TgADH was analyzed. Enzyme activity analysis of mutants with either one of the mutations, C212S or C305S, under oxidative conditions will help in understanding the role of the Cys residues behind the loss of enzyme activity due to oxidation.

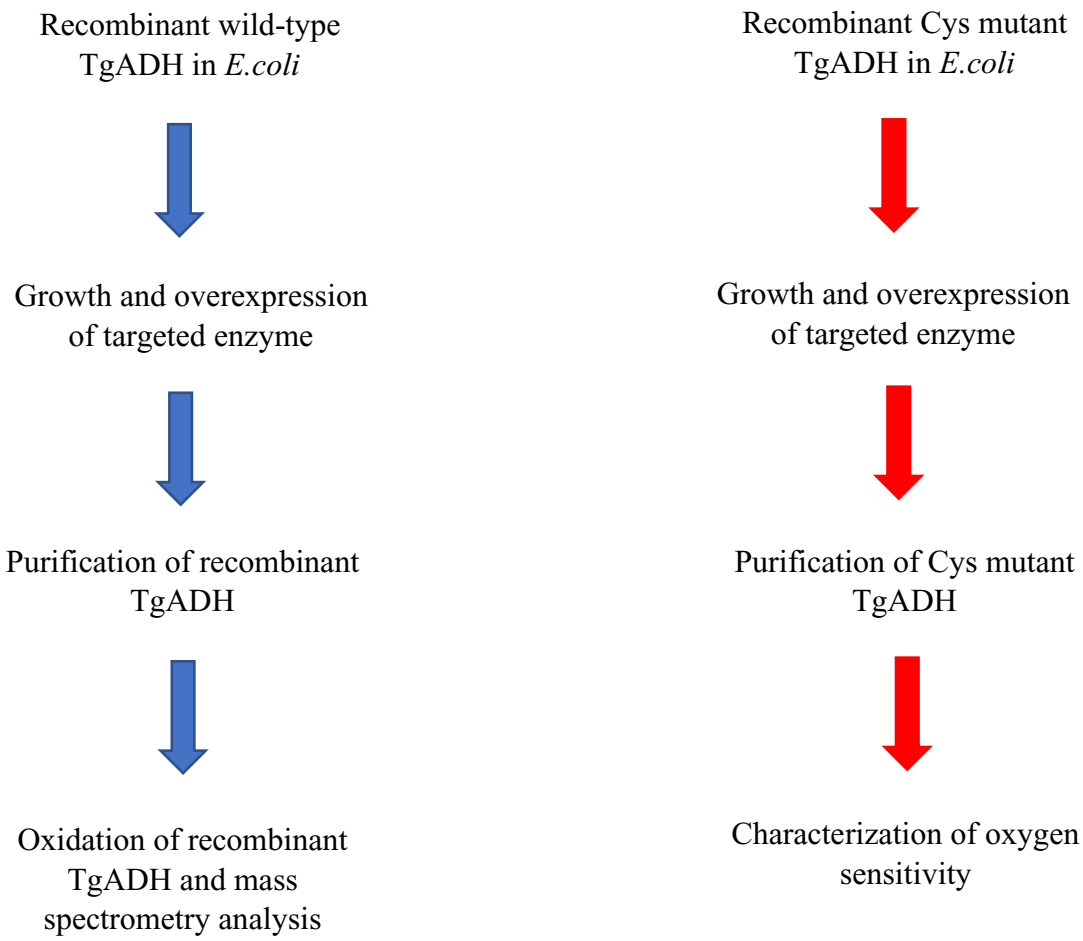


Figure 1-5: Brief strategy of this study

2. Materials & Methods

2.1 Chemicals used

All the chemicals in this research were purchased from commercially available sources

Table 2-1: Chemicals used

Chemical	Supplier
Acryl/Bis solution (29:1), 30% (w/v)	Bio Basic Inc. (ON, Canada)
Ammonium persulfate	Bio Basic Inc. (ON, Canada)
Tetramethylethylenediamine	Bio Basic Inc. (ON, Canada)
Tris base	Fisher Scientific (ON, Canada)
N-cyclohexyl-3-aminopropanesulfonic acid	Bioshop (ON, Canada)
Dithiothreitol	Bio Basic Inc. (ON, Canada)
Sodium dithionite	Unspecified
Lysozyme	Bio Basic Inc. (ON, Canada)
DNase I	Roche (ON, Canada)
ß- nicotinamide adenine dinucleotide (NADP)	Roche (ON, Canada)
2-Butanol	Sigma Aldrich (ON, Canada)
Glacial acetic acid (99%)	Sigma Aldrich (ON, Canada)
Methanol	Sigma Aldrich (ON, Canada)
Glycerol	Fisher Bioreagents (ON, Canada)
Isopropyl β-D-1-thiogalactopyranoside	Calbiochem (ON, Canada)
Sodium dodecyl sulfate	Fisher Scientific (ON, Canada)
Coomassie Brilliant Blue R250	EM Science (ON, Canada)
Hydrogen peroxide (30% v/v)	Caledon Laboratory Chemicals (ON, Canada)
Ammonium sulfate	VWR Life Sciences (ON, Canada)
Yeast extract	Bio Basic Inc. (ON, Canada)
Tryptone	Bio Basic Inc. (ON, Canada)
Sodium chloride	Bio Basic Inc. (ON, Canada)
Tryptic soy broth	Becton, Dickinson and Company (MD, USA)

2.2 Instruments used

Table 2-2: Instruments used

Instrument	Manufacturer
Centrifuge (Sorvall RC6-Refrigerated Superspeed Centrifuges)	Thermo Fisher Scientific (Canada)
FPLC	Amersham Biotech (QC, Canada)
C24 incubator shaker	New Brunswick Science (NJ, USA)
Mini-Protean III SDS-PAGE system	Bio-Rad Laboratories, Inc. (ON, Canada)
Spectrophotometer (GENESYS 10 UV)	VWR Canlab (ON, Canada)
Microcentrifuge	Eppendorf (ON, Canada)
Vortex (Mini vortexer)	VWR Canlab (ON, Canada)
Waterbath (Isotemp 215)	Fisher scientific company (ON, Canada)
Centrifuge (Sorvall Legend XTR)	Thermo Fisher Scientific (Canada)
Chart recorder	Servogor
Autoclave (rrA1)	Priorclave (MI, USA)
Chemidoc MP Imaging System	Bio-Rad Laboratories, Inc. (ON, Canada)
Autoflex Speed	Bruker (ON, Canada)

2.3 Microorganism(s)

E. coli BL 21(DE3)-RIL [F^-, *ompT, hsdS*$_B$ (r_B^- m_B^-) *gal dcm lacY1, pRARE (kamR)*] (Stratagene, CA, USA)

2.4 Growth and storage of microorganism(s)

Recombinant *E. coli* was grown in LB media with 10 g/L of tryptone, 5 g/L of yeast extract, and 10 g/L of NaCl. Deionized water was added to a total volume of 1 L and autoclaved for 30 minutes at 121°C. *E. coli* strains were grown in 500 or 1000 ml LB media at 37°C, shaken at 180-200 rpm in shake flasks loosely covered with aluminium foil. The ratio between the volume of the media and the volume of the shake flask was 1:2 (e.g. 1000 ml media was used for cultivation in a shake flask having a volume of 2000 ml). The inoculum for culture was 12-14 colonies from a 1.5 % LB agar plate with selective antibiotics or with a previous culture at late exponential phase in 1% (v/v). The recombinant *Tg*ADH was made by inserting the TgADH encoding gene to a pET-30a vector that was double digested using *EcoRI* and *XhoI*, and driven by the T7 lac promotor (For complete details on construction of recombinant plasmid, refer to Masters dissertation of Liangliang Hao, 2009, University of Waterloo, Canada). Therefore, IPTG was used as the inducer and was added at a final concentration of 0.2 mM when OD_{600} of the culture was between 0.6-0.8. The incubation was then continued for 12-16 hours followed by harvesting the cells by centrifuging the culture at 8,000 rpm, 4°C for 15 minutes.

2 x YT and tryptic soy broth were used to optimize overexpression of recombinant TgADH in *E. coli* (Sambrook et al., 1989). 2 x YT media was prepared using 16 g/L of tryptone, 10 g/L of yeast extract, and 5 g/L NaCl. Deionized water was added to 1 L. The solution was autoclaved for 30 minutes at 121°C. Tryptic soy broth was prepared by suspending 30 g/L of dehydrated media

(Becton, Dickinson and Company, USA) in 1 L of deionized water. The solution was autoclaved for 30 minutes at 121°C. The inoculation, incubation and overexpression methods were the same as when grown in LB medium.

For the recombinant stock of *E. coli*, a single colony was picked from a LB agar plate containing 0.5 mg/ml kanamycin and inoculated in 5 ml LB medium with the same selective antibiotic. It was then incubated at 37°C with vigorous shaking until OD_{600} reached 0.6-0.8. Then an aliquot (0.7 ml) of the bacterial culture was transferred to a sterilized 1.5 ml Eppendorf tube containing 0.3 ml of sterilized 100 % glycerol. After gently mixing the culture, it was stored at -20°C or -80°C for short and long term storage, respectively.

Stock solutions of antibiotics and reagents

Kanamycin: 50 mg/ml in deionized water and filter sterilized through 0.2 μm filter membranes. Stored at – 20°C. Used at a final concentration of 0.5 mg/ml.

IPTG (Isopropyl β-D-1-thiogalactopyranoside): 1 M in deionized water and filter sterilized through 0.2 μm filter membranes. Stored at –20°C. Used at a final concentration of 0.2 mM.

2.5 Enzyme purification

2.5.1 Preparation of cell-free extract

All procedures for the preparation of cell-free extracts were carried out anaerobically. Frozen *E. coli* cells (5 grams, wet weight) carrying the recombinant construction *TgADH-pET-30a* was transferred to a serum bottle and immediately sealed with a grey butyl stopper and an aluminum cap. The bottle was immediately degassed for 30 minutes and pressurized by 3 psi N_2 gas.

Degassed lysis buffer (50 mM Tris buffer containing 5 % (v/v) glycerol, 2 mM dithiothreitol (DTT), 2 mM sodium dithionite (SDT), 0.1 mg/ml lysozyme, and 0.01 mg/ml DNase I, pH 7.8) was anaerobically transferred, using a syringe, into the degassed serum bottle containing the cells. The volume of the anaerobic lysis buffer was about five times the weight of the cells (v/w). The cell suspension was incubated under 180 rpm stirring for 2 hours at 37°C. After centrifugation at 25,000 × g for 30 minutes, the supernatant was collected as cell-free extract and stored anaerobically for further use.

2.5.2 Purification of recombinant *Tg*ADH

Purification of the recombinant enzyme from *E. coli* was carried out anaerobically using the AKTA™ FPLC system. Since the enzyme was thermostable, a step of heat precipitation was applied prior to the liquid chromatography (See section 2.5.3). After incubation the solution turned gel-like. The denatured proteins and cell debris in the cell-crude extract were removed by centrifugation at 25,000 × *g* for 30 min at room temperature.

The supernatant containing the recombinant TgADH was loaded on to a phenyl-sepharose column (2.6 × 10 cm) equilibrated with 1 CV of buffer A (50 mM Tris-HCl at pH 7.8, 5 % (v/v) glycerol, 2 mM DTT, and 2 mM SDT) followed by 1 CV of buffer B (0.8 M ammonium sulfate in buffer A). A decreasing gradient, either segmented (0.82 M – 0.25 M, 1.5 CV followed by 0.25 M – 0.00 M, 2 CV) or linear (0.82 M – 0.00 M, 6 CV), was applied at a flow rate of 2 ml/min and 0.5 MPa pressure limit. The buffers used were filtered using filter paper, degassed and maintained at 3 psi using N_2 gas. The eluted samples were collected anaerobically in degassed serum bottles using a needle.

2.5.3 Heat treatment conditions

All procedures were carried out anaerobically. CFE was prepared as mentioned in section 2.5.1. In order to find the optimal incubation temperature, 1 ml of CFE in 1.8 ml degassed and gassed vial was heat treated at 80°C, 85°C, 90°C, 95°C, and 100°C for 10 minutes. The incubation time was optimised by heat treating 1 ml of CFE for 10, 20, 30, and 40 minutes at optimal temperature. The optimal concentration of the CFE was deduced by heat treating 0.5 ml of CFE at the concentrations of 24.35 mg/ml (undiluted), 15 mg/ml, 10 mg/ml, 5 mg/ml, 2.5 mg/ml, and 0.25 mg/ml, at the optimal temperature and duration. Dilutions were done anaerobically using 50 mM Tris-HCl buffer, pH 7.8, containing 5% (v/v) glycerol, 2 mM dithiothreitol (DTT), and 2 mM sodium dithionite (SDT).

2.6 Determination of protein concentration

Bradford method was used to measure the protein concentration of solutions using a spectrophotometer (Bradford, 1976). Two hundred µl of Bio-Rad reagent was mixed with 800 µl of protein solution (total volume was set to be 1 ml), and a control was set by mixing 200 µl of Bio-Rad reagent with 800 µl deionized water. A linear calibration curve was constructed with known concentrations of standard protein bovine serum (BSA, albumin fraction V) from 1 to 20 mg/ml standard solution. The protein reading was taken at 595 nm based on the λ_{max} of Bio-Rad reagent binding to proteins.

2.7 Determination of enzyme activity

The activity of ADH was determined spectrophotometrically by measuring the rate of production of the cofactor NADPH. The *in vitro* enzyme assays were done anaerobically at 80°C by measuring the substrate-dependent absorbance change of NADPH at 340 nm (ε_{340} = 6.3 $mM^{-1}cm^{-1}$). Unless specifically stated, the enzyme assay was carried out in duplicate using the standard reaction mixture (2 ml) for alcohol oxidation, which contained 100 mM 3-(cyclohexylamino)-1-propanesulfonic acid (CAPS) buffer (pH 10.5), 90 mM 2-butanol, and 0.4 mM NADP. The reaction was initiated by adding enzyme. One unit (U) is defined as the production of 1 μmol of NADPH per minute.

2.7.1 Oxygen sensitivity

The effect of oxygen on enzyme activity was investigated by exposing the enzyme samples to air at room temperature, under stirred conditions, and determining the residual activity after exposure. The exposure was performed in the presence or absence of 2 mM DTT and SDT. The residual activities of each sample at different time intervals were tested parallelly under the standard assay conditions.

2.7.2 Inactivation by H_2O_2

The effect of H_2O_2 was investigated by treating the enzyme samples with 1:10, 1:40, 1:160, 1:400, and 1:1600 mol/mol protein subunit:oxidant. The total protein in an enzyme sample, partially purified or purified, was considered to be TgADH. The samples were incubated for 10 minutes at room temperature and the enzyme activity was measured immediately after.

2.8 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using a Bio-Rad Mini-Protean III system, based on the methods by Laemmli (1970). Protein samples for SDS-PAGE were prepared by incubating at 100°C for 5 minutes in 1x sample loading buffer (0.35 M Tris-HCl pH 6.8, 0.35 M SDS, 30% v/v glycerol, 0.6 M DTT, and 0.175 mM Bromophenol Blue). The protein samples were loaded onto a 12.5 % SDS-PAGE gel and run under constant voltage of 150 V for 80 minutes. The gel was then stained with Coomassie Blue staining solution (0.1 % Coomassie Brilliant Blue R-250, 20% v/v methanol, 10 % v/v glacial acetic acid, topped to 1 L using deionized water) for 30 minutes on shaker at approximately 100-150 rpm. The gel was then de-stained with de-staining solution (15% methanol, 10% glacial acetic acid and 75% deionized water) until protein band(s) started to become visible. The molecular mass of the enzyme under denaturing conditions was estimated using EZ-Run™ Protein Marker (Fisher Scientific, ON, Canada) which contains the following purified proteins; β-galactosidase (116.0 kDa), Bovine serum albumin (66.2 kDa), Ovalbumin (45.0 kDa), Lactate dehydrogenase (35.0 kDa), Restriction endonuclease *Bsp*981 (25.0 kDa), β-lactoglobulin (18.4 kDa), and Lysozyme (14.4 kDa).

2.9 MALDI-TOF mass spectrometry

MALDI–TOF measurements were performed on a Bruker Autoflex Speed instrument equipped with a Nd:YAG UV laser. Positive ion spectra from 15,000–70,000 m/z were acquired in the linear mode with a 20 kV acceleration voltage. Aliquots of the sample solution and a saturated solution of either sinapinic acid (3,5-Dimethoxy-4-hydroxycinnamic acid) or dihydroxybenzoic acid (DHB) matrix in 1:1 deionized H_2O/ACN with 0.1 % TFA (4 µl each) were mixed and air-dried

on a stainless steel MALDI plate. 5,000 scans were accumulated and then processed manually using the flex Analysis software (Version 3.4, Bruker daltonics). The instrument was calibrated with protein standard II (Bruker) which contains a mixture of singly charged species; Trypsinogen (23,982 Da), Protein A (44,613 Da), and Albumin-Bovine (≈ 66.5 kDa) and doubly charged species; Protein A (22,307 Da), and Albumin-Bovine (≈33.3 kDa).

The effect of exposing TgADH to air was investigated by exposing the enzyme sample to air for 16 hours at 4°C. The sample was then mixed with the matrix at a ratio of 1:1 and air-dried on a stainless-steel MALDI plate.

The effect of H_2O_2 was investigated by treating the enzyme samples with 1:10, 1:40, 1:160, 1:400, and 1:1600 mol/mol protein subunit:oxidant. The total protein in an enzyme sample, partially purified or purified, was considered to be TgADH. The samples were incubated for 10 minutes at room temperature, then mixed with the matrix at a ratio of 1:1 and air-dried on a stainless-steel MALDI plate.

2.10 Enzyme structure modelling

The amino acid sequence of wild-type and Cys mutant TgADH was used to find a suitable model template through SWISS-MODEL homology-modelling server (Waterhouse et al., 2018). The modelled structure was downloaded as a PDB file and analysed using UCSF Chimera (Pettersen et al., 2004). The distance between zinc and Cys residue was measured from zinc to the sulfur atom of the Cys residue.

3. Results

3.1 Over-expression of *T. guaymasensis* ADH in *E. coli*

The TgADH encoding gene was inserted to a pET-30a vector that was double digested using *EcoRI* and *XhoI* (Figure 1-2), and overexpressed in *E. coli* BL21 cells (Hao, unpublished). Expression of the enzyme was induced using IPTG, which can be seen from the distinct thick and dark band around 45 kDa, in a 12.5 % SDS-PAGE (Figure 3-1). 0.2 mM IPTG showed approximately the same level of overexpression as 0.5 mM and 1.0 mM (Figure 3-1).

The recombinant *E. coli* cells were incubated in LB medium with 0.5 mg/ml kanamycin, until the optical density at 600 nm wavelength reached 0.8, which took 5 - 6 hours. Then IPTG was added and incubation was continued until OD_{600} of the culture reached ≈1.5, which took 15 - 16 hours. During the small-scale testing, the recombinant *E. coli* cells exhibited optimal levels of expression at 0.2 mM IPTG. Therefore, this concentration was used for large-scale (1 liter) incubation as well.

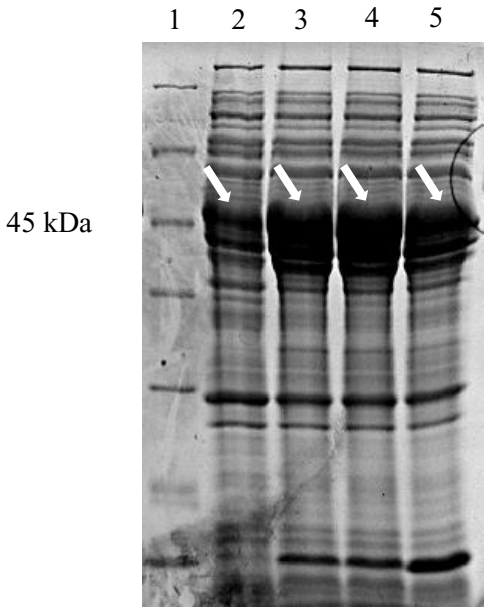


Figure 3-1: Analysis of over-expression level of *Tg*ADH in *E. coli* induced by IPTG, using SDS-PAGE (12.5 %)

Lane 1, Protein marker; Lane 2, No IPTG; Lane 3, 0.2 mM IPTG; Lane 4, 0.5 mM IPTG; Lane 5, 1.0 mM IPTG. White arrow denotes TgADH. 15 µl of CFE was loaded per lane.

3.2 Selection of media for growth and overexpression of recombinant *E. coli* cells

In order to select the type of media that would give optimal levels of overexpression of TgADH, the recombinant cells were grown in three different media broths; 2 x YT, LB, and TSB. The recombinant *E. coli* cells were incubated in respective medium with 0.5 mg/ml kanamycin, until the optical density at 600 nm wavelength reached 0.8, which took 5 - 6 hours. Then IPTG was added to a final concentration of 0.2 mM and incubation was continued until optical density at 600 nm wavelength reached 1.5.

CFE was prepared using the overexpressed cells and protein concentration was measured. Cells cultivated and overexpressed in 2 x YT and LB showed 50 % more protein concentration compared to those grown and overexpressed in TSB (Figure 3-2). Even though 2 x YT is richer than LB, cells grown and overexpressed in the former showed approximately similar levels of overexpression to those in the latter. Therefore, taking into consideration that richer media increases oxidative stress and higher mutation frequencies in *E. coli* cells (Kram and Finkel, 2015), LB was chosen over 2 x YT.

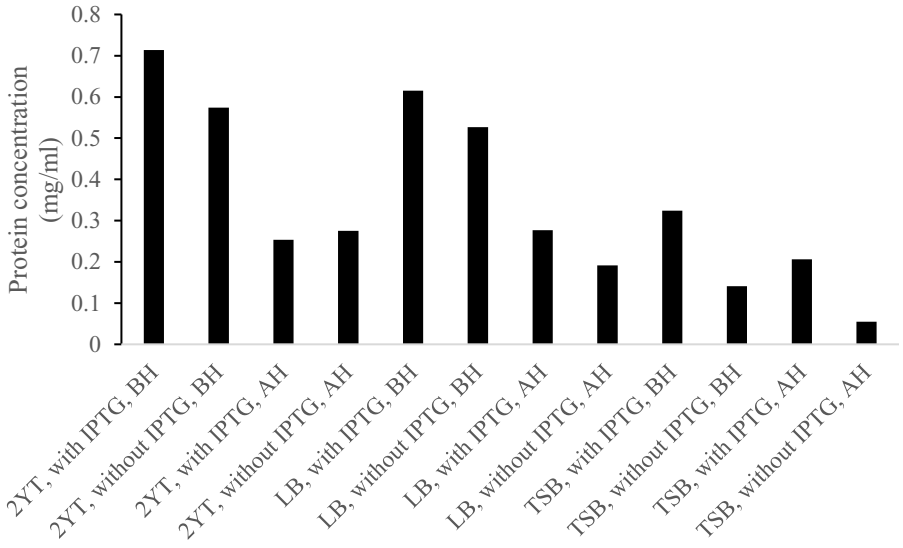


Figure 3-2: Analysis of over-expression of *Tg*ADH in *E. coli*, in 2YT, LB and TSB media, induced by IPTG

Protein concentration (mg/ml) was measured in CFE prepared using cells grown and overexpressed in 100 ml of respective media. 0.2 mM IPTG was used for overexpression. Individual data points were single experiments and not done in replicates.

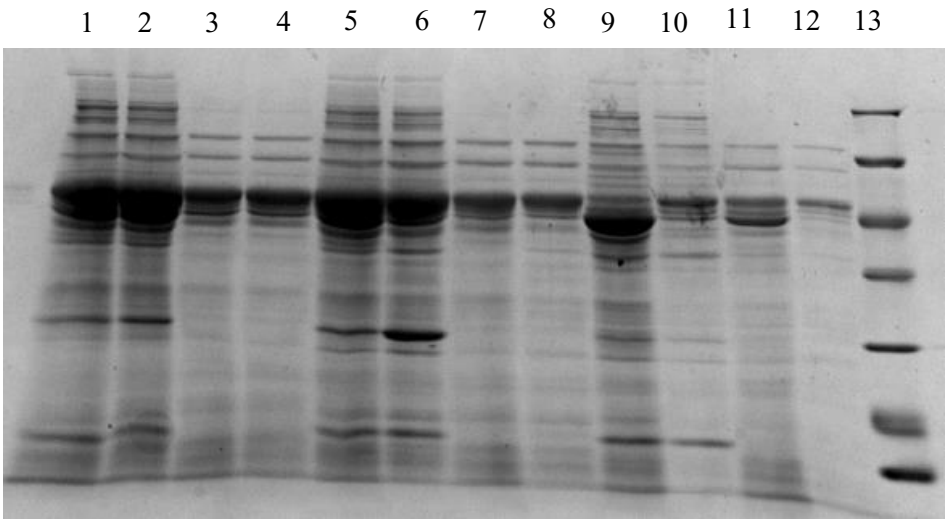


Figure 3-3: Analysis of over-expression of *Tg*ADH in *E. coli*, in 2YT, LB and TSB media, induced by IPTG

A 12.5 % SDS-PAGE was used to show the levels of overexpression in CFE prepared using cells grown in respective media. Lanes 1, 5 and 9: CFE BH, with IPTG, in 2YT, LB and TSB, respectively. Lanes 2, 6 and 10: CFE BH, without IPTG, in 2YT, LB and TSB, respectively. Lanes 3, 7 and 11: CFE AH, with IPTG, in 2YT, LB and TSB, respectively. Lanes 4, 8 and 12: CFE AH, without IPTG, in 2YT, LB and TSB, respectively. Lane 13: Protein marker. 15 µl of CFE was loaded per lane.

3.3 Heat treatment conditions

The CFE prepared using recombinant *E. coli* cells containing *pET-30a-TgADH* construct was subjected to a step of heat treatment before purification using liquid chromatography. The CFE was incubated at 60°C for 1 hour with shaking every fifteen minutes (Hao, unpublished). This only removed approximately 35-40 % of the protein (Figure 3-4) and not 60 – 65 % as reported. Therefore, the heat treatment step was optimized.

3.3.1 Incubation temperature

The optimal temperature was found by incubating at five different temperatures. Approximately 60 % of the unwanted protein in the CFE was removed at 100°C (Figure 3-6). This was the highest compared to the reduction of protein at other temperatures. No loss in enzyme activity was seen at 100°C (Figure 3-5).

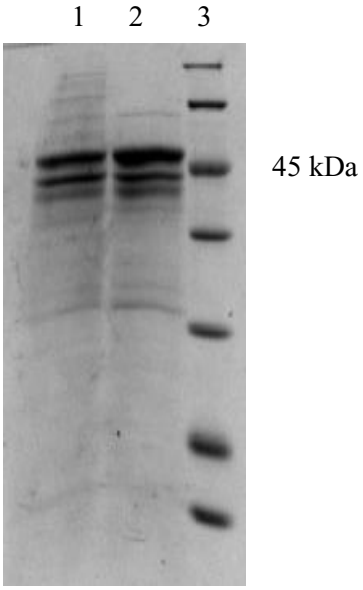


Figure 3-4: Heat treatment of CFE at 60°C for 1 hour

12.5% SDS-PAGE was used to show the difference in protein content after incubation at five different temperatures. Lane 1, CFE BH; Lane 2, CFE AH; Lane 3, Protein marker. 10 µl of CFE was loaded per lane.

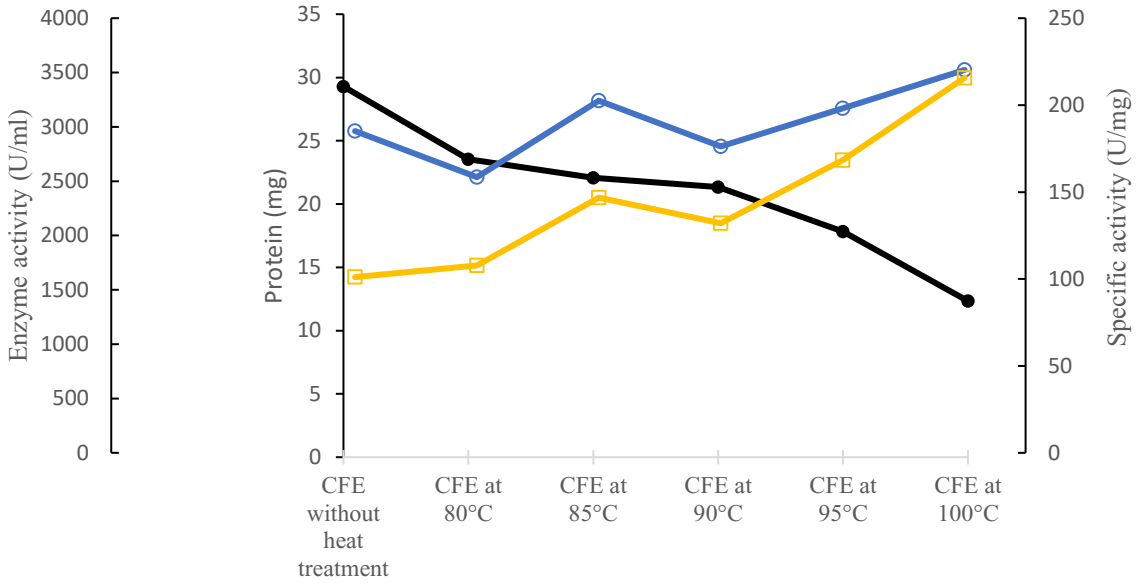


Figure 3-5: Heat treatment of CFE at five different temperatures

CFE was heat treated at 80°C, 85°C, 90°C, 95°C, and 100°C for 10 minutes. Filled circles, protein (mg) of CFE after heat treatment; Open circles, enzyme activity (U/ml) of CFE after heat treatment; Open squares, specific activity of CFE after heat treatment (U/mg). Individual data points were single experiments and not done in replicates.

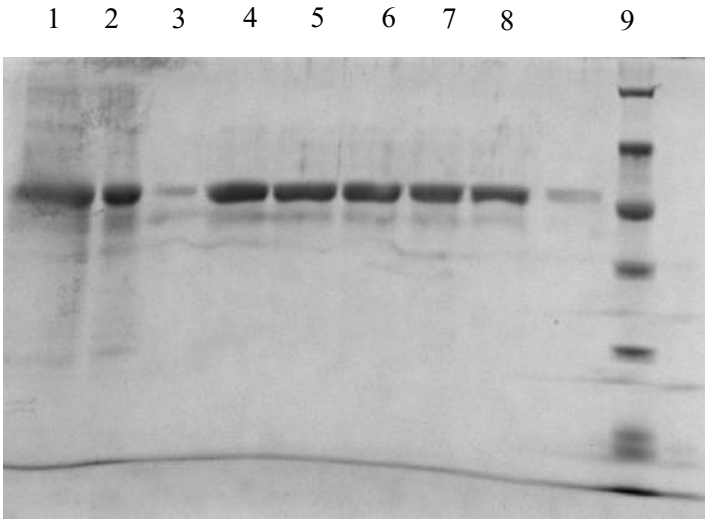


Figure 3-6: Heat treatment of CFE at five different temperatures

12.5% SDS-PAGE was used to show the difference in protein content after incubation at five different temperatures. Lane 1, CFE BH; Lane 2, CFE BH after centrifugation; Lane 4, CFE incubated at 80°C; Lane 5, CFE incubated at 85°C; Lane 6, CFE incubated at 90°C; Lane 7, CFE incubated at 95°C; Lane 8, CFE incubated at 100°C; Lane 9, Protein marker. 10 µl of CFE was loaded per lane.

3.3.2 Incubation period

The optimal incubation period at 100°C was deduced by incubating CFE for 10, 20, 40, and 50 minutes. The incubation was followed by centrifugation and the supernatant was used for determination of protein concentration. The maximum amount of protein was removed at 10 minutes of incubation (≈ 60 %). This did not change after 20, 40 and 50 minutes of incubation (Figures 3-7 and 3-8).

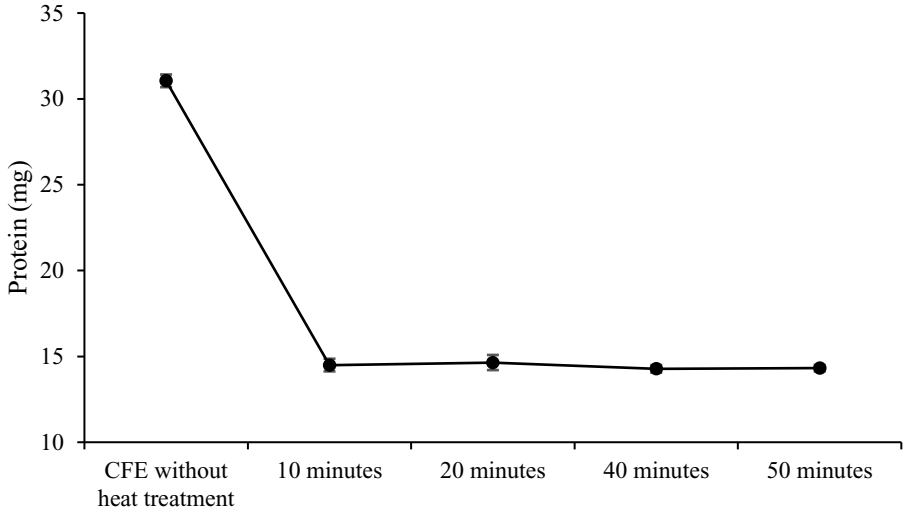


Figure 3-7: Heat treatment of CFE for different durations at 100°C

CFE was heat treated for 10, 20, 40, and 50 minutes at 100°C. Protein concentration was measured after incubation followed by centrifugation. The error bars show the deviation between duplicate experiments.

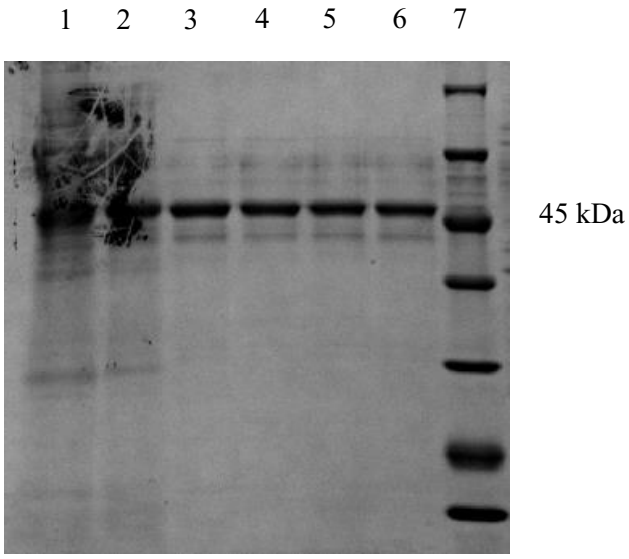


Figure 3-8: Heat treatment of CFE for different durations at 100°C

12.5% SDS-PAGE was used to show the difference in protein content after incubation for different durations at 100°C. Lane 1, CFE BH; Lane 2, CFE BH after centrifugation; Lanes 3, CFE incubated for 10 minutes; Lane 4, CFE incubated for 20 minutes; Lane 5, CFE incubated for 40 minutes; Lane 6, CFE incubated for 50 minutes; Lane 7, Protein marker. 15 µl of CFE was loaded per lane.

3.3.3 Concentration of CFE for incubation

In order to investigate if the efficiency of the heat treatment step depended on the concentration of the CFE, various concentrations of CFE were heat treated at 100°C for 10 minutes. Approximately 55% of protein was removed from the undiluted CFE (24.35 mg/ml). The same amount of protein reduction was seen across CFE with the concentrations 15 mg/ml, 10 mg/ml, and 5 mg/ml. Only 42 % and 11 % of protein was removed when the concentrations of the CFE were 2.5 mg/ml and 0.25 mg/ml, respectively (Figure 3-9). As expected, enzyme activity was approximately the same before and after heat treatment (Figure 3-9).

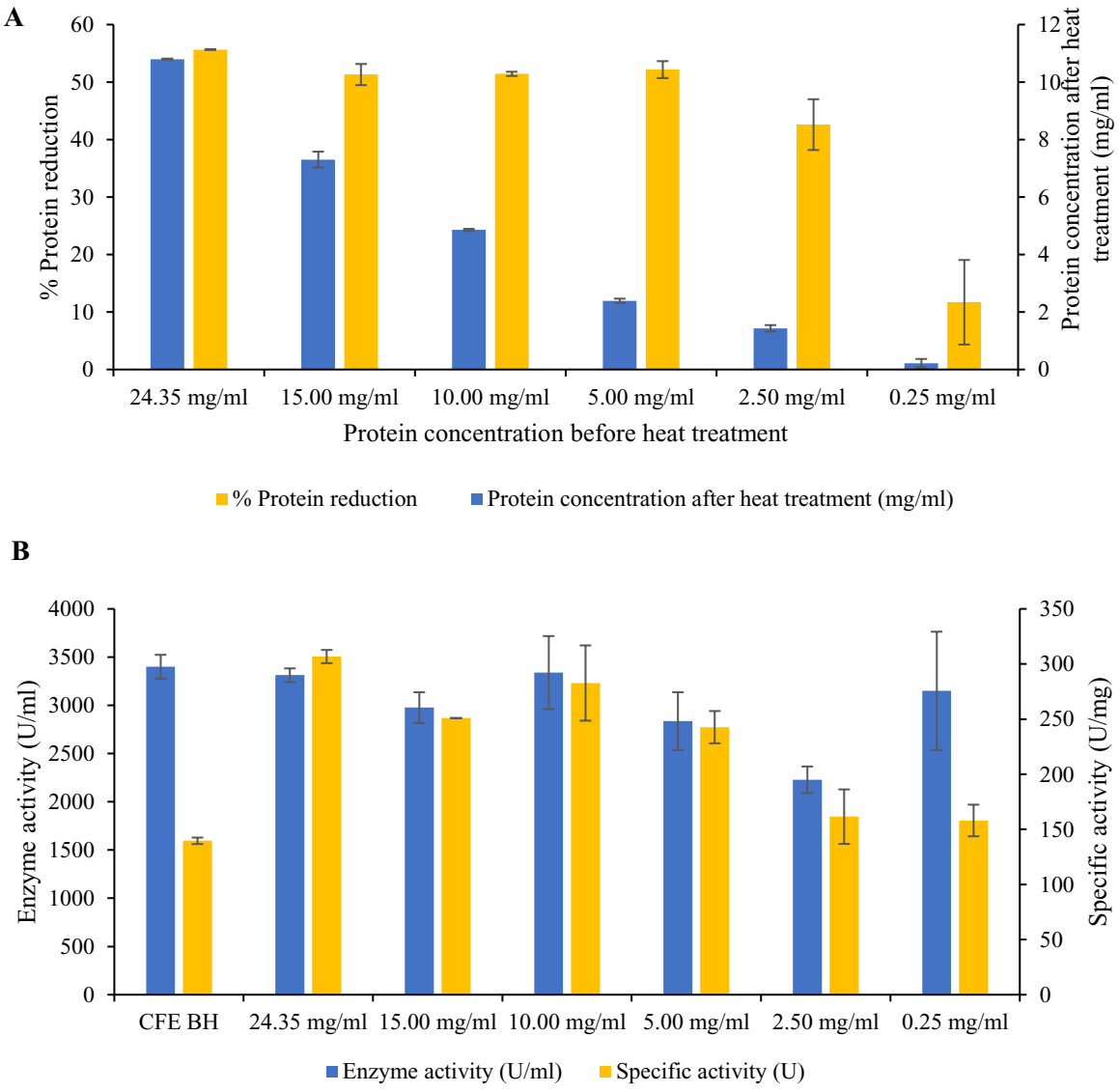


Figure 3-9: Heat treatment of different concentrations of CFE at 100°C

CFE was incubated at concentrations of 24.35 mg/ml (undiluted), 15.00 mg/ml, 10.00 mg/ml, 5.00 mg/ml, 2.50 mg/ml, and 0.25 mg/ml, at 100°C for 10 minutes. Protein concentration (A) and enzyme activity (B) were measured after incubation followed by centrifugation. The error bars show the deviation between duplicate experiments.

3.4 Purification of the recombinant wild-type *T. guaymasensis* ADH

The recombinant TgADH was purified from *E. coli* using a modified procedure. Prior to liquid chromatography, the cell free extract was heat treated. Heating at 100°C for 10 mins caused no loss of enzyme activity but significantly reduced the protein concentration to 40 % (Figure 3-14). The TgADH activity was dominant in the cell-free extract after heat treatment. Following heat treatment, the supernatant was loaded on to a phenyl-sepharose column and eluted using either linear or segmented gradient.

The segmented gradient was run for three and a half column volumes (230 ml) and was divided into two segments (Figure 3-10). The first segment ran from 41 % (0.82 M) to 12.5 (0.25 M) % of the gradient, followed by the second segment, which ran from 12.5 % (0.25 M) to 0 % (0.00 M) of the gradient. The recombinant TgADH started eluting at 0.2 M of the gradient. However, the enzyme was not purified, as can be seen from the SDS-PAGE (Figure 3-11). In fact, many of the fractions containing TgADH accumulated other protein contaminants, which in turn reduced the specific activity (Table 3-1).

The linear gradient was run for six column volumes (390 ml), starting from 0.82 M to 0.00 M of buffer B (Figure 3-12). Similar to the segmented gradient purification, the recombinant TgADH started eluting around 0.2 M of the gradient. Again, the enzyme was not purified, as can be seen from the SDS-PAGE (Figure 3-13) and the specific activity was 65 % lower than that obtained after heat treatment (Table 3-2).

Since purification using phenyl-sepharose column did not yield a pure sample, the recombinant TgADH was partially purified using only heat treatment (Figure 3-14).

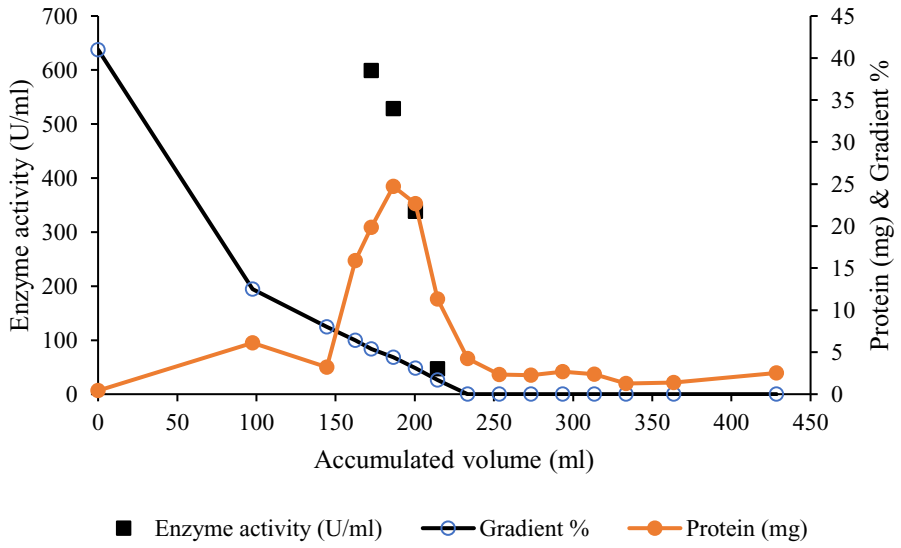
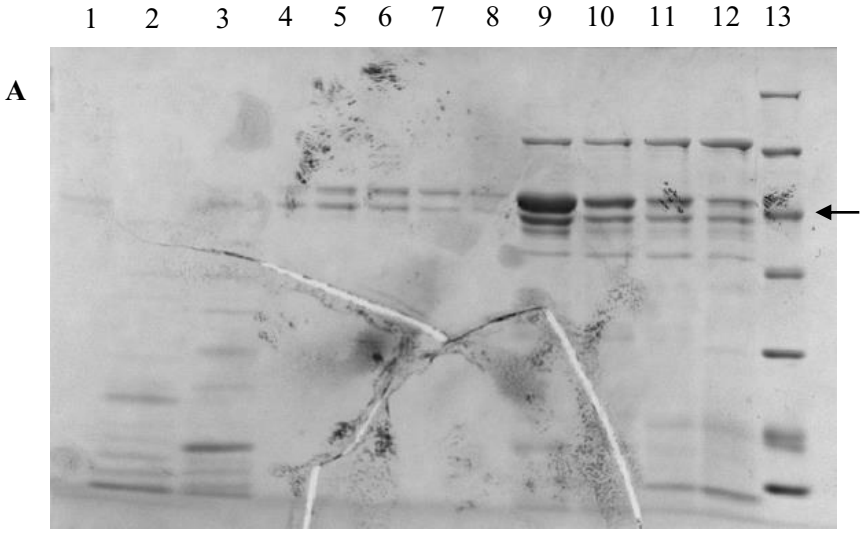


Figure 3-10: Purification of recombinant TgADH using a phenyl-sepharose column, with segmented gradient

Open circles, two segmented gradient (41 % to 0 % of buffer B); Filled circles, protein content of each fraction eluted; Filled squares, TgADH activity. Individual data points were single experiments and not done in replicates.

Table 3-1: Purification of recombinant TgADH using a phenyl-sepharose column, with segmented gradient

Step	Total protein (mg)	Enzyme activity (U/ml)	Total activity (U)	Specific activity (U/mg)
CFE BH	474.50	1339.04	26780.95	56.44
CFE AH	191.61	1973.32	68076.68	355.29
Phenyl sepharose (fraction 05)	19.86	599.04	5990.47	301.51



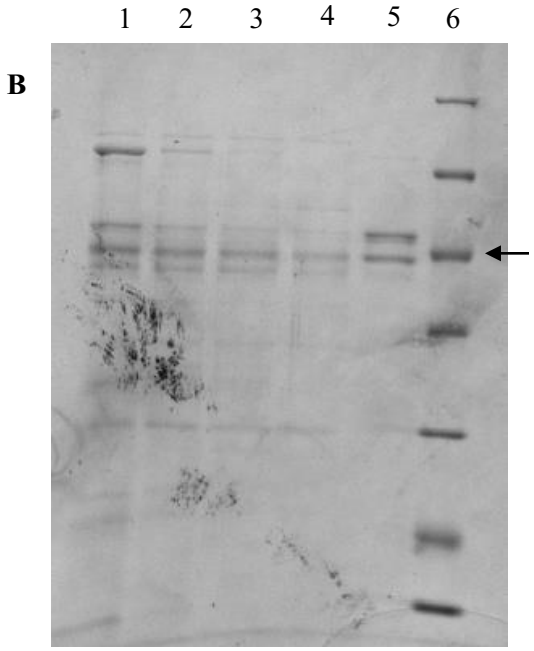


Figure 3-11: Fractions collected from phenyl-sepharose column when segmented gradient was used

12.5% SDS-PAGE was used to show the purity of the fractions (the amount of protein in each lane is given in brackets, µg)

A: Lane 1, Fraction 1 (0.11); Lane 2, Fraction 2 (0.93), Lane 3, Fraction 3 (1.02); Lane 4, Fraction 4 (0.13); Lane 5, Fraction 5 (0.29); Lane 6, Fraction 6 (0.26); Lane 7, Fraction 7 (0.24); Lane 8, Fraction 8 (0.12); Lane 9, Fraction 9 (3.35); Lane 10, Fraction 10 (1.76), Lane 11, Fraction 11 (1.68); Lane 12, Fraction 12 (1.99); Lane 13, Protein marker

B: Lane 1, Fraction 13 (1.78); Lane 2, Fraction 14 (0.94), Lane 3, Fraction 15 (0.69); Lane 4, Fraction 16 (0.58); Lane 5, Fraction CFE AH (0.48); Lane 6, Protein marker

Black arrow denotes the 45 kDa band of the protein marker

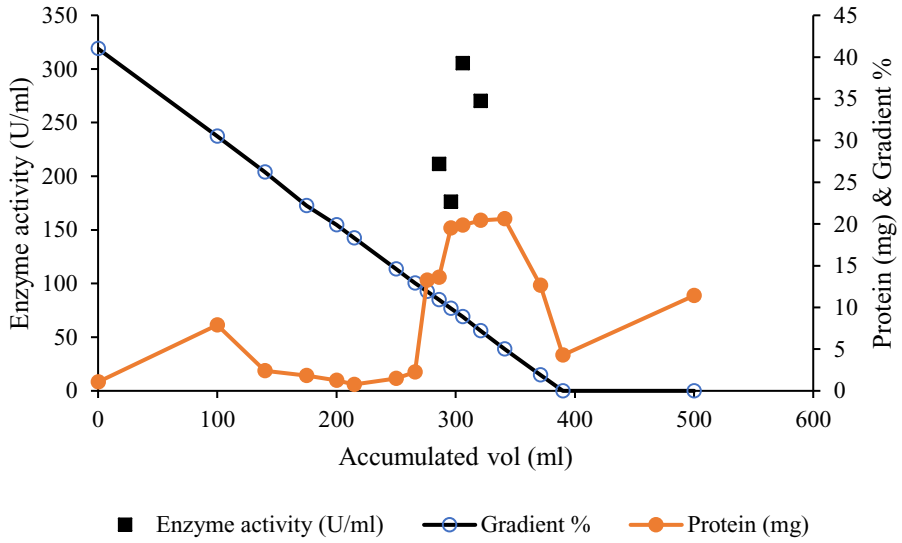
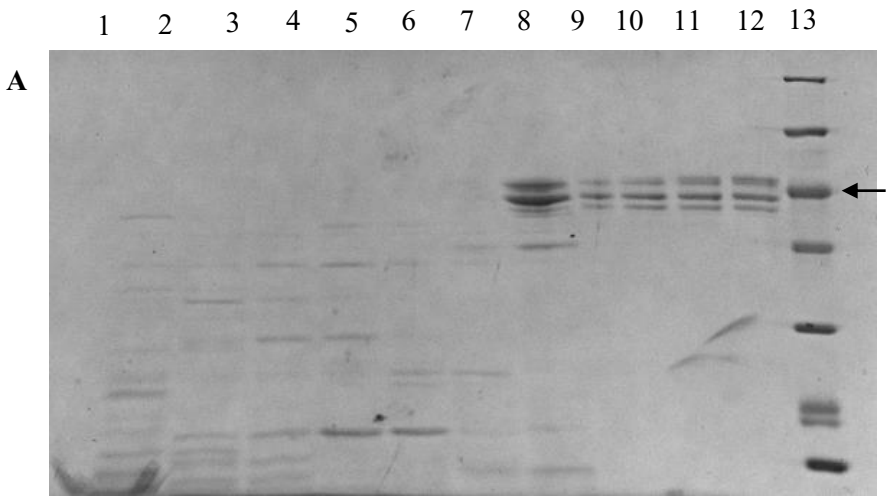


Figure 3-12: Purification of TgADH using a phenyl-sepharose column, with a linear gradient

Open circles, linear gradient (41 % to 0 % of buffer B); Filled circles, protein content of each fraction eluted; Filled squares, TgADH activity. Individual data points were single experiments and not done in replicates.

Table 3-2: Purification of recombinant TgADH using a phenyl-sepharose column, with a linear gradient

Step	Total protein (mg)	Enzyme activity (U/ml)	Total activity (U)	Specific activity (U/mg)
CFE BH	576.25	3976.97	99424.25	172.53
CFE AH	210	1242.32	93174.00	443.68
Phenyl sepharose (fraction 12)	19.86	305.39	3053.9	153.77



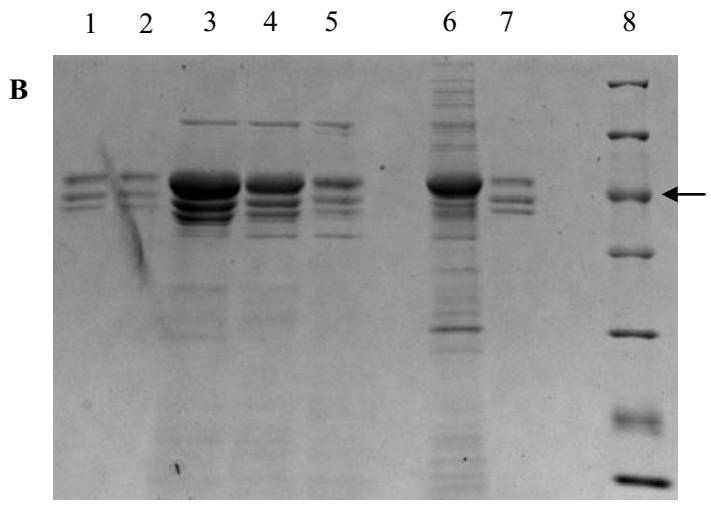


Figure 3-13: Fractions collected from phenyl-sepharose column when linear gradient was used

12.5% SDS-PAGE was used to show the purity of the fractions (the amount of protein in each lane is given in brackets, µg)

A: Lane 1, Fraction 1 (0.26); Lane 2, Fraction 2 (1.18), Lane 3, Fraction 3 (0.91); Lane 4, Fraction 4 (0.78); Lane 5, Fraction 5 (0.76); Lane 6, Fraction 6 (0.77); Lane 7, Fraction 7 (0.62); Lane 8, Fraction 8 (2.11); Lane 9, Fraction 9 (0.39); Lane 10, Fraction 10 (0.41), Lane 11, Fraction 11 (0.58); Lane 12, Fraction 12 (0.59); Lane 13, Protein marker

B: Lane 1, Fraction 13 (0.41); Lane 2, Fraction 14 (0.31), Lane 3, Fraction 15 (6.31); Lane 4, Fraction 16 (3.36); Lane 5, Fraction 17 (1.55); Lane 6, CFE BH (3.45); Lane 7, CFE AH (0.84); Lane 8, Protein marker

Black arrow denotes the 45 kDa band of the protein marker

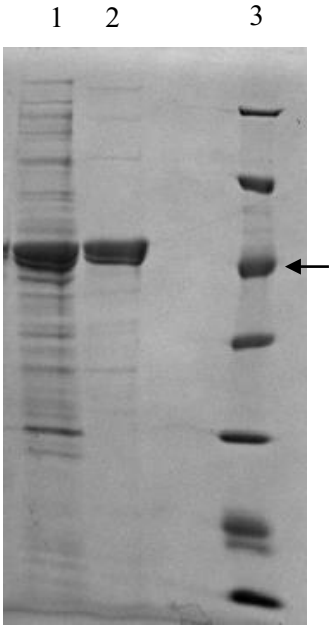


Figure 3-14: Partial purification of TgADH using heat treatment at 100°C for 10 minutes

12.5% SDS-PAGE was used to show the amount of protein removed before and after heat treatment. 15 µl of CFE was loaded per lane.

Lane 1, CFE BH; Lane 2, CFE AH; Lane 3, Protein marker

Black arrow denotes the 45 kDa band of the protein marker

Table 3-3: Partial purification of recombinant TgADH using heat treatment

Step	Total protein (mg)	Enzyme activity (U/ml)	Total activity (U)	Specific activity (U/mg)
CFE BH	599.71	2819.04	70476.19	117.51
CFE AH	290.66	3523.80	88095.23	303.08

3.5 Oxygen sensitivity of C212S and C305S TgADH mutants

Wild-type TgADH is known to lose its activity under oxidative conditions. The half-life ($t_{1/2}$) of recombinant TgADH when exposed to air was found to be 4 hours (Hao, unpublished; Figure 3-14). It is hypothesized that the loss of activity under aerobic conditions is due to oxidation of certain amino acids, especially cysteine residues. TgADH has four Cys residues in each of its subunits, at positions, 39, 56, 212, and 305 (Ying and Ma, 2011). Cys39 coordinates with zinc while mutation studies of Cys56 have eliminated it being the cause for the loss of activity when exposed to air (Ying and Ma, 2011; Ma and Ying, 2013). Here, the oxygen tolerance of Cys mutants, C212S and C305S, was analyzed in this study.

Both C212S and C305S mutants showed very similar enzyme activity as recombinant wild-type TgADH, under anaerobic conditions (Table 3-4). Heat treated CFE of both the mutants were exposed to air under stirred condition and enzyme activity was measured over time, with and without reducing agents (2 mM SDT and 2 mM DTT). C305S showed a half-life ($t_{1/2}$) of 1.75 hours and lost activity at a higher rate, compared to C212S, which showed a half-life ($t_{1/2}$) of 3.5 hours (Figures 3-16 and 3-17). The loss of activity was slightly protected in the presence of 2 mM SDT and 2 mM DTT (Figures 3-16 and 3-17). Plotting the natural log of loss of activity over time showed that the reactions were first order in nature (Figure 3-18). The half-lives were re-calculated for C212S, C305S and wild-type TgADH using their respective inactivation constant (j) and the equation $t_{1/2} = \ln 2 / j$ (Petrucci et al., 2007). The newly calculated half-lives were within the range of those deduced previously.

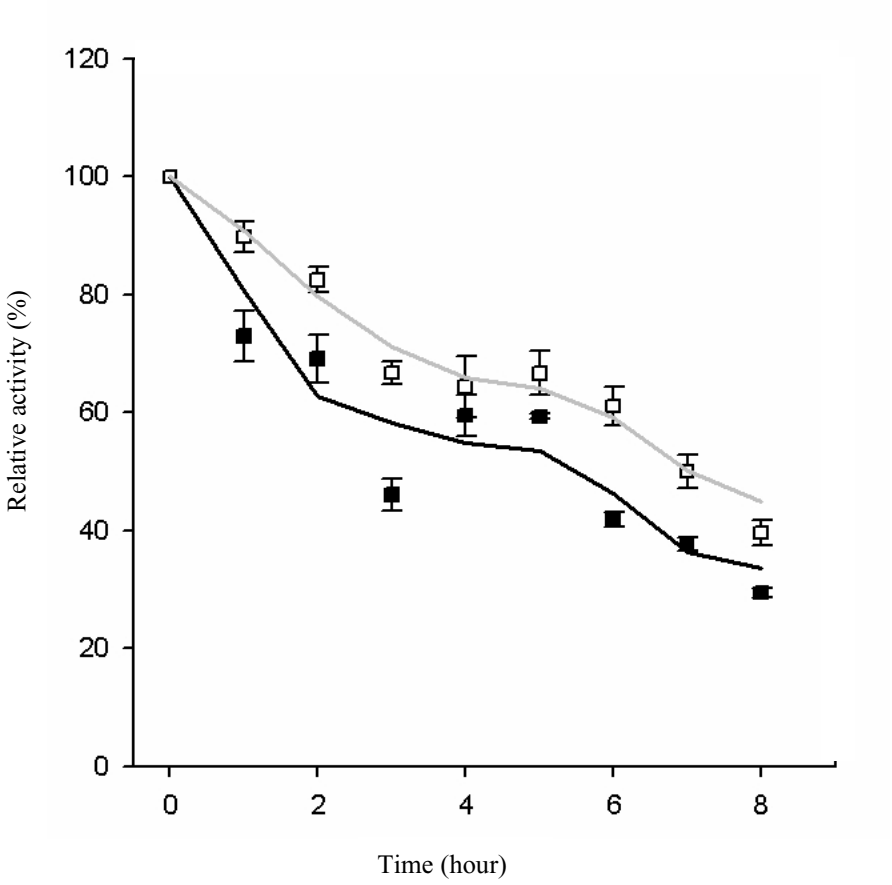


Figure 3-15: Oxygen sensitivity of purified recombinant wild-type TgADH

Open squares, in the presence of 2 mM DTT and 2 mM SDT; Filled squares, in the absence of 2 mM DTT and 2 mM SDT. The relative activity of 100% equals to the ADH activity prior to exposure to air (1029 U/mg). Standard deviations of the measurements are indicated using error bars. (Taken from Hao, unpublished)

Table 3-4: Partial purification of recombinant Cys mutants

Mutant	Step	Protein concentration (mg/ml)	Enzyme activity (U/ml)	Specific activity (U/mg)
C305S	CFE	26.41	3026.12	114.58
	Heat treatment	11.92	2748.57	230.58
C212S	CFE	21.92	3641.35	166.12
	Heat treatment	10.52	3805.71	361.75

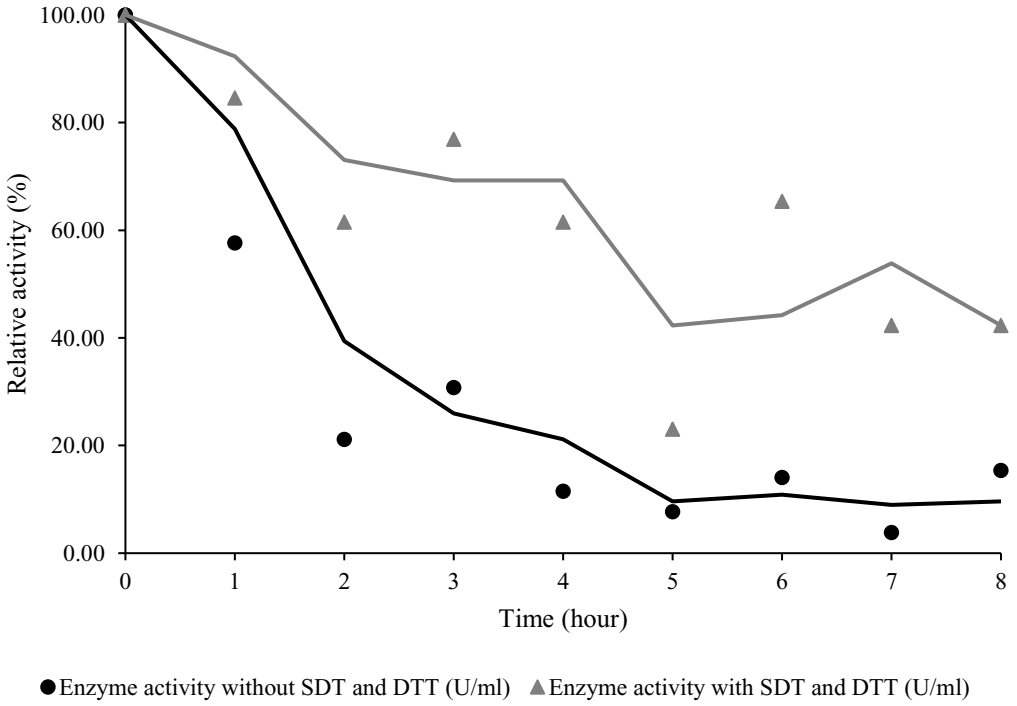


Figure 3-16: Oxygen sensitivity of partially purified TgADH Cys mutant C305S

Filled triangles, in the presence of 2 mM DTT and 2 mM SDT; Filled circles, in the absence of 2 mM DTT and 2 mM SDT. The relative activity of 100 % equals to the ADH activity prior to exposure to air (230.58 U/mg). Individual data points were single experiments and not done in replicates.

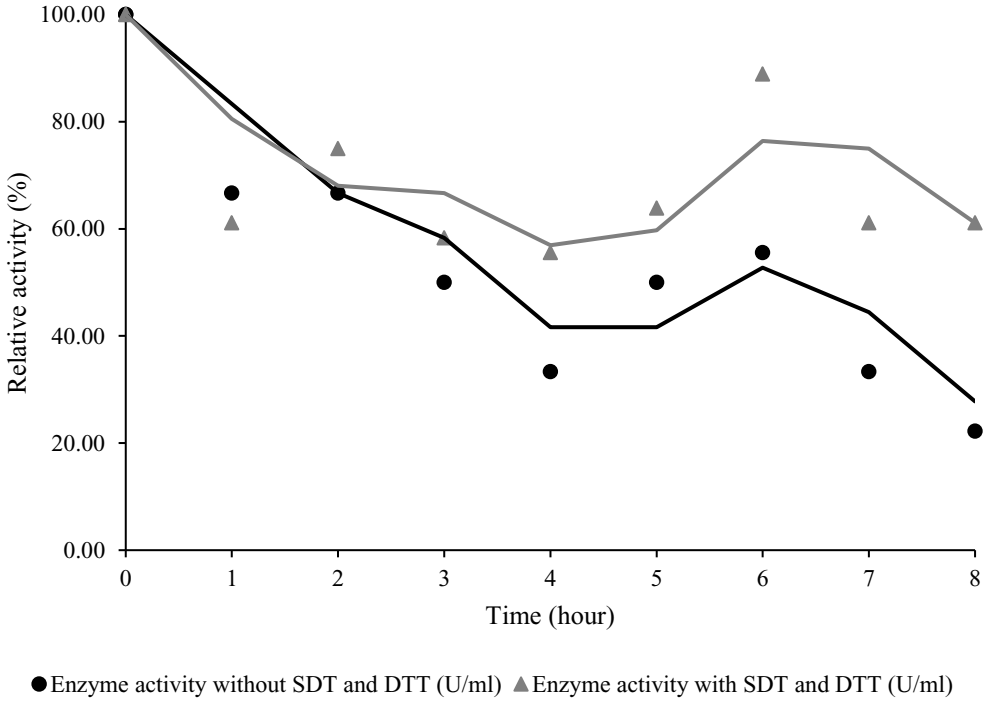


Figure 3-17: Oxygen sensitivity of partially purified TgADH Cys mutant C212S

Filled triangles, in the presence of 2 mM DTT and 2 mM SDT; Filled circles, in the absence of 2 mM DTT and 2 mM SDT. The relative activity of 100 % equals to the ADH activity prior to exposure to air (361.75 U/mg). Individual data points were single experiments and not done in replicates.

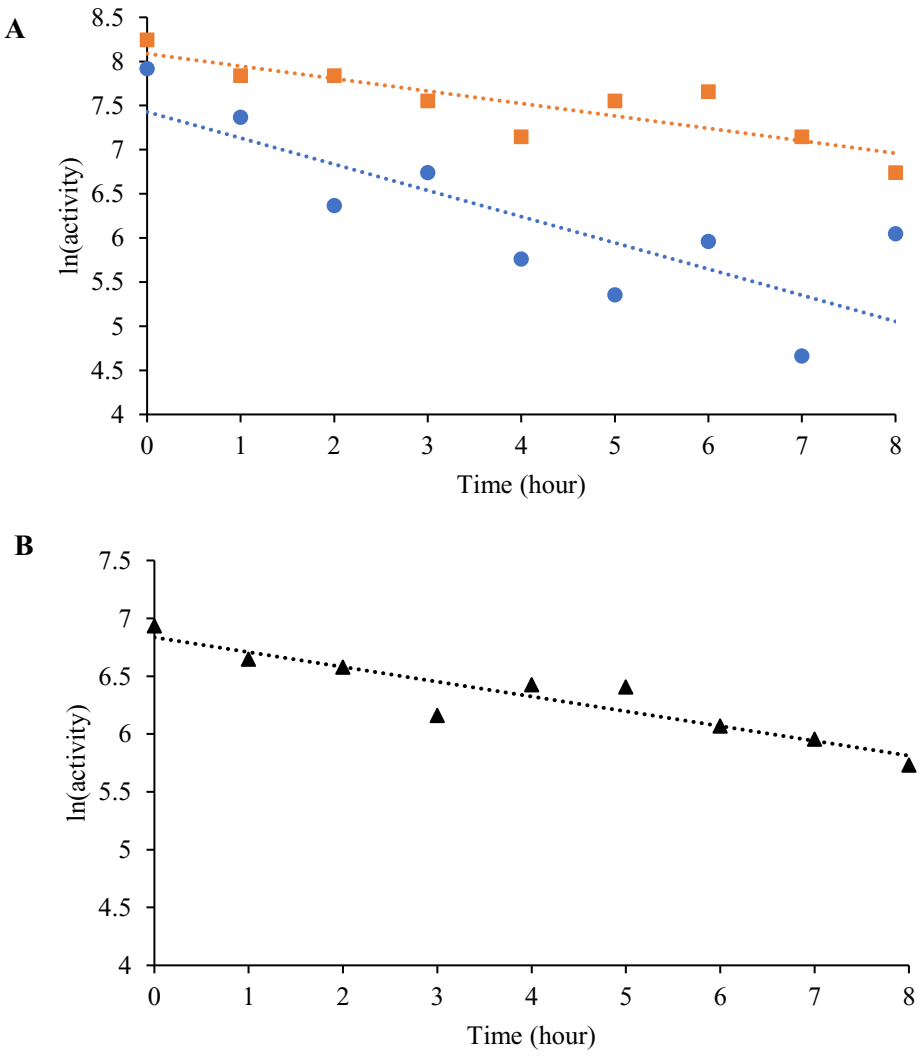


Figure 3-18: Inactivation of C212S, C305S and wild-type TgADH

Inactivation of **(A)** partially purified C212S mutant (Filled squares) and C305S mutant (Filled circles) in the absence of 2 mM DTT and 2 mM SDT, **(B)** purified wild-type recombinant TgADH (adopted from Hao, unpublished) in the absence of 2 mM DTT and 2 mM

3.6 Mass spectrometry analysis

3.6.1 Calibration

External calibration of the mass spectrometer was done using a standard sample (Bruker, ON, Canada) mixed with DHB at 1:1 ratio. The calibration process was repeated until the spectrum obtained with the measured m/z of each protein standard was as close as to its respective reported molecular mass (Figure 3-19). Protein A $[M+H]^{2+}$ and Protein A $[M+H]^{+}$ showed an error of 0.57 Da and 0.59 Da, respectively, when compared to the reported mass. Trypsinogen $[M+H]^{+}$ showed an error of 20.01. Albumin-Bovine (BSA) $[M+H]^{2+}$ and Albumin-Bovine (BSA) $[M+H]^{2+}$ showed the highest error of 64.68 Da and 80.65 Da, respectively (Table 3-5). This was expected because as the size of the protein gets bigger the respective peak gets broader, making it difficult to pick the highest point for calibration. The overall accuracy range achieved from the external calibration was 0.0013-0.19 % (Figure 3-20).

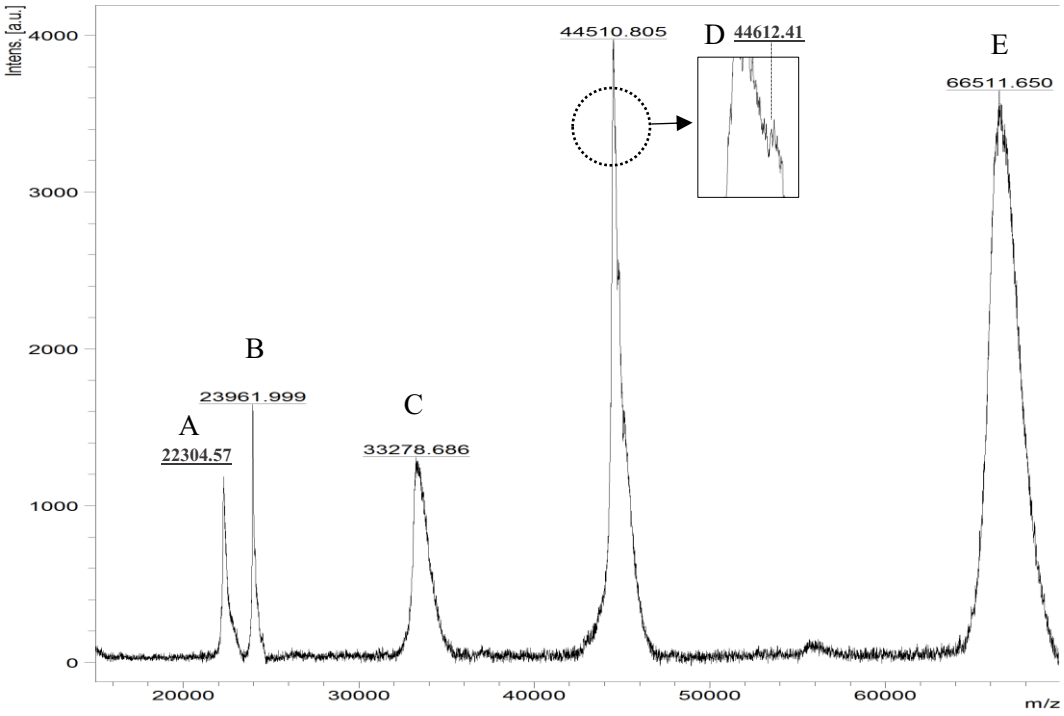


Figure 3-19: External calibration of mass spectrometer using a standard mixture

A: Protein A $[M+H]^{2+}$; **B:** Trypsinogen $[M+H]^{+}$; **C:** Albumin-Bovine (BSA) $[M+H]^{2+}$; **D:** Protein A $[M+H]^{+}$; **E:** Albumin-Bovine (BSA) $[M+H]^{+}$

The portion inside the dotted circle is enlarged (inset) to show the corresponding peak of Protein A $[M+H]^{+}$ that was used for calibration.

Table 3-5: External calibration of mass spectrometer using a standard mixture

Standard	Reported mass (Da)	Measured mass (Da)	Difference (Da)
Protein A $[M+H]^{2+}$	22304	22304.57	0.57
Trypsinogen $[M+H]^{+}$	23982	23961.99	20.01
Albumin-Bovine (BSA) $[M+H]^{2+}$	33214	33278.68	64.68
Protein A $[M+H]^{+}$	44613	44612.41	0.59
Albumin-Bovine (BSA) $[M+H]^{+}$	66431	66511.65	80.65

Components of the standard mixture (Bruker, ON, Canada), their reported and measured masses and the difference between the reported and measured masses are shown.

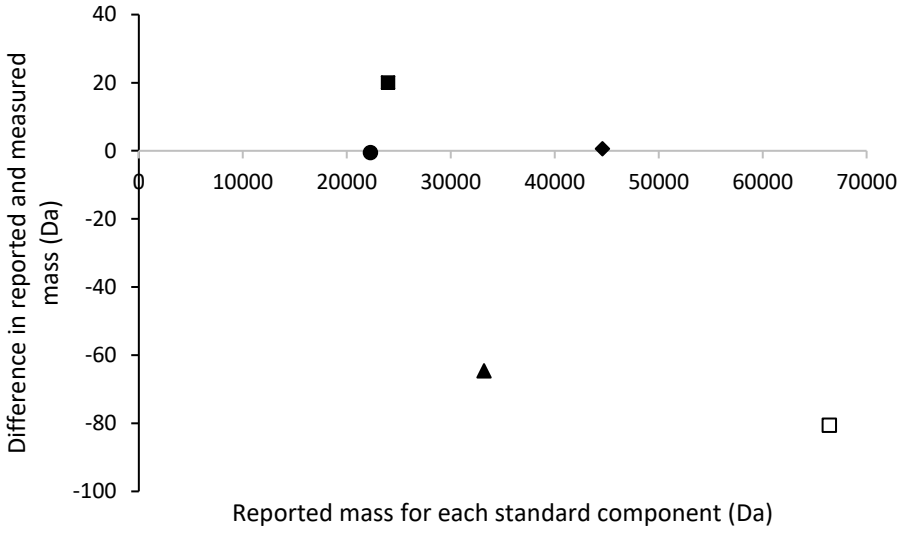


Figure 3-20: The difference between reported and measured masses of each component in the standard mixture

Filled circle, Protein A $[M+H]^{2+}$; Filled square, Trypsinogen $[M+H]^{+}$; Filled triangle, Albumin-Bovine (BSA) $[M+H]^{2+}$; Filled diamond, Protein A $[M+H]^{+}$; Open square, Albumin-Bovine (BSA) $[M+H]^{+}$

3.6.2 Optimal matrix

In order to find a suitable matrix, the active recombinant TgADH was analyzed with two different matrices, SA and DHB. The matrices were mixed with the sample at 1:1 ratio, spotted on the MALDI sample plate and air dried. 5,000 shots were acquired for each spot and the results were combined to increase the intensity of the signal and the quality of the spectrum. The sample with SA yielded a spectrum low in sensitivity and with two major peaks at 45,000 Da range (45,167 and 45,258) at a resolution of 273 (Figure 3-21). The recombinant TgADH being 45,260.14 Da should only give rise to one major peak at the 45,000 Da range. The sample with DHB yielded a spectrum of higher sensitivity and only one major peak at the 45,000 Da range at a resolution of 147 (Figure 3-21). Therefore, DHB was used for all further analyses.

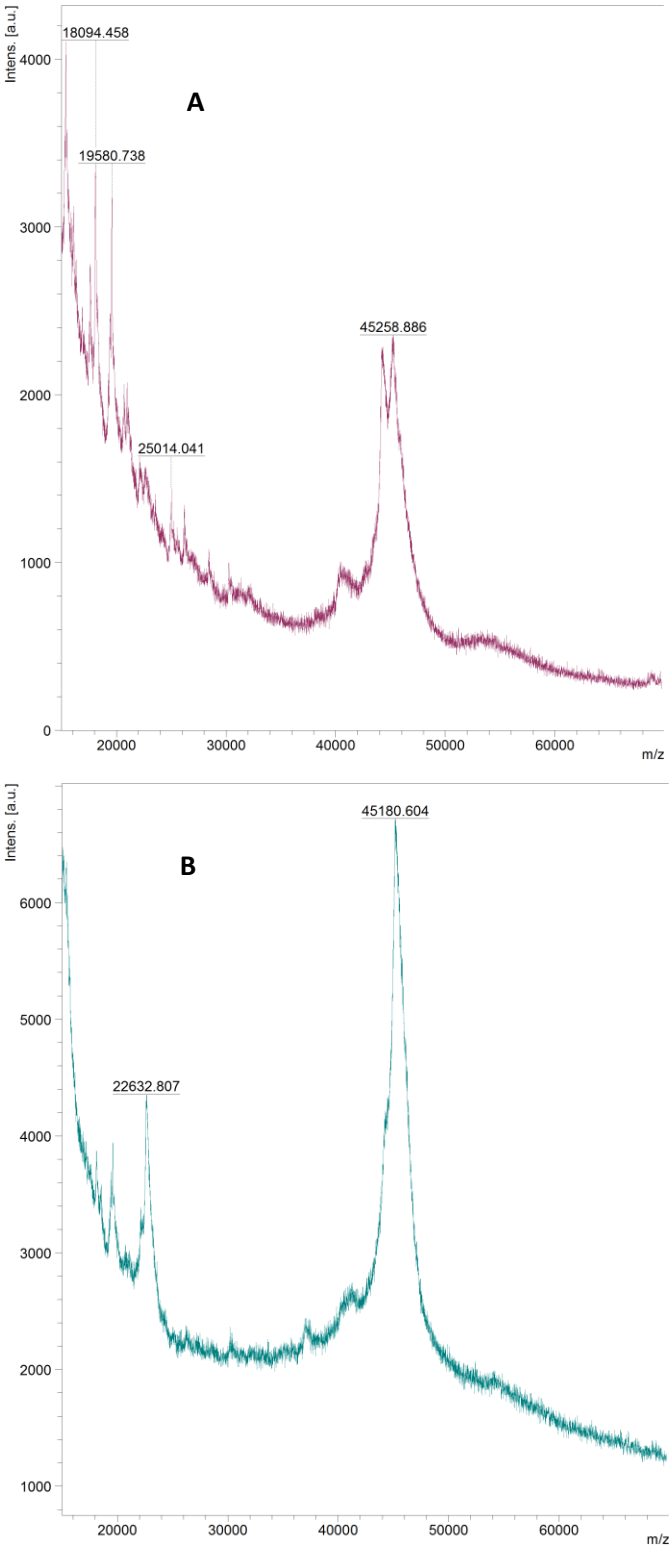


Figure 3-21: Selecting a matrix for MALDI-TOF analysis

Saturated sinapinic acid (A) and dihydroxybenzoic acid (B) were used with the active recombinant TgADH sample.

3.6.3 Analysis of oxidized recombinant TgADH

To investigate the correlation between oxidation of -SH groups and increase in mass, the recombinant TgADH was subjected to oxidative conditions, followed by MALDI-TOF mass spectrometry. Heat treated CFE of recombinant TgADH was either exposed to air or different concentrations of H_2O_2.

It was expected that the measured m/z of unoxidized active TgADH would be accurate up to 0.0013 – 0.19 % of the reported mass, as deduced from the external calibration. However, the measured m/z was 45,260.13, 47 Da less than reported mass (Table 3-6). The air and H_2O_2 oxidized TgADH were expected to show an increase in mass (due to the addition of oxygen atom(s)), but the air oxidized and H_2O_2 oxidized (1:10, 1:40, 1:160, and 1:400) samples showed a decrease in mass, compared to the reported mass of TgADH. A measured m/z of more than 0.19 % of the reported mass of TgADH was expected for the oxidized samples. Anything less than that will not reflect the increase due to the post-translational oxidation of the enzyme, since the change would be within the nominal range of error of the calibration. The H_2O_2 oxidized sample with a concentration of 1:1600 mol/mol (protein subunit/H_2O_2) was the only sample that showed an increase in measured m/z over 0.19 % of the reported mass of TgADH. The measured mass was 45,619.84 Da, 359 Da more than the reported mass of unoxidized TgADH.

Table 3-6: Comparison of protein masses measured by MALDI-TOF/MS vs reported values

Sample	Reported mass (Da)	Measured mass from MALDI-TOF (average) (Da)	Difference between duplicates (Da)	Difference between reported and measured mass (Da)
Active (unoxidized)	45260.13	45213.12	±32.65	-47.015
Air oxidized		45248.11	±8.52	-12.02
H_2O_2 1:10		45236.62	±21.39	-23.51
H_2O_2 1:40		45210.76	-	-49.37
H_2O_2 1:160		45235.99	-	-24.14
H_2O_2 1:400		45243.2	±46.92	-16.93
H_2O_2 1:1600		45619.84	-	+359.71

The activity of oxidized TgADH was monitored over increasing concentrations of. TgADH gradually lost its activity upon incubation with increasing concentrations of H_2O_2 (Figure 3-22). The enzyme activity was inhibited by approximately 50% at a molar ratio of 1:30 mol/mol (protein subunit/H_2O_2) and the activity was almost completely abolished when the ratio reached 1:1600 mol/mol. This correlates well with the increase in mass due to post-translational oxidation, where the highest increase in mass is seen when the ratio of protein/H_2O_2 was 1:1600 mol/mol. This indicates that at 1:1600 mol/mol (protein subunit/H_2O_2) all the cysteines and most of the methionine residues were oxidized, leading to almost completely inactivating the enzyme. The number of oxygens added at 1:10, 1:40, 1:160, and 1:400 mol/mol (protein subunit/ H_2O_2) cannot be deduced accurately since the difference in measured m/z is not more than the range of error for unoxidized TgADH. However, an increasing trend in mass can be seen over increasing concentrations of H_2O_2.

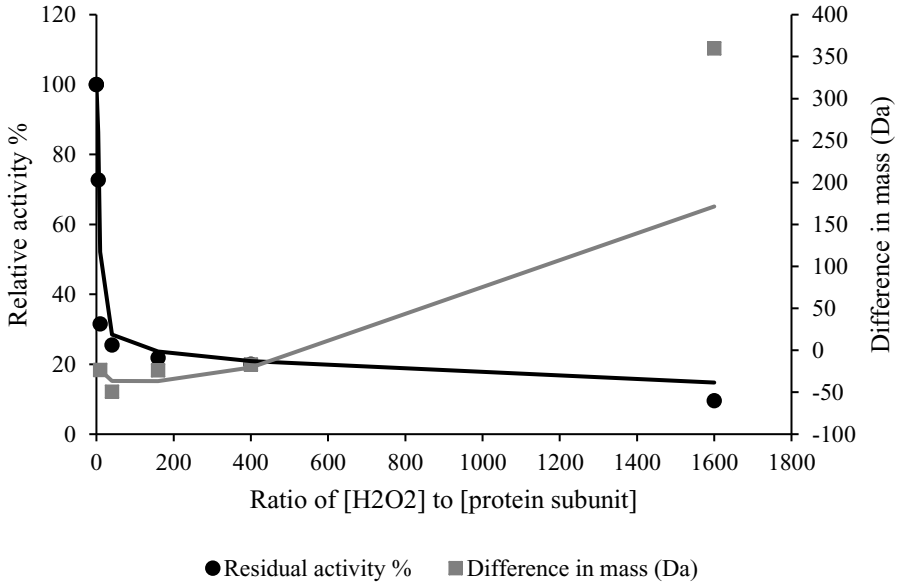


Figure 3-22: Correlation between enzyme activity and change in mass due to oxidation by H_2O_2

Filled squares, difference between reported mass of unoxidized TgADH and measured mass after treating with different concentrations of H_2O_2; Filled circles, relative activity of TgADH at different concentrations of H_2O_2. Individual data points were single experiments and not done in replicates.

3.7 Oxidative degradation of recombinant TgADH

When fractions containing recombinant TgADH that were collected from phenyl-sepharose purification were stored at 4°C under anaerobic conditions, degradation of the enzyme was observed over time (Figure 3-23). It was hypothesized that the degradation could be due to the oxidation of the enzyme when the anaerobic conditions are not maintained. In order to investigate if different storage/oxidative conditions would cause the recombinant TgADH to degrade, heat treated CFE was stored under anaerobic conditions at room temperature, 4°C, -20°C, and under aerobic conditions at 4°C. The enzyme activity and degradation were monitored over time.

As expected, the CFE stored under aerobic conditions at 4°C showed a steady decrease in enzyme activity over time, with a half-life ($t_{1/2}$) of 48 hours, reconfirming the oxygen sensitivity of TgADH (Figure 3-24). It should be noted that this half-life cannot be compared with the half-life of TgADH recorded by Hao (unpublished) since the experiment conditions were different. CFE stored under anaerobic conditions at room temperature and 4°C showed about 28 % decrease in enzyme activity over 48 and 96 hours., whereas CFE stored at -20°C showed only 25 % decrease after 144 hours. This difference can be attributed to the way the CFE was stored. Only one anaerobic vial containing the CFE was stored under room temperature and 4°C while three different anaerobic vials with the CFE were stored under -20°C. Therefore, the vials at room temperature and 4°C had a higher chance of getting contaminated with oxygen when sampling at each time point whereas the use of fresh vials for sampling from -20°C, reduced the chance of oxygen contamination.

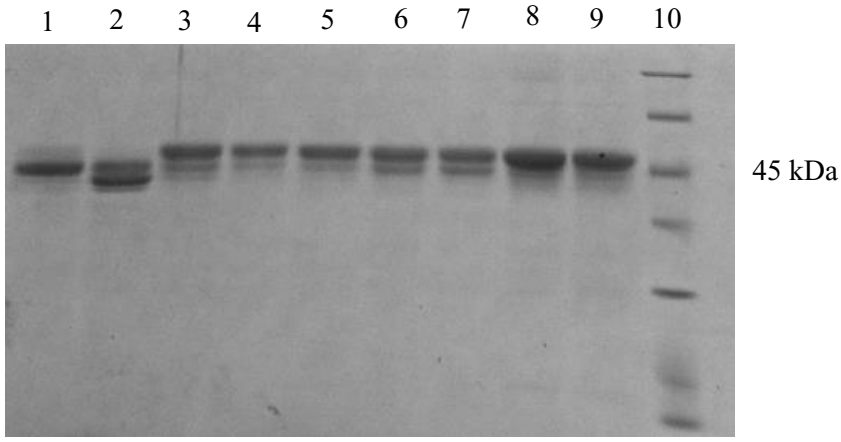


Figure 3-23: Degradation of recombinant TgADH in phenyl-sepharose fractions

When phenyl-sepharose fractions were stored at 4°C under anaerobic conditions, some of them showed enzyme degradation over time. Lane 1, Fraction with degraded enzyme that has lost ≈ 2,000 Da; Lane 2, Fraction with degraded enzyme that has lost ≈ 2,000 and ≈ 3,000 Da; Lane 3-7, Fractions without any enzyme degradation; Lane 8-9, Freshly heat treated CFE; Lane 10, Protein marker

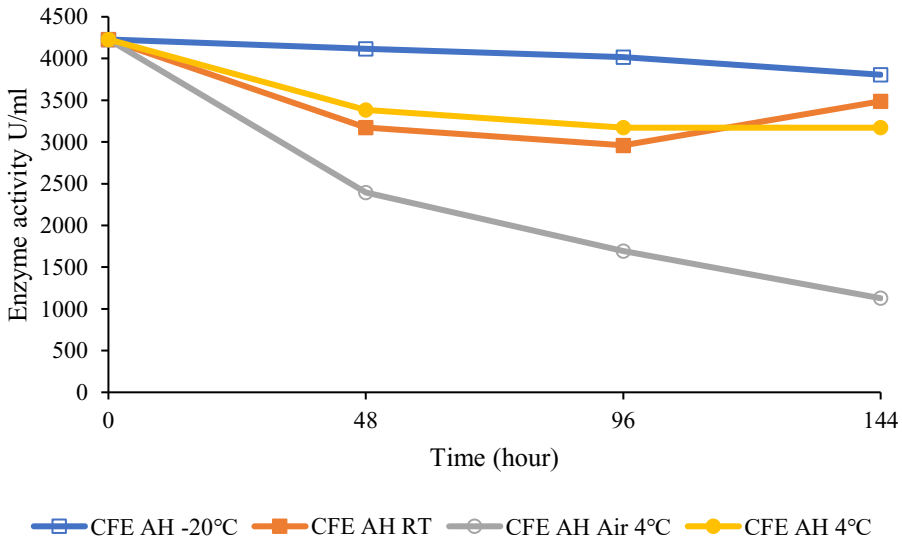


Figure 3-24: Enzyme activity of TgADH under different storage conditions

Open squares, anaerobic at -20°C; Filled squares, anerobic at room temperature; Open circles, aerobic at 4°C; Filled circles, anaerobic at 4°C. Individual data points were single experiments and not done in replicates.

An unusual phenomenon was observed when monitoring enzyme degradation over time. No change in structural conformation was seen after 48 and 96 hours, but after 144 hours the CFE sample stored at room temperature showed some precipitation. TgADH in that specific CFE was found to have lost approximately 2,000 kDa while the others still remained undegraded (Figure 3-25). It was expected that the enzyme sample exposed to aerobic conditions would show degradation, if any, but not the samples stored under anaerobic conditions. Furthermore, the degradation did not affect the enzyme activity drastically. Therefore, further microbial analyses were done to check if the degradation was due to the presence of certain microorganisms.

All four CFE samples were streaked on LB-agar plates. Only the CFE sample stored under anaerobic conditions at room temperature showed growth, while the remaining samples did not (Figure 3-26). Three morphologically different colonies were identified from the plate (colonies A, B and C). All three appeared to be a mix of rod-shaped cells and spores under the microscope. The three different colonies were used to inoculate freshly heat treated CFE to see if the degradation process occurred. The samples were incubated under anaerobic conditions at room temperature, for 24 hours. As expected, only the CFE samples inoculated with colonies A, B or C showed degradation, while CFE inoculated only with *E. coli* cells did not (Figure 3-27). This confirmed that the degradation is not due to *E. coli* or *E. coli* protein.

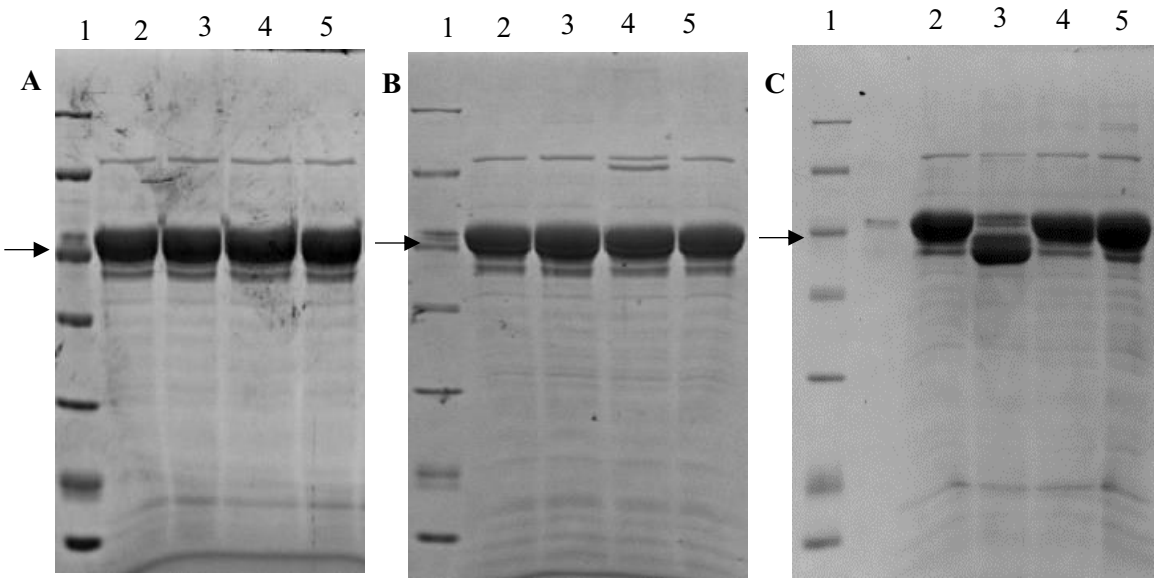


Figure 3-25: Degradation of TgADH under different storage conditions

A: After 48 hours, Lane 1, protein marker; Lane 2, anaerobic at -20°C; Lane 3, aerobic at 4°C; Lane 4, anaerobic at 4°C; Lane 5, anaerobic at room temperature

B: After 96 hours, Lane 1, protein marker; Lane 2, anaerobic at -20°C; Lane 3, aerobic at 4°C; Lane 4, anaerobic at 4°C; Lane 5, anaerobic at room temperature

C: After 144 hours, Lane 1, protein marker; Lane 2, anaerobic at -20°C; Lane 3, anaerobic at room temperature; Lane 4, anaerobic at 4°C; Lane 5, aerobic at 4°C

Black arrow denotes the 45 kDa band of the protein marker

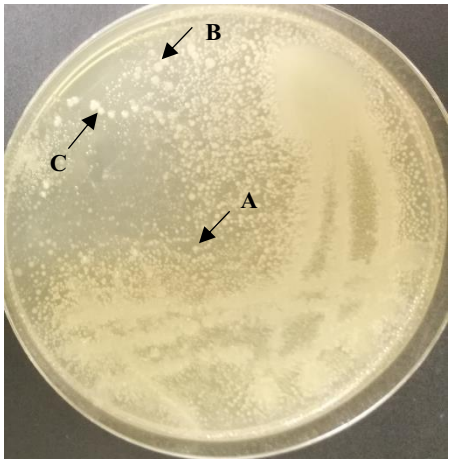


Figure 3-26: Streak plate of heat treated CFE stored under anaerobic conditions at room temperature, after 144 hours

Three morphologically different types of colonies were observed, labelled as colonies A, B, and C

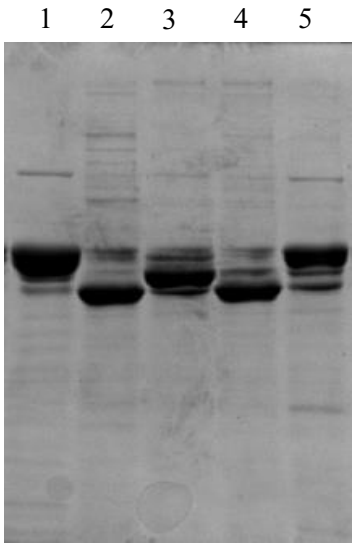


Figure 3-27: Degradation of TgADH after inoculating with colony A, B or C

Lane 1, Heat treated CFE; Lane 2, Heat treated CFE inoculated with colony A; Lane 3, Heat treated CFE inoculated with colony C; Lane 4, Heat treated CFE inoculated with colony B; Lane 5, Heat treated CFE inoculated with *E. coli*. 15 µl of CFE was loaded per lane.

Even though the TgADH in the three CFE samples inoculated with colonies A, B or C showed degradation, the pattern of degradation was not similar among them. TgADH in CFE inoculated with colonies A and B showed a loss of approximately 3,000 kDa while the one inoculated with colony C showed a loss of 2,000 kDa (Figure 3-27). This led to the speculation that that there could be two different organisms involved, which secrete two different enzymes that have protease or peptidase activity. In order to completely rule out the possibility of oxidative degradation of TgADH, all three colonies were grown aerobically and anaerobically, and the culture was used to inoculate freshly heat treated CFE. All three strains grew aerobically but only strain C grew anaerobically. Enzyme degradation was only observed in the CFE samples inoculated with either aerobically grown strains A, B or C, or anaerobically grown strain C (Table 3-7). This proved that the organism from colony A and B are the same and strictly aerobic while the organism from colony C is different and possibly a facultative anaerobe. However, it should be noted that the enzyme degradation pattern observed after inoculating strains A, B and C are different compared to what was observed earlier (Figure 3-28). Samples inoculated with the strictly aerobic strains A and B showed a dense band around 42,000 Da, as expected, but the facultative anaerobe colony C did not show a dense band around 43,000 Da. Instead, it too showed a band around 42,000 Da (Figure 3-28). This could be due to contamination of strain C with strain A and/or B. However, this proves that TgADH does not go through degradation due to oxidation.

Table 3-7: Growth of colonies A, B and C under aerobic and anaerobic conditions and their ability to degrade TgADH

Colony	Growth condition	Growth	Degradation of TgADH after inoculation
A	Aerobic	✓	✓
	Anaerobic	x	x
B	Aerobic	✓	✓
	Anaerobic	x	x
C	Aerobic	✓	✓
	Anaerobic	✓	✓

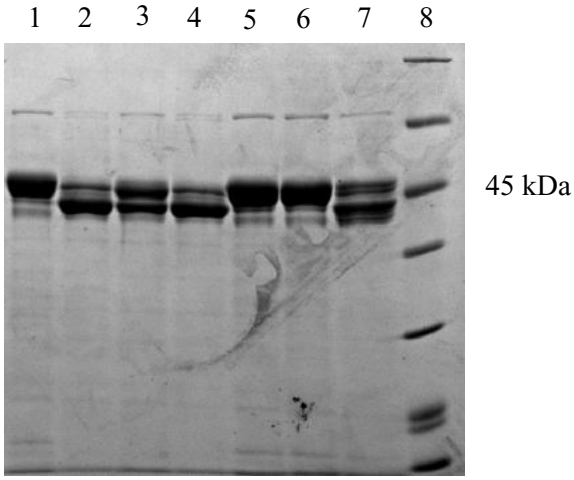


Figure 3-28: Degradation of TgADH by colonies A, B and C under aerobic and anaerobic conditions

Lanes 2-4: Incubation under aerobic conditions

Lanes 5-7: Incubation under anaerobic conditions

Lane 1, Heat treated CFE; Lane 2, Heat treated CFE inoculated with colony A; Lane 3, Heat treated CFE inoculated with colony C; Lane 4, Heat treated CFE inoculated with colony B; ; Lane 5, Heat treated CFE inoculated with colony A; Lane 6, Heat treated CFE inoculated with colony B; Lane 7, Heat treated CFE inoculated with colony C; Lane 8, Protein marker. 15 µl of CFE was loaded per lane.

3.8 Structural modelling and analysis of TgADH

No structural information exists about TgADH. Therefore, ADH from *Thermoanaerobacter brockii* (TbADH) (PDB ID 1PED) was used as template to model the structure of wild-type TgADH (Figure 3-29). TbADH is a zinc containing thermophilic homotetrameric enzyme. The crystal structure was solved at a resolution of 2.15 Å (Korkhin et al., 1996) and has a sequence similarity of 70 % with TgADH. The modelled structure of TgADH was used to understand the spatial arrangement of Cys residues relative to the zinc atom.

TgADH has four Cys residues per subunit; Cys39, Cys56, Cys212, and Cys305 (Ying and Ma, 2011). The zinc atom in each subunit is coordinated by Cys39, Ser41, His64, and Asp160 (Figure 3-29). The modelled structure showed Cys39 to be the closest to the zinc, at 2.387 Å, further confirming that it is one of the zinc-coordinating residues. Cys56, Cys212, and Cys305 were 11.814 Å, 16.218 Å, and 6.710 Å away from the zinc atom (Figure 3-29).

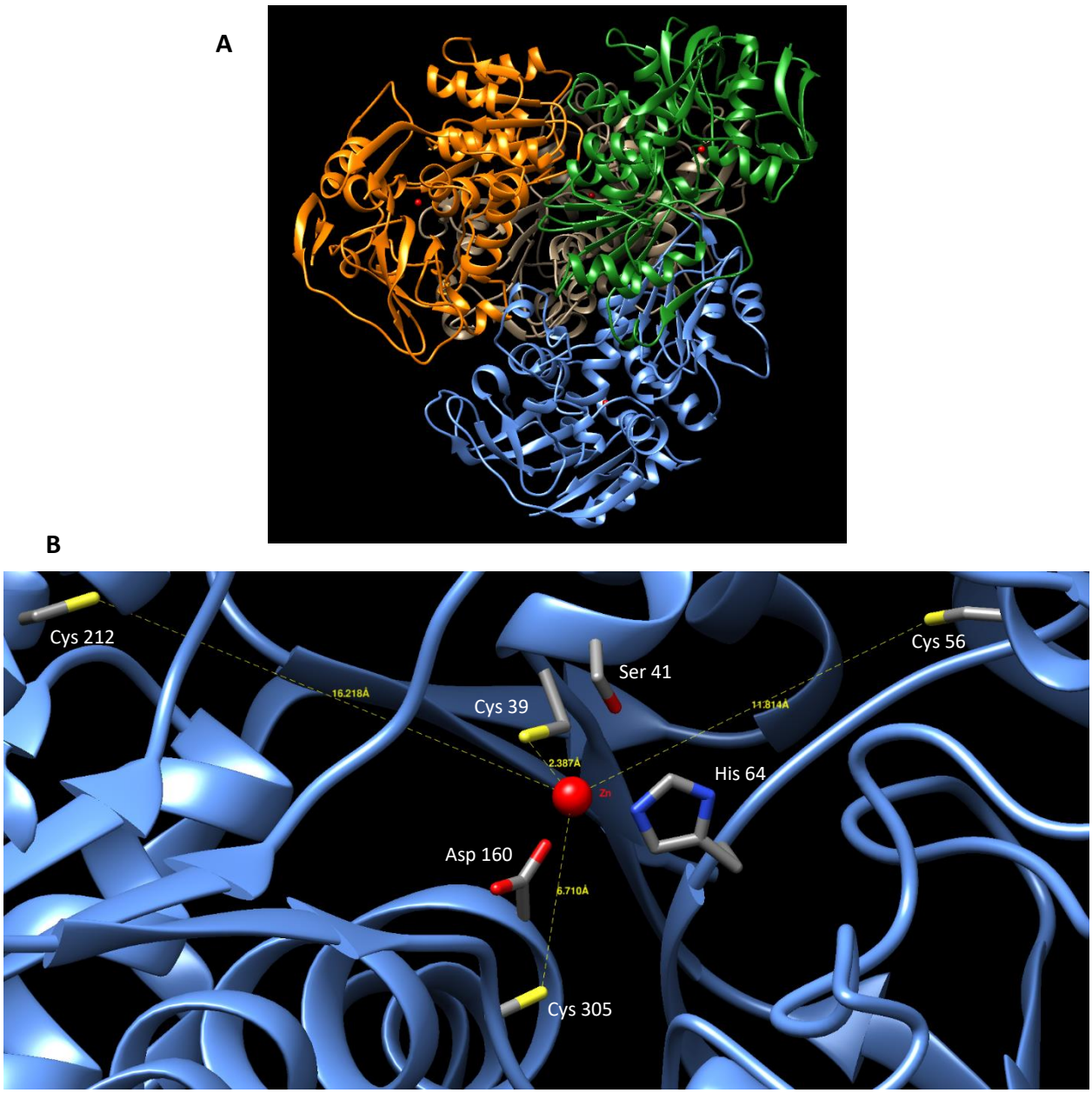


Figure 3-29: Modelled structure of TgADH

The structure of TgADH was modelled using *Thermoanaerobacter brockii* ADH as template. **A)** The modelled quaternary structure of TgADH showing the homotetrameric form. **B)** The four Cys residues along with the putative zinc-coordinating residues are shown. The distance between the Cys residues and zinc is shown in broken yellow line.

4. Discussion & Conclusion

4.1 Growth and expression of recombinant wild-type TgADH

T. guaymasensis is an obligate heterotroph and strictly anaerobic archaea that grows at 56°C to 90°C, with an optimum growth temperature of 88°C (Canganella et al., 1998). A NADP-dependent, zinc-containing alcohol dehydrogenase was characterised from *T. guaymasensis* (Ying and Ma, 2011). It was found to have many outstanding features; highest specific activity among the zinc-containing ADH family, active within a broad temperature range from 30°C to 95°C, high thermostability with a half-life of 26 hours at 95°C and stereospecificity (Ying and Ma, 2011). The TgADH encoding gene was inserted to a pET-30a vector and expressed in *E. coli*, specifically in *E. coli* BL21 Rosetta expression strains, in order to overcome codon bias (Hao, unpublished). The expression host was grown in LB medium with 0.5 mg/ml kanamycin and expression of the recombinant enzyme was induced using 0.2 mM IPTG. The recombinant enzyme was soluble, and no inclusion body was produced. The activity of recombinant TgADH did not require heat activation.

In the past the expression host has been grown in 2 x YT medium (Hao, unpublished). The only difference between LB and 2 x YT is that the latter is richer than the former. 2 x YT has 100% more yeast extract and 60% more tryptone, compared to LB medium. Most hyper/thermophilic recombinant expression vectors are grown in LB medium. For example, the *E. coli* expression host containing the recombinant ADH from *Thermococcus* strain ESI was grown and overexpressed in LB medium with ampicillin (Ying et al., 2008). The recombinant ADH from *Pyrococcus horikoshii* OT3 was also grown and expressed in LB medium with ampicillin (Sugimoto et al., 2016). It has also been reported that expression vectors such as *E. coli* grown in richer medium get subjected to higher oxidative stress from metabolizing the nutrients, leading to higher glycation levels, mutation frequency and death (Kram and Finkel, 2015).

4.2 Purification of recombinant wild-type TgADH

The CFE prepared using recombinant *E. coli* cells containing the *pET-30a-TgADH* construct was subjected to a step of heat treatment before purification using liquid chromatography. Heat treatment and denaturation is a unique method of purification that is exclusively used for purification of thermostable enzymes. This method denatures the other proteins in the mixture, which are removed by centrifugation, while the thermostable protein remains active in the supernatant. The incubation temperature and time depends on the thermostability of the protein. For example, recombinant ADH from *Thermotoga hypogea* was purified up to 60 % when heat treated at 60°C for 30 minutes (Hao, unpublished). The recombinant ADH from *Thermococcus* strain ESI was purified over 70 % after heating for 1 hour at 60°C (Ying et al., 2008). Recombinant ADH from *P. horikoshii* and recombinant xylanase from *Thermotoga thermarum* were partially purified by heating at 70°C for 20 and 30 minutes, respectively (Shi et al., 2013; Sugimoto et al., 2016). The examples are countless and would only increase in the future as more research is being focused on thermophilic organisms and due to the simplicity and effectiveness of the method.

Heat treatment step for TgADH was optimized. 1 hour at 60°C removed only 30-40 % of the unnecessary proteins (Figure 3-4). This percentage increased with increasing temperature. 10 minutes at 100°C removed the highest amount of protein of 60-65 % (Figure 3-5). Following heat treatment, the recombinant TgADH was purified using aromatic hydrophobic interaction chromatography (phenyl-sepharose). This method has previously been shown to purify the recombinant TgADH to homogeneity (Hao, unpublished), but in my experience phenyl-sepharose was unsuccessful in purifying it. In actuality, by looking at the SDS-PAGE and enzyme activity of fractions eluted from phenyl-sepharose purification (Figures 3-11 and 3-13; Tables 3-1 and 3-2), it could be inferred that either the phenyl-sepharose column or the pump/tubing of the AKTA-

FPLC machine might have harboured contaminants which may have transferred into the fractions while elution. This in turn could have reduced the specific activity of the enzyme.

4.3 Oxygen sensitivity, oxidative modification(s), and degradation of TgADH

Native and recombinant wild-type TgADH were found to lose activity under aerobic conditions, with a half-life ($t_{1/2}$) of 4 hours (Figure 3-15) (Ying and Ma, 2011; Hao, unpublished). Reports of ADH oxygen sensitivity/inactivation were almost always associated with iron-containing ADHs and not zinc-containing ADHs (Ying et al, 2007; Ying et al, 2008; Ying and Ma, 2011). The only other well-known example is the zinc-containing ADH from the mesophilic *S. cerevisiae* (YADH I). Buhner and Sund (1969) found that when YADH crystals were stored in ammonium sulfate at 4°C, the enzyme activity decreased over time. This was attributed to the oxidation of -SH groups of Cys residues. Therefore, it was hypothesised that the loss of activity in wild-type TgADH could also be due to the oxidation of Cys residues (Ying and Ma, 2011).

4.3.1 Oxygen sensitivity of Cys mutant TgADH

TgADH has four Cys residues in each subunit; Cys39, Cys56, Cys212, and Cys305 (Ying and Ma, 2011). Substitution of Cys39 to Ser caused complete inactivation of the enzyme showing that this Cys was coordinated with the zinc-ion and important for the catalytic activity. C56S mutant showed very similar catalytic properties to those of the wild-type and no improvement in oxygen tolerance (Ma and Ying, 2013). The oxygen tolerance of C212S and C305S mutants were characterized in this study.

Under anaerobic conditions both C212S and C305S mutants showed very similar enzyme activity to the recombinant wild-type (Table 3-4). However, their activities decreased over time when exposed to air. When exposed to air C305S mutant lost activity more rapidly than C212S mutant, with a half-life ($t_{1/2}$) of 1.75 hours for the former and 3.5 hours for the latter (Figures 3-16 and 3-17). The loss of activity when exposed to air behaved as a first order reaction, which was evident from the straight line observed when the natural log of activity was plotted against time (Figure 3-18). The inactivation constant (j) for each Cys mutant was deduced ($j_{C212S} = 0.141 \pm 0.03$, $j_{C305S} = 0.297 \pm 0.08$) from which the half-lives for each mutant were recalculated to be 5.15 ± 1.17 hours and 2.51 ± 0.67 hours for C212S and C305S mutants, respectively. This shows that the loss of enzyme activity when exposed to air is not due to any oxidative modifications that might/might not be occurring in residues Cys212 and Cys305. Since Cys56 has already been eliminated as a potential reason (Ma and Ying, 2013), it could be hypothesised that the loss of enzyme activity when exposed to air could be due to the oxidation of the metal coordinating cysteine, Cys39, which in turn causes the loss of zinc, thereby leading to the inactivation of the enzyme.

Kinetic studies of cysteine oxidation when treated with H_2O_2 showed a rate constant of 15.2 $M^{-1}s^{-1}$ for the formation of Cys-SOH, when the starting concentration of Cys-SH was 4 mM, at 25°C, pH 6.0 (Luo et al., 2004). Studies on oxidation of Cys232 in α1-Antitrypsin when treated with H_2O_2 showed that it had an oxidation rate constant of 7.0 $M^{-1}s^{-1}$ when the starting concentration of enzyme was 2 µM, at 25°C, pH 7.0. Cys232 oxidizes to Cys232-SOH, Cys232-SO_2H, and Cys232-SO_3H (Griffiths et al., 2002). However, it should be noted that these reactions are second order in nature and their half-lives cannot be calculated accurately. Therefore, the corresponding kinetics cannot be compared to the reaction corresponding to the loss of enzyme activity in TgADH when exposed to air.

4.3.2 Oxidative modification(s) in recombinant wild-type TgADH

MALDI-TOF mass spectrometry was used to investigate the modifications due to oxidation, if any, in recombinant wild-type TgADH. MALDI-TOF is usually used for the analyses of intact/whole proteins with an accuracy of 0.05 – 0.2 % (Lewis et al., 2000). Almost all the reported work on analysis of post-translational modifications use digested peptides of the protein. This method further increases the accuracy up to 0.01 – 0.1 % (Mann and Kelleher, 2008). However, this method could not be applied in this study for a number of reasons; 1) The TgADH samples used for mass spectrometry analysis were not pure since they were only partially purified using heat treatment method. This sample cannot be digested as this would give rise to many random mixtures of peptides from the contaminants. 2) Some studies have used a pure sample extracted from an electrophoresis gel, which is then subjected to peptide digestion. However, this cannot be done in this study since the objective was to investigate the modifications due to oxidation in an active enzyme sample. Due to these reasons it was decided to use the whole/intact recombinant TgADH for the analysis of oxidative modifications.

Since the whole/intact recombinant TgADH was used for analysis an accuracy of 0.05 – 0.2 % was expected (Lewis et al., 2000). It should be noted that for this study the lower end of this range is still too large (≈ 22 Da) for investigating the number of oxygens added under oxidative conditions. The molecular weight of an oxygen atom is 16 Da. 0.05 % accuracy of this method is greater than 16 Da, which can mask the modification and lead to false results. However, accuracy can change depending on the calibration of the instrument, matrix used, and the method used for sample application. The external calibration method used for this study showed an accuracy range of 0.0013 – 0.19 %. (0.57 – 85 Da). 0.0013 % is sufficient for the analyses of oxidative modifications in recombinant TgADH, but this is ≈ 145 times less than the higher end of the

accuracy range obtained from the external calibration (Figure 3-20). This large discrepancy could mask one or more of the expected oxidative modification(s), leading to false results. However, it should be noted that 0.0013 % of accuracy was recorded for Protein A $[M+H]^+$, which has very similar mass to the recombinant TgADH. Therefore, it was decided to use the same MALDI-TOF method that is used for the analysis of whole proteins to investigate oxidation related modifications.

It was unexpected when the measured mass of the active recombinant TgADH was 47 Da less than the reported mass. The air exposed sample was expected to show a higher measured mass compared to the active sample, due to oxidation and addition of any oxygen. However, the measured mass was 12 Da less than the reported mass (Table 3-6). The same trend was seen for TgADH treated with 1:10, 1:40, 1:160, and 1:400 mol/mol of protein subunit:H_2O_2, where the respective measured masses were less than the reported mass. The 1:1600 mol/mol protein subunit:H_2O_2 treated sample showed an increase of 360 Da. This was over the error range of the calibration making it a valid measurement. This large increase is possibly due to the oxidation of the Cys residues as well as the Met residues in the subunit, but the oxidation states of Cys and Met residues cannot be speculated due to the unreliable data. Since the difference between the measured masses of air exposed and active samples was ≈ 35 Da (Table 3-6), it could be argued that two oxygens are being added when the enzyme is exposed to air. However, due to the large discrepancy in the accuracy range and the unreliable data obtained for the H_2O_2 treated samples, this cannot be considered as conclusive evidence of any oxidative modification(s).

Cys oxidation in proteins can also be monitored using chemical methods. Disulfide bonds in a protein can be inferred using a trap/tag method combined with gel electrophoresis. First, the free thiols can be trapped/blocked using iodoacetamide (IAM). Then the disulfide bonds can be reduced

using a reducing agent like DTT. Finally, the newly exposed thiol groups can be tagged using a compound such as 4-acetamido-40-maleimidylstilbene-2,20-disulfonate (AMS) or methoxy-polyethylene glycol-maleimide (Jakob et al., 1999; Makmura et al., 2001) which would show a shift in the protein mass. This can be used to deduce the number of Cys residues that take part in disulfide bond formation. The same trap and tag method can also be combined with spectrophotometry using 5,5'-dithiobis-(2-nitrobenzoic)acid, also known as Ellman's reagent (Ellman, 1959), which can also be used to deduce the number of available -SH groups. The most popular and versatile approach to detect sulfenic acid (-SOH) is by using an enolisable 1,3-diketone, especially 5,5-dimethyl-1,3-cyclohexanedione (dimedone). This can react with sulfenic acids directly. Due to the lack of a detectable vector, many derivatives of dimedone have been developed that contain affinity probes, fluorophores or functionalisable groups such as azides and alkynes (Alcock et al., 2018). Sulfinic ($-SO_2H$) and Sulfonic ($-SO_3H$) acids are harder to detect, compared to sulfenic acids, due to their decreased reactivity. Carroll and co-workers (2015) developed a method to detect sulfinic acid, where the free thiols are first blocked and then sulfinic acids are detected using a biotin-labelled aryl nitroso compound containing an ester group.

4.3.3 Degradation of recombinant TgADH

It was investigated if oxidation causes the recombinant TgADH to degrade over time. A similar phenomenon was seen in ADH of *S. cerevisiae*, where the 141,000 Da homotetrameric enzyme dissociates into four polypeptide chains of 35,000 Da when inactivated due to the oxidation of free -SH groups (Buhner and Sund, 1969). TgADH samples were stored at various temperatures under aerobic and anaerobic conditions. As expected, the sample stored under aerobic conditions lost activity over time while the anaerobically stored samples still had over 90 % of activity after 144

hours. None of the samples showed any signs of degradation over time except the sample stored at room temperature under anaerobic condition which showed two bands at positions 43,000 Da and 42,000 Da after 144 hours of incubation (Figure 3-25). Further analysis of the sample showed that the degradation might not be due to oxidation, but due to the presence of some microorganism(s) that secrete a type of protease or peptidase. Three morphologically different colonies were identified from streak plating the sample on LB-agar. Further inoculation and growth under aerobic and anaerobic conditions proved to be successful in identifying two different microorganisms, a strict aerobe and a possible facultative anaerobe. The strictly aerobic microorganism cleaves off approximately 3,000 Da. Due to cross-contamination it wasn't possible to infer the size of the peptide cleaved by the facultative anaerobe. However, this proves that if not for external enzymatic digestion, TgADH does not undergo oxidative degradation.

4.4 Structure modelling of TgADH and analysis

In order to understand the spatial arrangement of the Cys residues in TgADH the structure of the enzyme was modelled using ADH from *Thermoanaerobacter brockii* (SMTL ID 1ped.1) as template (Korkhin et al., 1998). Cys56, Cys212, and Cys305 were located 11.814 Å, 16.218 Å and 6.710 Å from the zinc atom, respectively, while Cys39 was only 2.387 Å away from zinc (Figure 3-29). This further adds to the hypothesis that since Cys56, Cys212, and Cys305 are located further away from zinc and considering the fact that they are not coordinated with zinc, they would have no/very minimal effect on the catalytic activity of the enzyme. On the otherhand, the Cys39 is a zinc coordinating residue (Ying and Ma, 2011) and knowing that it is spatially located very close

to the metal ion further strengthns the hypothesis that the oxidation of Cys39 could possibly cause the loss of zinc, which in turn would lead to the inactivation of TgADH when exposed to air.

4.5 Conclusions

The goal of this study was twofold: 1) To purify recombinant TgADH and identify the modification(s) that occur due to oxidation when exposed to air, and 2) To characterize the oxygen sensitivity of C212S and C305S TgADH mutants. The following conclusions can be made based on experimental results obtained:

1. Growth and expression conditions for the recombinant *E. coli* host containing the *pET-30a-TgADH* construct were successfully optimized.
2. Heat treatment conditions were successfully optimized to remove 60 – 65 % of the mesophilic contaminants in the CFE when incubated at 100°C for 10 minutes, followed by centrifugation.
3. Phenyl-sepharose column used in this study failed to purify the recombinant TgADH due to technical issues. The rationale behind this behavior is unknown at this time.
4. C212S and C305S TgADH mutants did not show any increase in oxygen tolerance compared to wild-type. Therefore, taking into consideration that C56S did not show an increase in oxygen tolerance in previous studies, it can be hypothesized that the oxidation of Cys39 is the underlying cause for the loss of enzyme activity when exposed to air. It should be noted that further experiments might be required to revalidate the oxygen sensitivity of C212S and C305S as the data points shown in this study were not done in replicates.
5. Mass spectrometry results indicated that two oxygens were being added to the recombinant TgADH when exposed to air. However, due to the impure nature of the sample and the large discrepancy in the mass spectrometry method, this study was not able to conclusively identify

any oxidative modifications in recombinant wild-type TgADH. Further experiments are required to validate this phenomenon.

6. Recombinant wild-type TgADH does not undergo degradation due to oxidation, but can be degraded in the presence of certain contaminant microorganisms. Therefore, recombinant TgADH samples have to be handled and stored aseptically to eliminate the issue of contamination.

4.6 Recommendations for future study

1. The recombinant TgADH is known to be purified to homogeneity using phenyl-sepharose hydrophobic interaction chromatography. However, the phenyl-sepharose column used for this study failed to purify the enzyme due to technical issues. The column has to be either repacked with fresh medium or subjected to rigorous cleaning-in-place (CIP) to remove any precipitated protein and/or lipids. It should be noted that the recombinant TgADH has a 6 x His tag at the N-terminal. This can be used to purify the enzyme to homogeneity by using a column with nickel sepharose medium.
2. Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) can be used to quantify the number of free thiol groups in recombinant TgADH, before and after exposure to air. In order to identify the oxidative modification(s), purified recombinant enzyme sample can be subjected to MALDI-TOF mass spectrometry, under reflectron mode, after being enzymatically digested. The mass spectrometry results, along with the results from chemical quantification can be used to conclude the type of oxidative modification.
3. This study led to the hypothesis that oxidation of Cys39 of TgADH was the reason behind the loss of enzyme activity when exposed to air. This residue coordinates with zinc. Therefore, to

further validate this hypothesis the loss of zinc should be measured overtime after exposure to air.

4. Preliminary efforts to identify the contaminants that degraded recombinant TgADH showed them to be gram-negative spore formers. Further characterization of these microorganisms, including 16S rRNA analysis should be done to identify if they are novel.

5. References

Alcock, L. J., Perkins, M. V., Chalker, J. M. (2018). Chemical methods for mapping cysteine oxidation. Chemical Society Reviews, 47, 231-268.

Antoine, E., Rolland, J. L., Raffin, J.P., Dietrich, J. (1999). Cloning and overexpression in *Escherichia coli* of the gene encoding NADPH group III alcohol dehydrogenase from *Thermococcus hydrothermalis*. European Journal of Biochemistry, 264, 880-889.

Asokumar, N., Kim, S. D., Ma, K. (2018). Alcohol dehydrogenases catalyzing the production of ethanol at high temperatures. Innovative Energy and Research, 7(4). doi:10.4172/2576-1463.1000219

Auld, D.S., Bergman, T. (2008). The role of zinc for alcohol dehydrogenase structure and function. Cellular and Molecular Life Sciences, 65, 3961-3970.

Baker, P. J., Britton, K. L., Fisher, M., Esclapez, J., Pire, C., et al. (2009). Active site dynamics in the zinc-dependent medium chain alcohol dehydrogenase superfamily. Proceedings of the National Academy of Sciences, 106, 779-784.

Beauchamp, J., Gross, P. G., Vieille, C. (2014). Characterization of *Thermotoga maritima* glycerol dehydrogenase for the enzymatic production of dihydroxyacetone. Applied Environmental Microbiology, 98, 7039-7050.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry, 72, 248-254.

Buhner, M., Sund, H. (1969). Yeast alcohol dehydrogenase: -SH groups, disulfide groups, quaternary structure, and reactivation by reductive cleavage of disulfide groups. European Journal of Biochemistry. 11, 73 – 79.

Carroll, K. S., Conte, M. L., Lin, J., Wilson, M. A. (2015). A Chemical Approach for the Detection of Protein Sulfinylation. ACS Chemical Biology, 10(8), 1825–1830.

Castro, C., Millian , N. S., Garrow, T. A. (2008). Liver betaine-homocysteine S-methyltransferase activity undergoes a redox switch at the active site zinc. Archives of Biochemistry and Biophysics. 472, 26–33.

Chang, T., Yao, S. (2011). Thermophilic, lignocellulolytic bacteria for ethanol production: Current state and perspectives. Applied Microbiology and Biotechnology. 92, 13 – 27.

Chen, J. S. (1995). Alcohol dehydrogenase: Multiplicity and relatedness in the solvent-producing clostridia. FEMS Microbiology Reviews, 17, 263-273.

Dumont, S., Bykova, N. V., Khaou, A., Besserour, Y., Dorval, M., Rivoal, J. (2018) *Arabidopsis thaliana* alcohol dehydrogenase is differently affected by several redox modifications. PLoS ONE 13(9): e0204530.

Ellman, G. L. (1959). Tissue sulfhydryl groups. Archives of Biochemistry and Biophysics, 82(1), 70-77.

EMD Millipore (1998). pET30a-c(+) Vectors.

Eram, M. S., Ma, K. (2013). Decarboxylation of pyruvate to acetaldehyde for ethanol production by hyperthermophiles. Biomolecules, 3, 578 – 596.

Esposito, L., Sica, F., Raia, C. A., Giordano, A., Rossi, M., et al. (2002). Crystal structure of the alcohol dehydrogenase from the hyperthermophilic archaeon *Sulfolobus solfataricus* at 1.85 Å resolution. Journal of Molecular Biology, 318, 463-477.

Fiala, G., Stetter, K. O. (1986). *Pyrococcus furiosus* sp. nov. represents a novel genus of marine heterotrophic archaea bacteria growing optimally at 100°C. Archives of Microbiology, 145, 56-61.

Go, Y. M., Chandler, J. D., Jones, D. P. (2015). The cysteine proteome. Free Radical Biology and Medicine. 84, 227–245.

Griffiths, S. W., King, J., Cooney, C. L. (2002). The Reactivity and Oxidation Pathway of Cysteine 232 in Recombinant Human α1-Antitrypsin. The Journal of Biological Chemistry, 277(28), 25486 –25492.

Gupta, V., Carroll, K. S. (2014). Sulfenic acid chemistry, detection and cellular lifetime. Biochimica et Biophysica Acta. 1840(2014), 847–875.

Guy, J. E., Isupov, M. N., Littlechild, J. A. (2003). The structure of an alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix*. Journal of Molecular Biology, 331, 1041-10151.

Hao, L. (2009). Molecular characterization of zinc- and iron- containing alcohol dehydrogenases from anaerobic hyperthermophiles. (Masters dissertation). University of Waterloo, Waterloo, Canada.

Hattori, A., Unno, H., Goda, S., Motoyama, K., Yoshimura, T., Hemmi, H. (2015). In Vivo Formation of the Protein Disulfide Bond That Enhances the Thermostability of Diphosphomevalonate Decarboxylase, an Intracellular Enzyme from the Hyperthermophilic Archaeon *Sulfolobus solfataricus*. Journal of Bacteriology, 197, 3463-3471.

Hess, M., Antranikian, G. (2008). Archaeal alcohol dehydrogenase active at increased temperatures and in the presence of organic solvents. Applied Microbiology and Biotechnology, 77, 1003-1013

Hirakawa, H., Kamiya, N., Kawarabayashi, Y., Nagamune, T. (2004). Properties of an alcohol dehydrogenase from the hyperthermophilic archaeon *Aeropyrum pernix* k1. Journal of Bioscience and Bioengineering, 97, 202-206.

Hortin, G. L., Remalay, A. T (2006). Mass determination of major plasma proteins by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Clinical Proteomics. 2, 103-116.

Jakob, U., Muse, W., Eser, M., Bardwell, J. C. (1999). Chaperone activity with redox switch. Cell, 96(3), 341-52.

Jakubowski, H., (2016). Biochemistry online: An approach based on chemical logic. 2002. Protein Structure, Ch 2 https://employees.csbsju.edu/hjakubowski/classes/ch331/bcintro/default.html

Jin, L., Szeto, K.Y., Zhang, L., Du, W., Sun, H. (2004). Inhibition of alcohol dehydrogenase by bismuth. Journal of Inorganic Biochemistry. 98(8), 1331–1337.

Kanehisa, M., Goto, S. (2000). KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Research, 28, 27-3Kosjek, B., Stampfer, W., Pogorevc, M., Goessler, W., Faber, K., Kroutil, W. (2004). Purification and characterization of a chemotolerant alcohol dehydrogenase applicable to coupled redox reactions. Biotechnology and Bioengineering, 86, 55-62

Karas, M., Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10 000 Daltons. Analytical Chemistry. 60, 2299-2301.

Kinoshita, S., Kakizono, T., Kadota, K., Das, K., Taguchi, H. (1985). Purification of two alcohol dehydrogenases from *Zymomonas mobilis* and their properties. Applied Microbiology and Biotechnology, 22, 249–254.

Korkhin, Y., Frolow, F., Bogin, O., Peretz, M., Kalb, A. J., Burstein, Y. (1996). Crystalline alcohol dehydrogenases from the mesophilic bacterium *Clostridium beijerinckii* and the thermophilic bacterium *Thermoanaerobium brockii*: preparation, characterization and molecular symmetry. Acta Crystallogr D Biol Crystallogr, 1;52(Pt 4), 882-6.

Korkhin. Y., Kalb, A. J., Peretz, M., Bogin, O., Burstein, Y., et al. (1998). NADP-dependent bacterial alcohol dehydrogenases: Crystal structure, cofactor binding and cofactor specificity of the ADHs of *Clostridium beijerinckii* and *Thermoanaerobacter brockii*. Journal of Molecular Biology, 278, 967-981.

Kosjek, B., Stampfer, W., Pogorevc, M., Goessler, W., Faber, K., Kroutil, W. (2004). Purification and characterization of a chemotolerant alcohol dehydrogenase applicable to coupled redox reactions. Biotechnology and Bioengineering. 86, 55-62.

Kowalczyk, A., Serafin, E., Puchała, M., (2008). Inactivation of chosen dehydrogenases by the products of water radiolysis and secondary albumin and haemoglobin radicals. International Journal of Radiation Biology. 84(1), 15–22.

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

Larson, S. B., Jones, J. A., McPherson, A. (2019). The structure of an iron-containing alcohol dehydrogenase from a hyperthermophilic archaeon in two chemical states. Acta Crystallographica, 75, 217-226.

Lewis, J. K., Wei, J., Siuzdak, G. (2000). Matrix-assisted laser desorption/ionization mass spectrometry in peptide and protein analysis. In: R. A. Meyers (Ed.), 2000, Encyclopedia of Analytical Chemistry. pp. Chichester: John Wiley & Sons Ltd, 5880–5894.

Li, D., Stevenson, K. J. (1997). Purification and sequence analysis of a novel NADP (H)-dependent type III alcohol dehydrogenase from *Thermococcus* strain AN1. Journal of Bacteriology, 179, 4433-4437.

Li, D., Stevenson, K. J. (2001). Alcohol dehydrogenase from *Thermococcus* strain AN1. Methods in Enzymology, 331, 201-207.

Littlechild, J. A., Guy, J. E., Isupov, M. N. (2003). Hyperthermophilic dehydrogenase enzymes. Biochemical Society Transactions, 35, 255 – 258.

Liu, T., Belov, M. E., Jaitly, N., Qian, W., Smith, R. D. (2007). Accurate Mass Measurements in Proteomics. Chemical Reviews, 107(8), 3621–3653

Luo, D., Smith, S. W., Anderson, B. D. (2004). Kinetics and Mechanism of the Reaction of Cysteine and Hydrogen Peroxide in Aqueous Solution. Journal of Pharmaceutical Sciences, 94, 304-316.

Lynd, L. R., Wyman, C. E., Gerngross, T. U. (1999). Biocommodity engineering. Biotechnology Progress, 15(5), 777 – 793.

Ma, K., Adams, M. W. (1999). An unusual oxygen-sensitive, iron- and zinc-containing alcohol dehydrogenase from the hyperthermophilic archaeon *Pyrococcus furiosus*. Journal of Bacteriology, 181, 1163-1170

Ma, K., Adams, M. W. W. (2001). Alcohol dehydrogenases from *Thermococcus litoralis* and *Thermococcus* strain ES1. Methods in Enzymology, 331, 195-201.

Ma, K., Robb, F. T., Adams, M. W. W. (1994). Purification and characterization of NADP-specific alcohol dehydrogenase and glutamate dehydrogenase from the hyperthermophilic archaeon *Thermococcus litoralis*. Applied Environmental Microbiology, 60, 562-568

Ma, K., Tse, C. (2015). Alcohol dehydrogenases and their physiological functions in hyperthermophiles. In: F. Li, ed, 2015. Thermophilic Microorganisms. Beaverton: Roddington inc. Ch 6.

Machielsen, R., Uria, A. R., Kengen, S. W., van der Oost, J. (2006). Production and characterization of a thermostable alcohol dehydrogenase that belongs to the aldo-keto reductase superfamily. Applied and Environmental Microbiology, 72, 233-238.

Makmura, L., Hamann, M., Areopagita, A., Furuta, S., Munoz, A., Momand, J. (2001). Development of a sensitive assay to detect reversibly oxidized protein cysteine sulfhydryl groups. Antioxidants and redox signalling, 3(6), 1105-18.

Marino, S. M., Gladyshev, V. N. (2010). Cysteine function governs its conservation and degeneration and restricts its utilization on protein surfaces. Journal of Molecular Biology. 404(5), 902–916.

Martyniuk, C. J., Fang, B., Koomen, J. M., Gavin, T., Zhang, L., Barber, D. S., Lopachin, R. M. (2011). Molecular mechanism of glyceraldehyde-3-phosphatedehydrogenase inactivation by α,β-unsaturated carbonyl derivatives. Chemical Research in Toxicology, 24 (12), 2302–2311.

Men, L., Wang, Y. (2006). The oxidation of yeast alcohol dehydrogenase-1 by hydrogen peroxide *in vitro*. Journal of Proteome Research. 6, 216-225.

Miseta, A., Csutora, P. (2000). Relationship between the occurrence of cysteine in proteins and the complexity of organisms. Molecular Biology and Evolution. 17(8), 1232–1239.

Montella, C., Bellsolell, L., Perez-Luque, R., Badia, J., Baldoma, L. et al. (2005). Crystal Structure of an Iron-Dependent Group III Dehydrogenase That Interconverts L-Lactaldehyde and L-1,2-Propanediol in *Escherichia coli*. Journal of Bacteriology, 184, 4957–4966.

Moon, J., Lee, H., Park, S., Song, J., Park, M., et al. (2011). Structures of iron dependent alcohol dehydrogenase 2 from *Zymomonas mobilis* ZM4 with and without NAD^+ cofactor. Journal of Molecular Biology, 407, 413-424.

Morgan, P. E., Dean, R. T., Davies, M. J., (2002). Inhibition of glyceraldehyde-3-phosphate dehydrogenase by peptide and protein peroxides generated by singlet oxygen attack. European Journal of Biochemistry. 269, 1916–1925.

Mueller, M., Kratzer, R., Schiller, M., Slavica, A., Rechberger, G. et al. (2010). The role of Cys108 in *Trigonopsis variabilis* D-amino acid oxidase examined through chemical oxidation studies and point mutations C108S and C108D. Biochimica et Biophysica Acta. 1804(2010), 1483–1491.

National Research Council (2011). Renewable Fuel Standard: Potential Economic and Environmental Effects of U.S. Biofuel Policy. Washington, DC: The National Academies Press. https://doi.org/10.17226/13105.

Olson, D. G., Sparling, R., Lynd, L. R. (2015). Ethanol production by engineered thermophiles. Current Opinion in Biotechnology, 33, 130 – 141.

Pennacchio, A., Giordano, A., Pucci, B., Rossim M., Raia,C. A. (2010). Biochemical characterization of a recombinant short-chain NAD(H)-dependent dehydrogenase/reductase from *Sulfolobus acidocaldarius*. Extremophiles, 14, 193-204.

Petrucci, R. H., Harwood, W. S., Herring, G. E., Madura, J. (2007). General Chemistry: Principles and Modern Applications 9th Ed. New Jersey: Pearson Education Inc. 2007.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of Computational Chemistry. 25(13), 1605-1612.

Pikuta, E. V., Marsic, D., Itoh, T., Bej, A. K., Tang, J. et al. (2007). *Thermococcus thioreducens* sp. nov., a novel hyperthermophilic, obligately sulfur-reducing archaeon from a deep-sea hydrothermal vent. International Journal of Systemic and Evolutionary Biology, 57, 1612-1618.

Promega, (2018). Technical Manual-pGEM®-T and pGEM®-T Easy Vector Systems. Wisconsin.

Radianingtyas, H., Wright, P. C. (2003). Alcohol dehydrogenases from thermophilic and hyperthermophilic archaea and bacteria. FEMS Microbiology Reviews. 27, 593 – 616.

Raia, C. A., Giordano, A., Rossi, M. (2001). Alcohol dehydrogenase from *Sulfolobus solfataricus*. Methods in Enzymology, 331, 176-195.

Regalbuto, J. R. (2009). Cellulosic biofuels - Got gasoline?. Science, 325, 822-824.

Reid, M. F., Fewson, C. A. (1994). Molecular characterization of microbial alcohol dehydrogenases. Critical Reviews in Microbiology, 20(1), 13 – 56.

Rella, R., Raia, C. A., Pensa, M., Pisani, F. M., Gambacorta, A., et al. (1987). A novel archaebacterial NAD^+-dependent alcohol dehydrogenase. European Journal of Biochemistry, 167, 475-479.

Rodacka, A. (2016). The effect of radiation-induced reactive oxygen species (ROS) on the structural and functional properties of yeast alcohol dehydrogenase (YADH). International Journal of Radiation Biology. 92(1), 11–23.

Rodacka, A., Gerszon, J., Puchala, M., Bartosz, G. (2016). Radiation-induced inactivation of enzymes – Molecular mechanism based on inactivation of dehydrogenases. Radiation Physics and Chemistry. 128(2016), 112–117.

Sambrook, J., Fritsch, E. F., Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (second edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York,

Signor, L., Erba, E. B. (2013). Matrix-assisted Laser Desorption/Ionization Time of Flight (MALDI-TOF) Mass Spectrometric Analysis of Intact Proteins Larger than 100 kDa. Journal of Visualized Experiments. 79, e50635, doi:10.3791/50635

Singhal, N., Kumar, M., Kanaujia, P. K., Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. Frontiers in Microbiology. 6:791. doi: 10.3389/fmicb.2015.00791

Stekhanova, T. N., Mardonov, A. V., Bezudnova, E. Y., Gumerov, V. M., Ravin, N. V., et al. (2010). Characterization of a thermostable short-chain alcohol dehydrogenase from the hyperthermophilic archaeon *Thermococcus sibiricus*. Applied and Environmental Microbiology, 76, 4096-4098.

Stetter, K. O. (2006). Hyperthermophiles in the history of life. Philosophical Transactions Of The Royal Society Of Biology.361, 1837 – 1843.

Taylor, M. P., Eley, K. L., Martin, S., Tuffin, M. I., Burton, S. G., Cowan, D.A. (2009). Thermophilic ethanologenesis: Future prospects for second-generation bioethanol production. Trends in Biotechnology, 27, 398 – 405.

Tse, C., Ibrahim, N. E., Ma, K. (2017). Characterization of an alcohol dehydrogenase of the hyperthermophilic archaeon *Hyperthermus butylicus*. Journal of Applied Microbiology and Biochemistry, 2(1:2).

Tse, C., Ma. K. (2016). Growth and metabolism of extremophilic microorganisms. In: P.H. Rampelotto ed, 2016. *Biotechnology of Extremophiles*, Grand Challenges in Biology and Biotechnology 1, DOI 10.1007/978-3-319-13521-2_1

van der Oost, J., Voorhorst, W. G., Kengen, S. W., Geerling, A., Wittenhorst, V., et al. (2001). Genetic and biochemical characterization of a short-chain alcohol dehydrogenase from a hyperthermophilic archaeon *Pyrococcus furiosus*. European Journal Biochemistry, 268, 3062-3068.

Vitale, A., Thorne, N., Lovell, S., Battaile, K. P., Hu, X., et al. (2013). Physicochemical characterization of a thermostable alcohol dehydrogenase from *Pyrobaculum aerophilum*. PLoS ONE, 8, e63828.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research, 46, W296-W303.

Willies, S., Isupov, M., Littlechild, J. (2010). Thermophilic enzymes and their applications in biocatalysis: A robust aldo-keto reductase. Environmental Technology, 31, 1159-1167.

Wu, X., Zhang, C., Orita, I., Imanaka, C., Fukui, T. (2013). Thermostable alcohol dehydrogenase from *Thermococcus kodakarensis* KOD1 for enantioselective bioconversion of aromatic secondary alcohols. Applied Environmental Microbiology, 79, 2209-2217.

Yanai, H., Doi, K., Ohshima, T. (2009). *Sulfolobus tokodaii* ST0053 produces a novel thermostable, NAD-dependent medium chain alcohol dehydrogenase. Applied Environmental Microbiology, 75, 1758- 1763.

Yang, K., Wang, S. W., Woo, H. A., Hwang, S. C., Chae, H. Z., Kim, K., Rhee, S. G. (2002). Inactivation of human peroxiredoxin I during catalysis as the result of the oxidation of the catalytic site cysteine to cysteine-sulfinic acid. Journal of Biological Chemistry. 277(41), 38029–38036.

Ying, X., Grunden, A. M., Nie, L., Adams, M. W., Ma, K. (2009). Molecular characterization of the recombinant iron-containing alcohol dehydrogenase from the hyperthermophilic archaeon, *Thermococcus* strain ES1. Extremophiles, 13, 299-311.

Ying, X., Ma, K. (2011). Characterization of a zinc-containing alcohol dehydrogenase with stereoselectivity from the hyperthermophilic archaeon *Thermococcus guaymasensis*. Journal of Bacteriology, 193, 3009-3019.

Ying, X., Wang, Y., Badiei, H. R., Karanassios, V., Ma, K. (2007). Purification and characterization of an iron-containing alcohol dehydrogenase in extremely thermophilic bacterium *Thermotoga hypogea*. Archives of Microbiology, 187, 499 – 510.

Yoon, S. Y., Noh, H. S., Kim, E. H., Kong, K. H. (2002). The highly stable alcohol dehydrogenase of *Thermomicrobium roseum*: purification and molecular characterization. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology. 132, 415-422.

Zaldivar, J., Nielsen, J., Olsson, L. (2001). Fuel ethanol production from lignocellulose: A challenge for metabolic engineering and process integration. Applied Microbiology and Biotechnology, 56, 17 – 34.

Appendix

List of characterised hyperthermophilic ADHs, to date, and their biophysical and catalytical properties

Organism	Growth temp[a] (°C)	Cofactor/ metal	Type	Sub-unit (kDa α_n)	Opt. temp[b] (°C)	Thermo-stability [$t_{1/2}$, h[c] (at °C)]	Oxidation of alcohols			Reduction of aldehydes/ketones			Reference
							Opt. pH[d]	App. K_m (mM)[e]	App. k_{cat} (s^{-1})	Opt. pH[f]	App. K_m (mM)[g]	App k_{cat} (s^{-1})	
Thermococcus litoralis	88	NADP/Fe	III	48 (α_4)	85	5 (85), 0.3 (96)	8.8	11.0	26.0	nd	0.4	nd	Ma et al., 1994; Ma and Adams, 2001
Thermococcus. ES-1	91	NADP/Fe	III	46 (α_4)	>95	35 (85), 4 (95)	10.3-10.5	10.4	84.0	7	1.0	23.6	Ying et al., 2009
Thermococcus kodakarensis KOD1	85	NADP/Fe	III	31 (α1)	90	16.2 (85), 4.5 (95)	9.0	61.3 (meso-2,3-butanediol)	43.8	3	30.2 (acetoin)	13.7	Wu et al., 2014
Thermococcus zilligii	75-80	NADP/Fe	III	46 (α4)	85	0.27 (80)	6.8-7.0	10.0	nd	nd	nd	nd	Li and Stevenson, 1997; Li and Stevenson, 2001

Thermococcus hydrothermalis	85	NADP/Fe	III	45 (α2/α4)	80	0.5 (80), 2.5 (70)	10.5	2.0 (benzyl-alcohol)	23.0	7.5	0.01 (benzaldehyde)	1.7	Chen, 1995
Thermococcus thioreducens	83-85	NADP/Fe	III	41.5	nd	nd	nd	nd	nd	nd	nd	nd	Pikuta et al., 2007; Larson at el., 2019
Thermotoga hypogea	70	NADP/Fe	III	40 (α2)	>95	10 (70), 2 (90)	11.0	1.9 (1-butanol)	48	8.0	0.45 (butyraldehyde)	14	Ying et al, 2007
Pyrococcus furiosus (AdhA)	100	NADP/ non-metal	II	26 (α2/α4/ α6)	90	150 (80), 22.5 (90), 0.42 (100)	10-11	60.8 (2-pentanol)	18.2	7.5	3.3 (pyruv-aldehyde)	13.8	Fialla and Stetter, 1986; van der Oost et al., 2001
Pyrococcus furiosus (AdhC)	100	NADP/Zn/ Fe		48 (α6)	>90	160 (85), 7(95)	9.4-10.2	29.4	19.2	nd	0.2	5.5	Fialla and Stetter, 1986; Ma and Adams, 1999
Pyrococcus furiosus (AdhD)	100	NAD/ non-metal	II	32 (α1)	>100	2.16 (100)	8.8	86.8 (2,3-butanediol)	60.7	6.1	6.5 (acetoin)	11.7	Fialla and Stetter, 1986; Machielsen et al., 2006
Sulfolobus acidocaldarius (Adh1)	80	NAD/ non-metal	II	29 (α4)	75	0.5 (90)	8.2	6.4 ((S)-indanol)	13.7	5.1	0.7 (isatin)	22.0	Antoine et al., 1999; Pennacchio et al.., 2010

Sulfolobus acidocaldarius (Adh2)	80	NAD/ non-metal	II	27 (α4)	80	0.5 (88)	10	0.8 (isoborneol)	16.6	5	0.1 (2,2-dichloro-acetophenone)	0.6	Chen, 1995; Antoine et al., 1999
Thermococcus sibiricus	78	NADP/ non-metal	II	26.2 (α2)	>100	2 (90), 1 (100)	10.5	54.4 (D-xylose)	0.7	7.5	17.5 (pyruval-dehyde)	2.0	Stekhanova et al., 2010
Sulfolobus tokodaii	80	NAD/Zn	I	38 (α4)	85	2 (70)	10.5	9.1	0.8	7	29.7	10.2	Yanai et al., 2009
Sulfolobus solfataricus 1617		NAD/Zn	I	37.5 (α4)	>95	5 (70), 3 (85)	8.8-9.6	0.7	0.5	8.8-9.6	nd	nd	Raia et al., 2001; Esposito et al., 2002; Pennacchio et al., 2010
Sulfolobus solfataricus MT4		NAD/Zn	I	35.5 (α2)	>95	20 (60), 5 (70)	8.5	0.26	1.8	7.5	0.0025 (anisalde-hyde)	0.9	Rella et al., 1987
Thermococcus guaymasensis	88	NADP/Zn	I	40 (α4)	>95	24 (95), 70 (80)	10.5	0.4 (2-butanol)	833	7.5	0.2 (diacetyl)	202.0	Ying and Ma, 2011
Thermotoga maritima	80	NAD/Zn	I	39.7 (α4)	80	7 (50)	7.9	39.0 (glycerol)	15.8	6	30.0 (DHA)	8.8	Willies et al., 2010; Beauchamp et al., 2014

Aeropyrum pernix	90-95	NAD/Zn	I	39.5 (α4)	90	0.5 (90)	10.5	13.7	0.23	8	nd	nd	Hirakawa et al., 2004
Pyrobaculum aerophilum	100	NADP/Zn	I	38.5 (α4)	90	nd	na	na	na	7.5	2.4 (α-tetralone)	1.7×10^2	Vitale et al., 2013
Hyperthermus butylicus	95-106	NAD/Zn	I	33 (α4)	60	3 (60)	8.5	1.7 (butanol)	4.3	5.0	0.4 (propanal)	6.6	Tse et al., 2017

[a]Optimal temperature for organism growth

[b]Optimal temperature for enzyme activity

[c]Thermostability expressed as $t_{1/2}$, which is the time required to decrease 50% of its activity at the specified temperature

[d]Optimal pH for alcohol oxidation

[e]Apparent *K*m values with ethanol, unless specified

[f]Optimal pH for aldehyde/ketone reduction

[g]Apparent *K*m values with acetaldehyde, unless specified

Opt.=optimal; App.=apparent; h=hour; Ref.=reference; nd=not determined; na=not applicable