BIOCHEMICAL AND MOLECULAR STUDIES OF TRANSKETOLASE FROM RHODOBACTER SPHAEROIDES AND ITS INACTIVATION BY OXYGEN

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By
Cedric E. Bobst, B.S.

*****

The Ohio State University
2004

Dissertation Committee:
Professor F. Robert Tabita, Adviser
Professor Mark P. Foster
Professor Russ Hille
Professor Richard P. Swenson

Approved by

Adviser
Ohio State Biochemistry Program
ABSTRACT

Transketolase is a ubiquitous thiamin diphosphate (ThDP)-dependent enzyme that plays an important role in the oxidative pentose phosphate pathway (OPPP) of virtually all organisms; this enzyme also functions as a key catalyst in the Calvin-Benson-Bassham (CBB) reductive pentose phosphate cycle of autotrophic organisms. Most organisms maintain multiple isoforms of transketolase within their genome; however, all reported research comparing different forms of this enzyme within the same organism originated from genetic and cellular biological studies. In addition, most biochemical studies on transketolase have been focused on eukaryotic or OPPP forms of the enzyme. We therefore set out to biochemically characterize both isoforms of transketolase from the non-sulfur purple bacterium *Rhodobacter sphaeroides*. The OPPP (*tkt*) and CBB (*cbbT*) transketolase genes were over-expressed in *E. coli* and recombinant proteins were purified to homogeneity. Techniques applied to compare these enzymes included substrate kinetics, molecular modeling, and stability studies. Biochemical differences suggest that each enzyme has a preferred environment for catalysis. Similar to the *E. coli* enzyme, both enzymes displayed inactivation in the presence of oxygen. Characterization of the inactivated enzyme was performed in order to better understand this type of inactivation. In addition to kinetic studies, spectroscopic methods such as circular dichroism and fluorometry, suggested that inactivation occurred in an irreversible
manner and was accompanied by a small conformational change. To investigate the role of a loosely conserved cysteine near the active site of CbbT, this residue was targeted for site-directed mutagenesis. Various methods of quantifying ThDP binding were applied, establishing an important function for this cysteine in the binding and orientation of this cofactor. However, no link in oxygen inactivation could be established for this residue. To further probe the conformational changes that occur upon oxygen inactivation, the kinetics of hydrogen/deuterium exchange was studied utilizing Fourier transform ion cyclotron resonance mass spectrometry. Though confident identification of the CbbT peptides produced by pepsin digestion was not possible based on \( m/z \) values alone, differences in exchange rates were observed for several peptides from oxygen inactivated CbbT. The magnitudes of the changes were small, indicating a change in protein dynamics rather than a global conformational change.
To my mother and father, Elizabeth and Albert
ACKNOWLEDGMENTS

I wish to express my gratitude to all those that contributed to making this work an enjoyable and fulfilling experience. I thank my adviser Bob Tabita for his superb guidance and generous support over the years. I am also thankful to the members of my dissertation committee, Mark Foster, Russ Hill, and Richard Swenson.

I thank the members of the Tabita laboratory who I have worked with for providing a comfortable and intellectually stimulating environment. Several members of the lab have played a particularly rich role in my development, selflessly sharing their knowledge and assistance: Ki-Sook Yoon for distilling his vast knowledge of protein purification techniques; Janet Gibson in developing my molecular biology skills and project guidance throughout; Jon-David Sears as my rotation mentor in the lab and later as director of the fermentation facility; Tom Hanson for his critical analysis and insightful discussion.

I thank Professor Michael Freitas and members of his lab for providing guidance and instrumentation to conduct the mass spectrometry utilized in my research, Donald Snavely for teaching me the basics of operating the instrument, Liwen Zhang for her help in setting up the LCMS experiments, and Mitchell Meade for his hand in maintaining the instrument.
VITA

April 12, 1975 Born - Cincinnati, Ohio.

1995 Undergraduate Summer Research Fellowship, American Heart Association.

1995 Undergraduate Summer Research Fellowship, American Society of Pharmacology and Experimental Therapeutics.

1997 B.S. Chemistry, University of Cincinnati.

1997-present Graduate Research Associate, The Ohio State University.

PUBLICATIONS


FIELD OF STUDY

Major Field: Biochemistry
TABLE OF CONTENTS

Abstract ........................................................................................................... ii

Dedication ........................................................................................................ iv

Acknowledgments ............................................................................................ v

Vita ................................................................................................................... vi

List of Tables .................................................................................................. xi

List of Figures .................................................................................................. xii

List of Abbreviations ....................................................................................... xv

Chapters:

1. Introduction .................................................................................................... 1

   1.1 Transketolase ............................................................................................ 1

       1.1.1 Background ...................................................................................... 1
       1.1.2 Structure .......................................................................................... 4
       1.1.3 Thiamine Diphosphate and Catalysis ................................................. 7
       1.1.4 Biological Significance ....................................................................... 12
       1.1.5 Industrial Applications ....................................................................... 14

2. Characterization of Transketolase Isozymes from *Rhodobacter sphaeroides* …. 16

   2.1 Introduction ............................................................................................... 16
   2.2 Materials and Methods ............................................................................. 18
       2.2.1 General Chemicals ............................................................................ 18
       2.2.2 Cloning of the Transketolase Genes .................................................... 18
       2.2.3 Over-Expression of Tkt and CbbT ...................................................... 22
2.2.4 Growth of *Rhodobacter sphaeroides* ........................................ 22
2.2.5 Purification of CbbT ................................................................. 23
2.2.6 Purification of Tkt ................................................................. 25
2.2.7 SDS and Native PAGE of Tkt and CbbT ................................. 26
2.2.8 Transketolase Activity Assay ................................................. 27
2.2.9 Determining Optimal pH ......................................................... 27
2.2.10 Kinetic Analysis of CbbT and Tkt ......................................... 28
2.2.11 Molecular Modeling of CbbT and Tkt .................................... 28
2.2.12 Synthesis of Xu5P ................................................................. 29

2.3 Results and Discussion ............................................................. 29
2.3.1 Cloning of CbbT and Tkt into pET11a .................................... 29
2.3.2 Over-Expression of CbbT and Tkt Genes in *E. coli* .................. 33
2.3.3 Purification of CbbT ................................................................. 34
2.3.4 Purification of Tkt ................................................................. 37
2.3.5 Enzyme Stability ................................................................. 41
2.3.6 Oxygen-Mediated Inactivation ............................................... 45
2.3.7 Thermal Degradation of CbbT and Tkt ................................... 48
2.3.8 Transketolase Optimal pH ..................................................... 54
2.3.9 Kinetics ................................................................................ 57
2.3.10 Applications of Transketolase for Substrate Preparations ..... 60
2.3.11 Transketolase Expression in *R. sphaeroides* ......................... 64

2.4.1 Conclusions ............................................................................ 68

3. Oxygen Inactivation of Transketolase ............................................. 74

3.1 Introduction .............................................................................. 74
3.2 Materials and Methods ............................................................ 76
   3.2.1 Low Oxygen Fermentations ................................................. 76
   3.2.2 Anaerobic Over-Expression in Enhanced Media ....................... 77
   3.2.3 Time Course of Inactivation ................................................. 78
   3.2.4 Circular Dichroism ............................................................. 78

5.1 Introduction ............................................................................. 136
5.2 Materials and Methods .............................................................. 140
  5.2.1 Preparation of Inactivated CbbT .......................................... 140
  5.2.2 Proteolytic Digestion of CbbT ............................................ 141
  5.2.3 HD Exchange Conditions .................................................. 142
  5.2.4 LCMS ............................................................................ 142
  5.2.5 FT-ICR MS Data Collection and Analysis ............................ 145
  5.2.6 HD Exchange Analysis and Rate Calculations ...................... 146
5.3 Results and Discussion ............................................................ 147
  5.3.1 Preparation of Oxygen-Inactivated CbbT ............................. 147
  5.3.2 Preparation of Peptidase-Generated Peptides of CbbT .......... 150
  5.3.3 Assignment of Peptidase-Generated Peptides of CbbT .......... 154
  5.3.4 Online and Offline MS Analysis of HD Exchange ................. 161
  5.3.5 Analysis of HD Exchange Rates by LCMS .......................... 164
5.4 Conclusions ............................................................................ 192

6. Summary and Future Direction .................................................. 196

References .................................................................................... 204
LIST OF TABLES

2.1 Location of the nucleotides corrected by sequencing .................................. 32
2.2 CbbT purification table .................................................................................. 37
2.3 Tkt purification table ..................................................................................... 41
2.4 Kinetic constants calculated for CbbT at three separate pH values .......... 58
2.5 Comparison of catalytic properties of CbbT and Tkt with transketolases from
   various sources ................................................................................................. 59
3.1 Gas flows and measured parameters at variable oxygen levels in 2 l fermentor
   vessels ............................................................................................................ 86
4.1 Primers used for mutagenesis at residue C160 ............................................ 101
4.2 Kinetic properties of CbbT and C160 mutant proteins ................................. 114
4.3 CbbT and C160D fluorometrically determined ThDP dissociation constants ..... 118
4.4 Enzymatically determined affinity for ThDP of CbbT and C160 mutant proteins .. 121
4.5 Fluorometrically determined binding rates for ThDP of CbbT and C160D ....... 126
5.1 Assignment of 20 of the 53 well-resolved tryptic CbbT peptides by the offline
   method ........................................................................................................... 157
5.2 Assignment of 22 of the 47 well-resolved tryptic CbbT peptides by LCMS ....... 159
5.3 Comparison of HD exchange measured online versus offline ...................... 162
5.4 Exchange rate constants for 38 peptides of active and inactive CbbT .......... 184
LIST OF FIGURES

1.1 Scheme displaying the role of transketolase in the pentose phosphate pathway, the links to glycolysis, and aromatic amino acid biosynthesis ........................................ 3

1.2 Scheme emphasizing the role of transketolase in the Calvin-Benson-Bassham cycle of CO2 fixation ............................................................... 2

1.3 Higher order structure of transketolase .................................................. 5

1.4 Structure of thiamine diphosphate ......................................................... 8

1.5 The reaction mechanism of transketolase .............................................. 10

1.5 Reaction scheme of transketolase ....................................................... 15

2.1 Arrangement of the cbb gene cluster containing the transketolases cbbT and the region containing tkt in \textit{R. sphaeroides} ........................................... 17

2.2 Purification of CbbT, SDS-PAGE .......................................................... 36

2.3 Purification of Tkt, SDS-PAGE ............................................................. 39

2.4 Sulfate effect on CbbT activity versus time ........................................... 43

2.5 Oxygen inactivation of transketolase from \textit{R. sphaeroides} ....................... 46

2.6 Protection of thermal cleavage of CbbT mutant C160A by alkylation .......... 50

2.7 Thermal degradation of CbbT ................................................................. 51

2.8 Activity levels at different pH values for CbbT and Tkt ........................... 55

2.9 Scheme showing the synthesis of Xu5P ............................................... 61

2.10 Plot of HCl consumption during the synthesis of Xu5P ......................... 63

2.11 Preferential CbbT accumulation under autotrophic conditions in \textit{R. sphaeroides} .. 66
2.12 Molecular models of CbbT and Tkt displaying the electrostatic potential mapped to
the surface ................................................................. 70

3.1 CbbT activity under variable aeration conditions ............................................. 82
3.2 Growth curves of *E. coli* 2 l fermentations cultures at variable O₂ levels ........ 84
3.3 Time course inactivation of CbbT in the presence of argon vs. O₂ .................... 88
3.4 Time course inactivation of CbbT under variable O₂ ...................................... 89
3.5 SDS PAGE analysis of CbbT samples incubated under argon or O₂ .............. 90
3.6 Fluorescence emission spectra of O₂ inactivated CbbT ............................... 92
3.7 CD spectra of CbbT incubated under a headspace of either N₂ or air .............. 93
3.8 Urea denaturing curves for active and inactivated CbbT ............................. 96

4.1 Molecular models of the C160 mutants of *R. sphaeroides* CbbT .................. 110
4.2 SDS/PAGE analysis of C160E through different stages of purification .......... 113
4.3 Activity versus pH plots for CbbT and three C160 mutant proteins ................. 115
4.4 Fluorometric determination of the K_D for ThDP ...................................... 117
4.5 Enzymatic determination of the K_m for ThDP ....................................... 120
4.6 Hill Plot of C160A reconstitution data .................................................. 122
4.7 Fluorescence reconstitution kinetics of CbbT and C160D ........................... 124
4.8 Circular dichroism spectra of apo and holo CbbT and C160 mutant proteins .... 128
4.9 Time course inactivation of CbbT and the C160D mutant ........................... 131
4.10 Region of residue C160 and part of the cofactor ThDP overlaid with the electron
density map ....................................................................................... 133

5.1 Scheme of LC-MS setup utilized for measurement of HD exchange ............ 144
5.2 Preparation of CbbT samples for analysis by mass spectrometry ............... 149
5.3 Effect of additives on CbbT digestion by pepsin ................................. 152
5.4 Chromatographic separation of a CbbT pepsin digest .......................... 165
5.5 Mass spectra demonstrating HD exchange raw data ............................. 169
5.6 Representative peptides displaying similar deuterium exchange kinetics ...... 172
5.7 CbbT peptides displaying different deuterium exchange kinetics .............. 176
5.8 CbbT peptides with potentially different deuterium exchange kinetics ......... 181
5.9 Peptide Data displaying similar exchange due to large standard deviation ...... 187
5.10 Mapping peptides to the CbbT structure ........................................... 189

6.1 Sulfate location in the E. coli TktA structure solved by crystallography ....... 200
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>Adenosine 5’-triphosphate</td>
</tr>
<tr>
<td>BME</td>
<td>β-Mercaptoethanol</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>CBB</td>
<td>Calvin-Benson-Bassham</td>
</tr>
<tr>
<td>CD</td>
<td>Circular Dichroism</td>
</tr>
<tr>
<td>D</td>
<td>Deuterium</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic Acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleotide 5’-triphosphate</td>
</tr>
<tr>
<td>dO₂</td>
<td>Dissolved Oxygen</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>E4P</td>
<td>Erythrose 4-phosphate</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic Acid</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray Ionization</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast Protein Liquid Chromatography</td>
</tr>
<tr>
<td>F6P</td>
<td>Fructose 6-phosphate</td>
</tr>
<tr>
<td>FTICR</td>
<td>Fourier Transform Ion Cyclotron Resonance</td>
</tr>
<tr>
<td>Gnd-HCl</td>
<td>Guanidinium Hydrochloride</td>
</tr>
<tr>
<td>HD</td>
<td>Hydrogen/Deuterium</td>
</tr>
<tr>
<td>HEPES</td>
<td>(N-[2-hydroxyethyl]piperazine-N’-[2-ethanesulfonic acid])</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>IAM</td>
<td>Iodoacetamide</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LC-MS</td>
<td>Liquid Chromatography-Mass Spectroscopy</td>
</tr>
<tr>
<td>MeOH</td>
<td>Methanol</td>
</tr>
<tr>
<td>MOPS</td>
<td>4-Morpholinepropanesulfonic acid</td>
</tr>
<tr>
<td>MS/MS</td>
<td>Tandem Mass Spectroscopy</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>QTOF</td>
<td>Quadrupole Time-of-Flight</td>
</tr>
<tr>
<td>R5P</td>
<td>Ribose 5-phosphate</td>
</tr>
<tr>
<td>ROS</td>
<td>Reactive Oxygen Species</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulfate</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic Acid</td>
</tr>
<tr>
<td>ThDP</td>
<td>Thiamine Diphosphate</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Xu5P</td>
<td>Xylulose 5-phosphate</td>
</tr>
<tr>
<td>WK</td>
<td>Wernicke-Korsakoff</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 TRANSKETOLASE

1.1.1 BACKGROUND

Thiamine diphosphate (ThDP)-dependent enzymes play a key role in carbohydrate metabolism. Members of this family of proteins include transketolase, pyruvate decarboxylase, pyruvate oxidase, acetolactate synthase, acetohydroxyacid synthase, α-ketoglutarate dehydrogenase, 1-deoxyxylulose-5-phosphate synthase and benzoylformate decarboxylase. Transketolase is a ubiquitous enzyme that plays an important role in the oxidative pentose phosphate pathway (Figure 1.1) of virtually all organisms; this enzyme also functions as a key catalyst in the Calvin-Benson-Bassham (CBB) reductive pentose phosphate cycle of autotrophic organisms (Figure 1.2). First isolated from yeast in 1953 (1), it was recognized as an important enzyme in metabolic reactions, creating interest in its transferase mechanism. Transketolase has since been characterized from a number of sources, with a more intensive focus on the Saccharomyces cerevisiae (TkI), E. coli (TktA), and human enzymes. Transketolase catalyzes the transfer of a two-carbon dihydroxyethyl group from a ketose donor substrate to an aldose acceptor substrate in a stereospecific manner. Most transketolases
can utilize substrates of variable backbone lengths (C$_3$-C$_7$), making this an ideal enzyme for use in synthetic organic chemistry (2, 3).

Figure 1.1. Scheme displaying the role of transketolase in the pentose phosphate pathway (I) and the links to glycolysis (II) and aromatic amino acid biosynthesis (III). Reactions catalyzed by transketolase are shown as **thick** arrows; other enzymatic steps are displayed as thin arrows; dashed lines indicate more than one step is involved in the reaction. The abbreviated substrates are as follows: G6P, glucose 6- phosphate; F6P, fructose 6- phosphate; G3P, glyceraldehyde 3- phosphate; PEP, phosphoenolpyruvate; Ru5P, ribulose 5- phosphate; Xu5P, xylulose 5- phosphate; R5P, ribose 5- phosphate; S7P, sedoheptulose 7- phosphate; E4P, erythrose 4- phosphate.
**Figure 1.2.** Scheme emphasizing the role of transketolase in the Calvin-Benson-Bassham cycle of CO₂ fixation. Reactions catalyzed by transketolase are shown as **thick** arrows, other enzymatic steps are displayed as thin arrows; dashed lines indicate more than one enzyme is involved in the reaction. The abbreviated substrates are as follows: F6P, fructose 6-phosphate; G3P, glyceraldehyde 3-phosphate; Ru5P, ribulose 5-phosphate; Xu5P, xylulose 5-phosphate; R5P, ribose 5-phosphate; S7P, sedoheptulose 7-phosphate; E4P, erythrose 4-phosphate; RuBP, ribulose 1,5-bisphosphate; 3PGA, 3-phosphoglycerate; 1,3bisPGA, 1,3-bisphosphoglycerate.
1.1.2 STRUCTURE

The x-ray crystal structure of yeast TkI1 represents the first solved structure of a ThDP-dependent enzyme and demonstrated that each transketolase subunit was comprised of three distinct domains. The first two were named for their interaction with ThDP and consist of a pyrimidine binding domain (Pyr) and a pyrophosphate binding domain (PP). The third domain was simply labeled the C-terminal domain (4) (Figure 1.3). Two molecules of ThDP bind in clefts located between the two subunits to form the active holoenzyme, containing two chemically equivalent active sites (Figure 1.3).

Previously, sequence analysis of a variety of ThDP-dependent enzymes identified a ThDP binding motif comprised of 30 amino acids, composing a GDGX_{26}D motif. The predicted secondary structure, a βαβ fold, resembled the fingerprint of a DNA binding fold (5). The crystal structure demonstrated that this conserved motif was different from the DNA binding fold, yet was responsible for various protein-ThDP contacts. This motif has since been termed the ThDP binding motif (6) and has been identified in all known ThDP-dependent enzymes. A number of interacting side-chains were identified that stabilize the dimer in the subunit-subunit interface of the Pyr and PP domains.

However, the C-terminal domain still has no recognized function in the yeast protein (7) or the enzyme from maize (8). Transketolase structures which have been determined by x-ray crystallography also include the *E. coli* enzyme; though the pdb coordinates have been made available, no supporting information has been published (9).
Figure 1.3. Higher order structure of transketolase. (A) Subunit domains demonstrated by the yeast structure (pdb ID 1GPU), the two subunits are shaded differently from top to bottom: PP, Pyr, and C. (B) CbbT theoretical model. Ribbon structure showing the subunit arrangement in the dimeric holoenzyme; the domains are arranged from top to bottom: PP, Pyr, and C. (C) The same as in Figure A, but after applying a molecular surface, the location of the substrate channel leading to the solvent accessible C2 of ThDP was revealed. (D) Same model as in Figure A following a 90° rotation around an axis parallel with the page, yielding an overhead view of the two active sites with the Pyr domain closest to the reader. In all figures, ThDP contains a hydroxyethyl intermediate at the C2 position; the location of the cofactor is given by arrows.
Figure 1.3
1.1.3 THIAMINE DIPHOSPHATE AND CATALYSIS

Thiamine diphosphate, a phosphorylated form of vitamin B\textsubscript{1}, is comprised of a thiazole and pyrimidine moiety (Figure 1.4 A) and plays a central role in catalysis of ThDP-dependent enzymes. It is capable of performing enzyme-like reactions in free solution, albeit at increased pH and at a much slower rate (10). When bound in its active site, ThDP adopts a unique strained ‘V’ conformation (Figure 1.4 B), which allows for the generation of the active carbanion at the C2 carbon of the thiazolium ring in the first step of catalysis (11).
**Figure 1.4.** Structure of thiamine diphosphate (ThDP) cofactor of transketolase. (A) Two-dimensional structure displaying the numbering of the ring atoms that compose the pyrimidine and thiazole moieties of the molecule. (B) Picture displaying a three-dimensional rendering of ThDP to demonstrate the V-shaped structure adopted between the two cyclic moieties upon binding in an enzyme’s active site. The structure was extracted from the pdb coordinates of the *E. coli* transketolase enzyme (PDB ID: 1QGD). The reactive anion forms at the C2 atom of the thiazole moiety, indicated by an arrow.
Current evidence supports that this V-shaped structure, as well as anion formation by deprotonation at the thiazole C2 carbon of ThDP, are common properties amongst all ThDP enzymes studied to date (12). The rate of deprotonation at C2 is believed to be a limiting factor in catalysis by pyruvate decarboxylase and transketolase, which display rates of approximately 1 and 61 sec⁻¹ respectively, as determined by NMR (13). It is generally accepted that the deprotonation is driven by the 4’-amino group of ThDP following tautomerization to an imino form. The proposed reaction scheme of transketolase is shown in Figure 1.5. A detailed kinetic analysis of the transketolase from *E. coli* (14), demonstrated the reaction is consistent with a Ping Pong Bi Bi mechanism (15). Essential residues important to the proposed mechanism of catalysis, subsequently implicated x-ray by crystallography (16-18) were later confirmed via site-directed mutagenesis (19) and NMR experiments (20).
Figure 1.5. The reaction mechanism of transketolase. The reaction involves the initial formation of an anion at the C2 position of the thiazole moiety of ThDP. Nucleophilic attack by the anion on the substrate forms a covalent adduct. The product is then released by cleavage of a carbon-carbon bond induced by proton abstraction. Following the release of the first product, this mechanism is repeated a second time utilizing the second substrate, followed by release of the second product, to complete one reaction cycle. The carbanions have been highlighted in **bold**. B₁ and B₂ represent basic groups in the active site, R₁ and R₂ represent the pyrimidine and pyrophosphate groups of ThDP, R and R’ represent the remaining chains of the ketose and aldose substrates.
Figure 1.5
1.1.4 BIOLOGICAL SIGNIFICANCE

Transketolase is found in all organisms examined to date, however the number of isoforms detected varies. For example, spinach apparently contains one transketolase as suggested by southern blot analysis (21), whereas three distinct transketolase genes were detected in the drought-tolerant plant *Craterostigma plantagineum* (22). In *C. plantagineum*, two of the isoforms were preferentially expressed during the rehydration phase of the plant, however these isoforms were not successfully purified for biochemical analysis (22). Studies on the regulation of the genes that encode two isoforms of yeast transketolase, Tkl1 and Tkl2, demonstrated that transcription is largely regulated by glucose catabolism (23-25). Examination of transketolase expression in *E. coli*, by comparing enzymatic activity of clarified extracts of single transketolase knockout strains, suggested *tktB* encoded for 30% of the total transketolase activity found in the wild-type organism (26).

Verification of the presence of different isoforms has been complicated by the apparent susceptibility of transketolase to posttranslational modification (PTM). The ratio of multiple variants, as observed by chromatographic separation of transketolase isoforms from yeast, rabbit, rat, and pig, was found to change based on the condition of the host (27). In these earlier studies however, no sequence analysis was conducted to differentiate genetic isoforms from transketolase that may have been subjected to PTM. Variants of human transketolase were first detected in patients suffering from the neurodegenerative disorder Wernicke-Korsakoff (WK) syndrome (28). WK is a combination of Wernicke’s disease, characterized by ataxy of gait, palsy, and mental confusion and Korsakoff’s psychosis, described as a large impairment of retentive
memory (29); both are caused by malnutrition, especially due to alcoholism. Transketolase from WK patients was shown to display a lower apparent $K_m$ for the cofactor ThDP (28, 30). Sequencing of cDNA encoding transketolase from WK patients compared to healthy individuals showed that genetic differences did not account for WK (31). Identifying modifications in WK forms of transketolase is therefore an important goal of current research on transketolase. At this time, there is no published information that demonstrates post translational modification of transketolase \textit{in vivo}. Establishing methods to detect and identify posttranslational modifications is crucial for identifying the nature of the transketolase abnormalities found in WK patients.

Transketolase was recognized as an important source of metabolic substrates for cell proliferation and this enzyme appears to possess a large flux control coefficient in erythrocytes (32) and liver cells (33). More recently, this transketolase-mediated flux control coefficient was recognized to be significant for the growth of tumor cells (34). In support of the role of transketolase in tumor proliferation, a recent study compared the two dimensional protein profiles of a poorly metastatic cell line and a highly metastatic cell line. The results identified transketolase as one of 9 proteins that accumulated to high levels in the highly metastatic cells (35). The importance of the nonoxidative pentose phosphate pathway, specifically transketolase, was further implicated by studies with oxythiamin (36), a noncompetitive inhibitor of transketolase. The addition of this compound caused a decrease in nucleic acid synthesis via decreased rates of ribose synthesis. Finally, it was shown that transketolase is a target for Avemar, the fermented wheat germ extract that exhibits anti-tumor properties by virtue of its inhibition of \textit{de}
*novo* nucleic acid biosynthesis (37). Precisely how Avemar inactivates transketolase is not known at present.

### 1.1.5 INDUSTRIAL APPLICATIONS

In addition to its importance in biological function, transketolase has received attention for its potential in enzymatic synthesis of asymmetric compounds. Considered valuable is the enzyme’s stereoselective specificity for (2R)-hydroxyaldehyde acceptor substrates resulting in specific generation of enantiomerically pure compounds that possess a (3S) configuration (38) ([Figure 1.6](#)). This selectivity is imparted by a conserved Asp residue in the active site of yeast transketolase, which was proposed to hydrogen bond to the C2 hydroxyl of the acceptor substrate. Mutation of this residue to Ala, disrupted the contact with acceptor substrate and led to a mutant enzyme with a slight reduction in catalytic efficiency and the loss of stereoselectivity (39). Also important is the ability of transketolase to accept a wide variety of substrates, though the donor substrate of choice is hydroxypyruvate, due to the release of CO₂ as the first product in the reaction to drive the equilibrium of the reaction. A large variety of α-hydroxy acceptor substrates have been shown to be utilized by transketolase (40). A number of compounds requiring an asymmetric carbon – carbon bond have been synthesized utilizing transketolase. Applications include the production of flavor and fragrance compounds (41), scientifically useful unnatural (42) as well as naturally occurring sugars to be utilized as precursors (43, 44), substrates (45), or analogues. The synthesis of specific isotopically labeled sugars to study metabolic processes (46). Methods applied to facilitate the use of transketolase include scale up procedures and improvement of
enzyme stability such as by immobilization to a support matrix (47). Studies furthering our understanding of transketolase stability, through the potential to guide modifications in attempt to improve enzyme function, would have significant applications in various fields of enzymatic synthesis.

**Figure 1.6.** The reaction scheme of transketolase. Scheme displays the two carbon unit that is transferred from the donor substrate (ketose) to the acceptor substrate (aldose). Also shown is the preference for the acceptor in the 2R configuration, resulting in an aldose product and a ketose product with a 3S configuration.
CHAPTER 2

CHARACTERIZATION OF TRANSKETOLASE ISOZYMES FROM RHODOBACTER Sphaeroides

2.1 INTRODUCTION

Transketolase has served as a model system to study the involvement of the cofactor ThDP in enzymatic catalysis. Much information has been gathered from studying various ThDP-dependent enzymes over the past decade, including elucidation of how this cofactor activates catalysis and other mechanistic considerations. However, much is unknown about enzymes that catalyze transketolase-like reactions and much more research is required. There is especially little known of the CBB pathway enzyme (encoded by the \textit{cbbT} gene) and how this transketolase differs from the oxidative pentose phosphate pathway enzyme (encoded by \textit{tkt}) of \textit{Rhodobacter sphaeroides}. The gene encoding CbbT had previously been identified within the form II CO$_2$ fixation (\textit{cbbII}) operon of \textit{Rhodobacter sphaeroides} (48). The \textit{R. sphaeroides cbbT} gene provided the first available deduced sequence of any transketolase protein. Its organization with other structural genes that encode enzymes of the CBB pathway allows for convenient and efficient regulation of gene expression (49). In contrast, the region containing \textit{tkt} remains largely unidentified, seemingly located independent of other known structural genes with the exception of the gene that encodes glyceraldehyde-3-phosphate dehydrogenase (\textit{gapdh}) (Figure 2.1). The DNA sequences in this region surrounding \textit{tkt} was searched to
confirm the annotation by the Joint Genome Institute (JGI), however no new matches were found in a recent Blast search of the surrounding annotated genes.

**Figure 2.1.** Arrangement of the cbb gene cluster containing *cbbT* (A) and the region containing *tkt* (B) in *R. sphaeroides*. Abbreviations are as follows: *cbbF* – fructose-1,6-bisphosphatase; *cbbP* – phosphoribulokinase; *cbbG* (*gapdh*) – glyceraldehyde-3-phosphate dehydrogenase; *cbbA* – fructose-bisphosphate aldolase; *cbbM* – ribulose-1,5-bisphosphate carboxylase/oxygenase.

Many organisms contain multiple isoforms of transketolase and much has been learned with regard to the differential expression of these isoforms in *Craterostigma plantagineum* (22) and to a lesser extent *Escherichia coli* (26) and *Saccharomyces cerevisiae* (23, 24). However, there is little information available of the biochemical properties of transketolase isoforms from the same organism. For these reasons, we
initiated studies of both the transketolases (CbbT and Tkt) from *R. sphaeroides*. The cloning, over-expression, purification of recombinant protein, and characterization of CbbT and Tkt are discussed in this chapter.

2.2 MATERIALS AND METHODS

2.2.1 GENERAL CHEMICALS

Ribose 5-P, xylulose 5-P, MgSO₄, ThDP, glyceraldehyde 3-phosphate dehydrogenase, and NAD⁺ were obtained from Sigma (St. Louis, MO). Oligonucleotides for mutagenesis were purchased from MWG Biotech. Sequencing was performed in house and by the Plant-Microbe Genomics Facility at The Ohio State University. Rabbit anti CbbT antibodies were obtained from Cocalico Biologicals, Inc (Reamstown, PA). N-terminal sequencing was performed by the Protein Structure Lab, University of California, Davis. Expression vector pET11a was from Novagen (Madison, WI). DEAE FF, QSHP, PSHP resins and the S200 column were from Pharmacia (currently Amersham Biosciences, San Francisco, CA).

2.2.2 CLONING OF THE TRANSKETOLASE GENES

The *cbbT* gene was PCR-amplified from plasmid pJG5, a vector containing a 4 kb segment of the *cbbII* operon from *R. sphaeroides*. This plasmid contains the complete *cbbT* gene (48). The primers designed for amplification allowed for the introduction of a 3′ *NdeI* and a 5′ *BglII* restriction site. The PCR product was cloned into pCRscript and then into pET11a to create the expression vector pET11CbbT. While in pET11a, *cbbT* was sequenced in both directions to confirm the wild type sequence.
To introduce the necessary NdeI and BglII restriction sites on either end of the cbbT gene, the following oligonucleotides obtained from Operon Technologies Inc. were used as primers: CbbTNdeI – 5’-GGGATTCCATATGAAGGACATTGGAGCCG-3’ and CbbTBglII – 5’-GAAGATCTCCCTCAGATCCGTTCTTTTGCC-3’. The cbbT containing plasmid pJG5 was used as the template. Each PCR reaction consisted of a 50 µl total volume and contained 10% DMSO, 1X Taq polymerase buffer (Invitrogen, Carlsbad, CA), 50 µM dNTPs, 1 mM MgCl2, 1 µM each of forward and reverse primers, 30 µg chromosomal DNA, and 2.5 U of Taq polymerase, with a titrating MgCl2 concentration from 1 to 5 mM in 1 mM increments. Thermocycling conditions were as follows, a 30 sec denaturing step at 95 °C, a 30 sec annealing step at 68 °C, and a 2 min extension step at 72 °C; this loop was repeated for a total of 25 cycles and was preceded by an initial 1 min denaturation at 95 °C and a final 4 min extension at 72 °C. Reactions containing at least 2 mM MgCl2 yielded a PCR product of the correct size (1.97 kbps). The PCR product was polished using T4 DNA polymerase to ensure blunt ends and increase ligation efficiency (50) then ligated into the pPCR-Script Amp vector (Stratagene, La Jolla, CA). This was transformed into E. coli strain JM109; resulting colonies displaying a white phenotype were selected and screened by digestion with BamHI, NdeI, and BglII. One of the plasmids that showed the correct digest pattern was selected for subsequent cloning into pET11a. The insert was prepared by sequential double digest with NdeI followed by BglII; the resulting 1.9 kb fragment was gel purified by agarose electrophoresis and isolated using the Gene Clean kit (QBiogene Inc.). This fragment was ligated into a pET11a vector that had been prepared by a sequential double digest using NdeI followed by BamHI. Ligation was performed overnight at 16 °C, in 20
µl total volume containing 1X ligation buffer, 0.04 µg vector, 0.2 µg insert, and 1 µl T4 ligase. The ligation was EtOH precipitated and the resulting DNA pellet resuspended by adding 3 µl of TEM. 1 µl of this was then transformed into JM109 to screen for correct inserts. Inserts were screened by digesting with EcoRI, SmaI, and BamHI. A colony that displayed the predicted fragment pattern was selected and a stock culture was prepared and stored at -80 °C. Confirmation of the cbbT DNA sequence was achieved using an ABI prism 310 Genetic Analyzer (Applied Biosystems, Foster City, CA). In addition to the standard T7 promoter and T7 terminator primers, several internal primers were designed to aid in the complete sequencing of both strands of cbbT. Template DNA was purified from a 2 ml overnight culture using the Qiagen mini prep kit. The sequencing reactions were performed as described in the manual of the ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit. Each sample consisted of 20 µl total volume and contained 0.5 – 1.0 µg of DNA (pET11CbbT), 0.5 µl of primer, 5% DMSO, and 1X of the provided ready mix. Standard PCR thermocycling conditions consisted of a 30 sec denaturation step at 95 °C, a 15 sec annealing step at 50 °C, and a 4 min extension step at 60 °C; this loop was repeated for a total of 25 cycles. The resulting PCR products were then EtOH precipitated, resuspended in the supplied sequencing buffer, and submitted for processing on the analyzer. The resulting sequences obtained were then assembled, manually aligned, and compared with the published CbbT sequence.

Tkt was PCR-amplified directly from genomic DNA. The PCR product was ligated into the pCR-TOPO vector. Single colonies were selected and tkt inserts generated by NdeI and BamHI sequential double digest. The inserts were then ligated into a prepared pET11a vector to generate the expression vector pET11Tkt. A single
colony showing the proper size insert was then selected for sequence analysis in both directions.

The following oligonucleotides were used to clone tkt and introduce the required NdeI and BglII restriction sites:

\[
\text{TKTNdeI} - 5'\text{-GGAATTCCATATGGACATACAGAGCCTGCGC-3',} \\
\text{TKTBglII} - 5'\text{-GAAGATCTTCAGCCCAGAAGGGCTTTGGCC-3'.}
\]

A blend of Pfu and Taq DNA polymerases in a 1:50 ratio was used to minimize replication errors. Each PCR reaction consisted of a 50 µl total volume and contained 10% DMSO, 1X Taq polymerase buffer (Invitrogen), 50 µM dNTPs, 1 mM MgSO\textsubscript{4}, 1 µM each of forward and reverse primers, 30 µg chromosomal DNA, 0.05 U of Pfu polymerase, and 2.5 U of Taq polymerase, with a MgCl\textsubscript{2} concentration titration from 0 to 5 mM in 1 mM increments. Thermocycling conditions were as follows, a 30 sec denaturing step at 95 °C, a 30 sec annealing step at 65 °C, and a 2 min extension step at 72 °C; this loop was repeated for a total of 25 cycles and was preceded by an initial 1 min denaturation at 95 °C and a final 30 min extension at 72 °C. Reactions containing 2, 3, and 4 mM MgCl\textsubscript{2} all yielded a PCR product of the correct size (2.02 kb). The PCR product from the 2 mM MgCl\textsubscript{2} reaction was ligated into the pCR2.1-TOPO cloning vector (Invitrogen) and transformed into JM109 by electroporation. Two resulting colonies were selected, grown in 2 ml overnight cultures, plasmid prepared by an alkaline lysis mini preparation procedure, and analyzed by a sequential double digest with NdeI for 24 h at 37 °C, followed by BglII for 20 h at 37 °C. The resulting 2 kb fragments were isolated by agarose gel electrophoresis and purified using the Qia gel purification kit (Qiagen, Valencia CA). The insert was ligated into gel purified pET11a vector prepared by an NdeI BglII
sequential double digest. This plasmid was transformed into *E. coli* JM109 and white colonies were selected and then screened for insert by restriction digest. One colony was selected for sequencing. A 10 ml overnight culture was mini prepped using the Qiagen kit and the resulting DNA submitted for sequencing at the Plant-Microbe Genomics Facility (PMGF) at The Ohio State University. Primers used for sequencing included T7 promoter and terminator primers as well as three 20-mer primers designed to anneal throughout the *tkt* gene.

### 2.2.3 OVER-EXPRESSION OF TKT AND CBBT

The vector pET11CbbT was used to over-express the *cbbT* gene and the vector pETTkt the *tkt* gene, in the salt-inducible *E. coli* strain GJ1158 (51). A 3.6 l culture of LB-NaCl medium (LB with no salt) in a 4 l Erlenmeyer flask was inoculated with 3 ml of an overnight culture. This was grown at 37 °C and shaken at 200 rpm until the culture reached an OD$_{600}$ of 0.8, at which time 5 M NaCl was added to bring the final NaCl concentration to 0.4 M. The temperature was reduced to 25 °C and the shaker rotational speed as also reduced to 120 rpm. Induction under these conditions was allowed to proceed overnight (12 – 16 h). The cells were harvested and washed two times with TEM (20 mM Tris-HCl, pH 7.0, 0.1% BME, 1 mM EDTA) buffer; the pellet was stored at -80 °C until further use.

### 2.2.4 GROWTH OF RHODOBACTER SPHAEROIDES

*R. sphaeroides* strains CAC (52) and 1884 (53), were grown under autotrophic and photoautotrophic conditions under a gas atmosphere of 1.5% CO$_2$/98.5% H$_2$. 

22
Aerobic chemoheterotrophic growth was performed in Omerod's medium (54) supplemented with 0.4% malate. For strain CAC, chemoautotrophic conditions were achieved by bubbling with a gas atmosphere of 5% CO₂, 45% H₂, 50% air and were grown in Omerod’s media in the dark. Anaerobic conditions for phototrophic growth were maintained by bubbling with argon and light was provided by four incandescent 60-W light bulbs. All media was supplemented with 15 µg biotin, 1 mg nicotinic acid and 1 mg thiamine-HCl per l. All cultures were grown in 500 ml rectangular bottles containing 300 ml of media, in 30 °C temperature controlled water baths. Inoculation was with malate cultures grown to the late logarithmic phase. Chemoheterotrophic cultures were harvested at an OD 600 nm of 0.4, all other cultures at an OD 600 nm of 0.8; resulting cell pellets were washed once with the suspension buffer (20 mM MOPS-NaOH pH 6.9, 0.1 mM ThDP, 2 mM MgCl₂, and 0.1% BME) and stored at -20 °C overnight. The cells were thawed on ice, resuspended in ice cold suspension buffer, and then lysed by sonication. These extracts were clarified by a 10 min centrifugation at 15,000 g at 4 °C and then filtered through a 0.4 µ filter.

2.2.5 PURIFICATION OF CBBT

All of the purification steps were conducted either at 4 °C or on ice. Cells were resuspended in 20 ml buffer A (20 mM MOPS-NaOH, pH 6.9, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME) and were lysed using a French press. This was followed by a low speed centrifugation (10,000g, 20 min). Nucleic acids were precipitated from the supernatant of the previous centrifugation by incubating extracts with 0.1% streptomycin sulfate on ice for 20 minutes. After ultracentrifugation (100,000 g, 1 h), followed by
filtration through a 0.45 µ filter, the extract was loaded onto a 2.6 x 14 cm DEAE fast-flow sepharose (DSFF) column and eluted using an increasing linear gradient of 2 – 15% buffer B (buffer A with 1 M ammonium sulfate). Active fractions eluted between 85 and 130 mM ammonium sulfate; these were pooled and then treated with solid ammonium sulfate to a final concentration of 1 M. Following a 1 h equilibration on ice, the solution was filtered and then loaded onto a 1.6 x 90 cm phenyl sepharose high performance (PSHP) column; elution was achieved using a decreasing linear gradient from 100% to 30% of buffer B. Active fractions eluting between 890 and 600 mM ammonium sulfate were pooled and diluted with buffer A to bring the final salt concentration to ~ 50 mM. This material was loaded onto a 1.6 x 11 cm Q-sepharose HP (QSHP) column and active enzyme eluted using a linear gradient ranging from 0 – 20% buffer B. Active fractions, eluting between 115 and 210 mM ammonium sulfate, were concentrated (Millipore spin concentrators, NMW 30 kDa) then loaded onto a 1.6 x 50 cm Superose 200 gel filtration column and eluted with 5% buffer B; CbbT peak elution was at 61.8 ml. For storage of purified CbbT, glycerol was added to 20%, the aliquots were flash frozen in liquid N2 and stored at -80 °C. For analytical gel filtration analysis, a 1.0 x 30 cm Superose-12 column calibrated with 5% buffer B was used. Under these conditions CbbT eluted at 12.2 ml corresponding to a molecular weight of 90,000; this value is not close to the theoretically predicted homodimer molecular weight of a of 138,000. However, when CbbT was eluted in buffer A containing 500 mM NaCl, CbbT eluted at 11.8 ml, corresponding to a molecular weight of 122,000. Following preparative gel filtration, glycerol was added to 20% and aliquots were flash frozen in liquid N2 and stored at -80 °C until needed; no noticeable loss in activity was observed after storage.
2.2.6 **PURIFICATION OF TKT**

All of the purification steps were conducted either at 4 °C or on ice. Cells were resuspended in 20 ml buffer A (20 mM MOPS-NaOH, pH 6.9, 0.1 mM ThDP, 2 mM MgSO₄, and 10 mM BME) and were lysed using a French press. This was followed by a low speed centrifugation (10,000g, 20 min). Nucleic acids were precipitated from the supernatant of the previous centrifugation by incubating extracts with 0.1% streptomycin sulfate on ice. After ultracentrifugation (100,000 g, 1 h), followed by filtration through a 0.45 micron filter, the extract was loaded onto a 2.6 x 14 cm DSFF column with most of the transketolase activity passing through the column. The flow through containing the majority of the total transketolase activity was then loaded onto 1.6 x 11 cm Q-sepharose HP column equilibrated with buffer A; all activity passed through this column as well. To this flow-through fraction, finely ground ammonium sulfate was slowly added to 2.5 M final concentration, while on ice with constant stirring. After 1 h of equilibration the sample was filtered using a 0.45 µ syringe filter. This was loaded onto a 1.6 x 90 cm phenyl sepharose HP column equilibrated with 100% buffer B (buffer A with 2 M ammonium sulfate), and eluted with an isocratic flow of 100% B. The 60 ml of eluent containing the activity was then concentrated to 1.2 ml for gel filtration (Millipore spin concentrator, 15 ml capacity, NMW 30 kDa) and loaded onto a 1.6 x 50 cm Superpose 200 gel filtration column and eluted with 5% buffer B. Peak elution occurred at 62.3 ml. The active peak was collected and diluted 1:4 with buffer C (20 mM MOPS-NaOH, pH 6.5, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME) was then loaded over the course of two runs onto a 1.0 x 9.4 cm column containing the positive ion exchange resin S-sepharose HP (SSHP) and eluted by a linear gradient of 0 – 18% buffer D (buffer C with
1 M ammonium sulfate). Active fractions eluted in a sharp peak from 80 to 110 mM ammonium sulfate. For storage of purified Tkt, glycerol was added to 20%, the aliquots were flash frozen in liquid N₂ and stored in a -80 °C freezer. For analytical gel filtration analysis, a 1.0 x 30 cm Superose-12 column calibrated with 5% buffer B was used. Under these conditions Tkt eluted at 12.1 ml corresponding to a molecular weight of 140,000.

2.2.7 SDS AND NATIVE PAGE OF TKT AND CBBT

All polyacrylamide gels were made from a 30% acrylamide stock containing 0.8% bis acrylamide. For SDS-PAGE, 10 to 12% acrylamide was typically used in the running gel in conjunction with a 4% stacking gel. The buffer used in the running gel contained 375 mM Tris-HCl pH 8.8, 0.1% SDS, 0.05% APS, and polymerization was initiated by the addition of TEMED to 0.05%. The stacking gel buffer was identical with the exception that 125 mM Tris-HCl of pH 6.8 was used. For native PAGE, stacking gels were also composed of 4% acrylamide; however the running gels contained 7% acrylamide. Gels were poured and run using the Mini-PROTEAN II System (BioRad). The running buffer consisted of 25 mM Tris, 200 mM glycine, 0.08% SDS with a pH of 8.8. Gels were run at constant current of 20 mA per gel at room temperature unless otherwise noted. Bands were visualized either by Coomassie R-250 or by rapid silver staining (55). Sample buffer consisted of 62.5 mM Tris-HCl pH 6.8, 2% SDS, and 2% BME. For carboxyalkylation of Cys, samples (~ 1 mg/ml protein, pH 7.8 20 mM HEPES-NaOH, 1% SDS) were first reduced by incubation with 2 mM DTT for 30
minutes at 37 °C and then alkylated with 5 mM iodoacetamide for 30 min at room temperature in the dark.

2.2.8 TRANSKETOLASE ACTIVITY ASSAY

The enzymatic assay was based on the NAD\(^+\) dependent glyceraldehyde-3-phosphate dehydrogenase coupled system developed by Kochetov (56). Each cuvette contained 100 mM buffer, 5 mM MgSO\(_4\), 1 mM ThDP, 2.5 mM NaAsO\(_4\), 0.4 mM NAD\(^+\), 4 mM R5P, 0.1% BME, 2 units of GAPDH, and transketolase (around 15 mU). The total volume was 1 ml, less the volume of the Xu5P to be added. The cuvette was mixed and equilibrated at 25 °C. The progress of the reaction was monitored at 340 nm. After approximately one min, upon establishing a stable linear baseline, the transketolase reaction was initiated by the addition of Xu5P (typically to a final concentration of 0.2 mM) the solution was quickly mixed and then data collection was resumed. Enzyme activity was calculated based on the molar absorbtivity of NAD\(^+\) with a value of 6.23 x 10\(^3\) M\(^{-1}\) cm\(^{-1}\). One unit of enzyme activity was defined as the reduction of 1 µmol NAD min\(^{-1}\). Protein concentration was determined using the Bio-Rad protein microassay procedure. All kinetic data were collected at 25 °C using a Varian Cary 100 spectrophotometer with temperature controller.

2.2.9 DETERMINING OPTIMAL PH

Dependence of pH on activity was measured using 3 different buffers at approximate 0.2 pH unit intervals with overlapping regions between different buffers. For this method the buffers used were MOPS-NaOH, HEPES-NaOH, and glycylglycine.
Tris-HCl was also used to measure the pH optima of Tkt, however not for CbbT due to an observed inhibition by Tris. The exact pH of the reaction was measured immediately following the assay using a calibrated pH probe; all buffers were used at a final concentration of 200 mM in the assay.

2.2.10 KINETIC ANALYSIS OF CBBT AND TKT

Substrate kinetics were determined in the presence of 0.1 mM ThDP and 5 mM MgSO₄ at 25°C. For CbbT the buffer HEPES was used at a concentration of 100 mM and for Tkt Tris-HCl was used at the same concentration and pH of 7.8. Additional measurements were made for CbbT using MOPS-NaOH (pH 6.9) and glycylglycine (pH 8.5). Xylulose 5-P (Xu5P) was used as the donor substrate and ribose 5-P (R5P) as the acceptor substrate. For determination of $K_m$ and $V_{max}$ values, concentrations of one of the substrates was varied over a range from 0 to 0.5 mM for Xu5P and 0 to 20 mM for R5P while the corresponding substrate was maintained in excess at a constant non-inhibitory concentration in the 1.0 ml assay mixture. The concentrations of the substrates were determined enzymatically by running the activity assay in a quantitative manner. SigmaPlot v8.0 was used for data analysis.

2.2.11 MOLECULAR MODELING OF CBBT AND TKT

Molecular models were built for visualization of the tertiary structure of both proteins. Theoretical models of CbbT and Tkt based on the crystal structures of yeast transketolase (pdb IDs: 1AYO, 1GPU, 1ITZ, 1NGS, 1TKA, 1TKB, 1TKC, and 1TRK) (13, 17-19, 57-59), E. coli transketolase (pdb 1QDG) (9), and maize transketolase (pdb
1ITZ) (8), were constructed using the program Swiss-Model. Alignments and mutations were analyzed using the programs SPDBV (60-63) and O (64).

2.2.12 SYNTHESIS OF XU5P

Synthesis of xylulose 5P was based on the method published by Zimmermann et al (45) and modified as discussed in section 2.3.8. Xu5P was purified from the reaction mixture using Dowex 50W-X cation (H\(^+\) form) exchange resin in a batch method. This was followed by a 2.2 X 14 cm Dowex 1-X8 anion exchange resin (formate form) column, eluting Xu5P with an increasing gradient of formic acid (from 0 to 4 M), peak elution occurred at 3 M formic acid. Fractions showing substantial Xu5P (determined enzymatically) were pooled and concentrated under vacuum using a rotovap. \(^{1}H\) NMR was performed by the Campus Chemical Instrument Center (CCIC) at The Ohio State University.

2.3 RESULTS AND DISCUSSION

2.3.1 CLONING OF CBBT AND TKT INTO PET11A

The pET expression system of Novagen had previously been used successfully in our lab and has been adopted as a standard method for the over-expression of bacterial genes in appropriate \textit{E. coli} strains. Target genes are cloned into the pET vector under the control of strong bacteriophage T7 promoter, and expression is induced by providing a source of T7 RNA polymerase in the host cell (65). The cloning step was performed with the use of the non-expression host cell, \textit{E. coli} strain JM109. Cloning into pET11a requires usage of the restriction nuclease \textit{NdeI} that contains the start codon
sequence ATG within its recognition sequence CATATG. The pET11 vector also contains the restriction nuclease BamHI for the insertion of the 3’ end of the gene, prior to a T7 terminator in the vector. However cbbT contains a BamHI site, thereby requiring either silent mutagenesis to remove this internal restriction site or the use of an alternate restriction nuclease in the cloning of cbbT into pET11a. We chose to introduce a BglII restriction site on the end of cbbT, as the BglII restriction yields a sticky end complimentary to a BamHI restriction fragment, with the resulting ligation product no longer recognized by either restriction enzyme.

The cbbT gene was PCR-amplified from plasmid pJG5, a pUC based vector containing a 4 kb segment of the cbbII operon from R. sphaeroides that contains the complete cbbT gene (48). The primers designed for amplification allowed for the introduction of a 3’ NdeI and a 5’ BglII restriction site. Initially a successful PCR reaction was elusive despite the application of various methods to improve the likelihood of a product. After standard PCR reaction conditions failed to yield product, a touchdown PCR method was attempted. The touchdown method of PCR begins with a higher annealing temperature than necessary for the primers in use and slowly decreases this annealing temperature with each subsequent cycle until the desired temperature has been reached (66). In theory this approach is useful if the primers are having a difficult time annealing to their target sequences. Less specific products are generated in the initial reactions and then more specific primer/template pairings are made as the annealing temperature decreases. In addition, a core sampling approach was applied in which the template is obtained by running a PCR product on agarose electrophoresis, and then isolated by coring the desired band or region based on the markers directly out of the
gel. This small core with a volume of approximately 10 µl is then added directly to the PCR reaction. This technique has been applied in cases where more than one PCR product is generated and allows for a rapid selection of the appropriate target for a subsequent round of PCR. However, these more exotic methods also did not provide product. It was not until DMSO was added to the reaction that an actual PCR product could be visualized on an agarose gel by ethidium bromide staining. DMSO is thought to reduce secondary structure and has been demonstrated to be useful for GC rich templates (67, 68), such as is found in *R. sphaeroides*. The PCR product of the expected size was first cloned into pCRscript and then into pET11a to create the expression vector pET11CbbT. This was done to utilize the high copy number of pCRscript in generating an ample supply of *cbbT* insert for subcloning into pET11a. While in pET11a, *cbbT* was sequenced in both directions with the aid of 7 primers designed to anneal throughout the gene. Interestingly, the sequence obtained from the pET11CbbT revealed four nucleic acids that deviated from the published sequence. All four of the changes were confirmed to be present on both forward and reverse strands of *cbbT*, and were confirmed by sequencing the template (pJG5) used for PCR. These four differences in nucleic acid sequence translated to three amino acid differences (*Table 2.1*), all of which matched the published amino acid sequence of the homologous yeast transketolase enzyme Tkl1. For these reasons the *cbbT* sequence obtained was considered to be that of authentic wild type *cbbT*. These three amino acids were also found to agree with the more recently available conceptual translation made available by the NCBI Microbial Genomes Annotation Project in September of 2003.
Table 2.1. Location of the nucleotides corrected by sequencing. The four nucleotides were distributed among three codons, all of which led to an alternate amino acid at these positions. Nucleotide numbering begins with the adenine of the start codon ATG.

<table>
<thead>
<tr>
<th>Published Nucleotides</th>
<th>Codon (Amino Acid) Published</th>
<th>Sequenced Nucleotides</th>
<th>Codon (Amino Acid) Corrected</th>
</tr>
</thead>
<tbody>
<tr>
<td>G 518, C 519</td>
<td>GGC (Gly 173)</td>
<td>C 518, G 519</td>
<td>GCG (Ala 173)</td>
</tr>
<tr>
<td>T 586</td>
<td>TCG (Ser 196)</td>
<td>A 586</td>
<td>ACG (Thr 196)</td>
</tr>
<tr>
<td>C 620</td>
<td>CCG (Pro 207)</td>
<td>G 620</td>
<td>CGG (Arg 207)</td>
</tr>
</tbody>
</table>

A similar methodology was applied to cloning the tkt gene into the pET11a expression vector. Due to the lack of any preexisting subclones, tkt was PCR-amplified directly from genomic DNA, using reaction conditions similar to those established for cbbT. Primers were designed based on the publicly available R. sphaeroides genomic sequence to prepare a PCR fragment containing tkt from chromosomal DNA. Similar to cbbT, tkt contained an internal BamHI restriction site; therefore a BglII site was introduced at the 5’ end to facilitate cloning into pET11a. The primers designed for tkt introduced the desired NdeI and BglII restriction sites. The PCR product was first ligated into the pCR-TOPO 2.1 vector. Single colonies were selected and tkt inserts generated by NdeI and BglII sequential double digest. The inserts were then ligated into prepared pET11a vector to generate the vector pET11TKT. A single colony showing insert was
then selected for sequence analysis in both directions. The resulting sequence was an exact match with the sequence publicly available by the NCBI for tkt.

### 2.3.2 OVER-EXPRESSION OF CBBT AND TKT GENES IN E. COLI

The over-expression studies with pET11CbbT was initially performed using *E. coli* strain BL21, inducing the *lac* promoter with IPTG. This method produced moderate levels of CbbT. As a comparison, over-expression using *E. coli* strain GJ1158 was also tested. GJ1158 carries a single integrated copy of the phage T7 polymerase under the control of the *proU* promoter. The *proU* operon encodes 3 genes that encode a high-affinity glycine betaine transport system. Transcription of *proU* in *E. coli* is activated several-hundredfold in cultures grown in high-osmolarity media (69). Induction of the T7 polymerase is thereby performed by increasing the salt concentration of the media. In addition to cost effectiveness, another reported advantage of this method is an increase in soluble over-expressed protein. Various concentrations of NaCl and various times of induction were explored. It was found that inducing with 0.4 M NaCl at 25 °C for 12 to 16 h yielded reproducible high levels of CbbT accumulation. In addition, it was found that cultures grown with low levels of oxygen contained higher overall transketolase activity. These conditions were also applied to the production of recombinant Tkt, resulting in the induction of large amounts of soluble protein. Typical specific activities obtained from crude extracts of cells grown under these conditions ranged from 3.5 to 5.5 U/mg.
2.3.3 PURIFICATION OF CBBT

From the deduced amino acid sequence, CbbT has a theoretical monomeric molecular weight of 69.3 kDa and a calculated pI of 5.65. Upon establishing a robust over-expression system, various column matrices were tested for interaction with CbbT from resulting crude extracts. Tested matrices included Cibracon Blue, DEAE FF, QSHP, Green A, PSHP, Butyl S, SSHP, Hydroxylapatite and a R5P linked agarose, in addition to several gel filtration columns. The search for effective purification matrices was complicated by the fact that CbbT became degraded after samples were heated for the usual SDS-PAGE procedure. Thus, SDS-PAGE gels falsely reflected the purification state of CbbT preparations. This apparent lack of purification led us to try some less common matrices such as dye-affinity columns as well as construct our own R5P-linked Sepharose affinity matrix. No significant interactions of CbbT were observed with the dye affinity columns and most proteins of the crude extract passed right through the columns. In contrast, CbbT did bind to the R5P agarose column; however this required removal of cofactors required to maintain activity, namely Mg$^{++}$ and ThDP, as well the absence of Tris buffers. Hydroxyapatite columns offered significant purification, however this column was eliminated due to its requirement for phosphate. Phosphate was found to inhibit CbbT, similar to other transketolases. In addition, it was communicated to us that *E. coli* transketolase exposed to phosphate during purification remained as a ligand in the resulting crystal structure. Not heating samples prior to SDS-PAGE lead to the realization that several of the column matrices tested had resulted in significant purification. The matrices subsequently chosen for optimization were DEAE FF, QSHP, and PSHP.
Upon final optimization of the DEAE FF column, it was found that only an additional PSHP column was necessary for CbbT of high purity as seen by SDS-PAGE. However, for most preparations the stronger anion exchanger QSHP and the gel filtration column Superdex-200 were also used to maximize reproducibility. SDS-PAGE gels demonstrated the purification through each step (Figure 2.2) and the salient features of the purification protocol adapted is presented in Table 2.2. Upon establishing the protective effect of ammonium sulfate, this salt was used throughout the purification. Based on the theoretical subunit size of 69.3 kDa, CbbT migration on SDS/PAGE gels (MW = 66 kDa) and native gel filtration chromatography (MW = 122 kDa) indicated that the CbbT protein was most likely a homodimer under non-denaturing conditions.
Figure 2.2. Purification of CbbT. 10% SDS-PAGE Coomassie-stained gel showing the stages of CbbT purification. Lanes: 1, crude extract following the high speed centrifugation; 2, DSFF anion exchange column; 3, PSHP hydrophobic interaction column; 4, QSHP anion exchange column. Approximately 8 µg of total protein was loaded on each lane. No visual change in purity was observed compared to lane 4, following a subsequent gel filtration column.
<table>
<thead>
<tr>
<th>Purification Step</th>
<th>Volume (ml)</th>
<th>Activity (U/ml)</th>
<th>Protein (mg/ml)</th>
<th>Activity (U/mg)</th>
<th>Total Units</th>
<th>Total Protein (mg)</th>
<th>% Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude</td>
<td>50.0</td>
<td>55.76</td>
<td>9.13</td>
<td>6.11</td>
<td>2790</td>
<td>457</td>
<td>100.0</td>
</tr>
<tr>
<td>DSFF</td>
<td>104.0</td>
<td>16.45</td>
<td>1.47</td>
<td>11.2</td>
<td>1710</td>
<td>153</td>
<td>61.4</td>
</tr>
<tr>
<td>PSHP</td>
<td>66.0</td>
<td>28.04</td>
<td>1.67</td>
<td>16.8</td>
<td>1850</td>
<td>110</td>
<td>66.4</td>
</tr>
<tr>
<td>QSHP</td>
<td>40.0</td>
<td>36.96</td>
<td>2.16</td>
<td>17.1</td>
<td>1480</td>
<td>86.3</td>
<td>53.0</td>
</tr>
<tr>
<td>S200</td>
<td>14.0</td>
<td>93.91</td>
<td>5.12</td>
<td>29.8</td>
<td>1410</td>
<td>71.7</td>
<td>50.5</td>
</tr>
</tbody>
</table>

**Table 2.2.** CbbT purification table. The results of a purification beginning with a 7.2 l harvest of *E. coli* GJ1158 containing plasmid pET11CbbT grown under microaerobic conditions. Activity measurements were performed in 100 mM HEPES-NaOH, pH 7.8, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME at 25 °C. One unit of activity (U) is defined as µmol substrate consumed min⁻¹; % yield was calculated based on total units of activity.

### 2.3.4 PURIFICATION OF TKT

In contrast to CbbT, Tkt exhibited no interaction with anion exchange resins at pH 7.5, including the strong quaternary amine anion exchangers QSHP and UnoQ. This was explained by the relatively high pI of 6.74 calculated for Tkt based on the deduced amino acid sequence. By comparison, the pI for CbbT is 5.65. In addition to differences in charge, Tkt displayed a much weaker affinity for the hydrophobic interaction resin PSHP,
requiring 2.5 M ammonium sulfate to bind. Though Tkt released from the column during
the 2 M ammonium sulfate isocratic wash, this column still offered an appreciable
purification step. Similar to the CbbT purification, a final gel filtration step with
Superdex-200 was added as a polishing step for the purification. Peak elution of Tkt
occurred at 62.3 ml using the same exact column used for CbbT. However, since the
anion exchange columns were relatively ineffective, active gel filtration fractions still
contained several contaminating protein bands (Figure 2.3 A). The higher pI of Tkt
suggested that the cation exchange resin sulfopropyl linked sepharose (SSHP) at pH 6.5
might be useful. Indeed, this proved to be an ideal column for Tkt purification, with
virtually all the contaminating protein bands removed in the flow through and wash, with
active Tkt elution occurring over a narrow salt range (Figure 2.3 B). This was the last
column employed in this extensive purification, which still led to a considerable yield of
72.6 mg of purified fully active protein (Table 2.3). Based on the theoretical subunit size
of 72.4 kDa, Tkt migration on SDS/PAGE gels (MW = 70 kDa) and native gel filtration
chromatography (MW = 140 kDa) the Tkt protein was most likely a homodimer under
nondenaturing conditions. Based on these studies, future purifications of this enzyme
could be optimized further to obviate the use of anion exchange columns, capitalizing on
the demonstrated effectiveness of the PSHP and SSHP columns. Nevertheless, the
purification protocol adopted for Tkt yielded high specific activity preparations which
could be effectively used to compare with the properties of CbbT.
**Figure 2.3.** SDS-PAGE of samples from each stage of the Tkt purification. (A) Gel demonstrating the first four columns utilized. Lane 1: crude extract; Lane 2: DSFF flow through; Lane 3: QSHP flow-through; Lane 4: PSHP high salt wash; Lane 5: S200 peak fractions; Lane MW: molecular weight markers, sizes are indicated to the right of the gel in kDa. (B) Gel demonstrating the effectiveness of the cation exchange SSHP column. Lane L; load sample, preparation initially placed onto the SSHP column; it is the same as the S200 sample from Figure 2.3 A, but with more sample loaded; Lane F; represent proteins eluted in the flow-through prior to the salt wash. The elution of Tkt over six dominant fractions can be seen in the subsequent lanes; the ammonium sulfate gradient was calculated based on the measured conductivity of the fractions and is indicated above these lanes.
Figure 2.3
Table 2.3. Tkt purification table. The results of a purification beginning with a 7.2 l
harvest of *E. coli* GJ1158 containing plasmid pET11Tkt grown under microaerobic
conditions. Activity measurements were performed in 100 mM MOPS-NaOH pH 6.8,
0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME at 25 °C. One unit of activity (U) is
defined as µmol substrate consumed min⁻¹; % yield was calculated based on total units of
activity.

<table>
<thead>
<tr>
<th>Purification Step</th>
<th>Volume (ml)</th>
<th>Activity (U/ml)</th>
<th>Protein (mg/ml)</th>
<th>Activity (U/mg)</th>
<th>Total Units</th>
<th>Total Protein (mg)</th>
<th>% Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crude</td>
<td>30.0</td>
<td>16.59</td>
<td>55.26</td>
<td>3.33</td>
<td>1660</td>
<td>498</td>
<td>100</td>
</tr>
<tr>
<td>DSFF</td>
<td>200</td>
<td>1.43</td>
<td>7.25</td>
<td>5.07</td>
<td>1450</td>
<td>286</td>
<td>87</td>
</tr>
<tr>
<td>QSHP</td>
<td>250</td>
<td>0.97</td>
<td>4.91</td>
<td>5.08</td>
<td>1230</td>
<td>242</td>
<td>74</td>
</tr>
<tr>
<td>PSHP</td>
<td>60.0</td>
<td>2.32</td>
<td>17.84</td>
<td>7.70</td>
<td>1070</td>
<td>139</td>
<td>65</td>
</tr>
<tr>
<td>S200</td>
<td>30.0</td>
<td>3.22</td>
<td>19.05</td>
<td>5.91</td>
<td>571</td>
<td>96.6</td>
<td>34</td>
</tr>
<tr>
<td>SSHP</td>
<td>30.0</td>
<td>2.42</td>
<td>26.15</td>
<td>10.80</td>
<td>784</td>
<td>72.6</td>
<td>47</td>
</tr>
</tbody>
</table>

**2.3.5 ENZYME STABILITY**

Enzyme stability also played a large role in designing the purification schemes. As mentioned previously, the discovery that ammonium sulfate protected activity allowed for purification with practically no loss in total units of activity. This effect of protection was explored further in an attempt to quantify the amount required for
protect. Using CbbT that was purified in the absence of ammonium sulfate (utilizing NaCl instead in the purification), various concentrations of ammonium sulfate was added from 0 to 500 mM. These samples were incubated on ice and activity was assayed over time. Concentrations as low as 5 mM ammonium sulfate were able to provide dramatic stabilization of enzymatic activity (Figure 2.4 A). In addition, samples containing higher concentrations of ammonium sulfate showed an increase in activity over time, suggesting that the enzyme purified in the absence of ammonium sulfate was not fully reconstituted. Ammonium sulfate had never previously been noted to play such an important role relative to transketolase stability; in fact it is a reported inhibitor for the yeast enzyme (70). To determine which ion is responsible for stability, samples were also incubated in NH₄Cl and Na₂SO₄, where it was found that SO₄²⁻ is responsible for the increase in stability (Figure 2.4 B). For this reason MgSO₄ was subsequently used to maintain the concentration of the divalent metal cofactor instead of MgCl₂ in all buffers.
Figure 2.4. Sulfate effect on CbbT activity versus time. (A) During incubation at 4 °C under variable concentrations of ammonium sulfate (AS). Concentrations of ammonium sulfate are indicated in the legend. (B) Incubation at 25 °C in the presence of 0.5 M ammonium sulfate, 1 M NH₄Cl, or 0.5 M Na₂SO₄. All activity measurements were performed in 100 mM MOPS-NaOH pH 6.8, 0.1 mM ThDP, 2 mM MgCl₂, and 0.1% BME at 25 °C. Salt was also added to the assay so that every assay was identical in chemical environment. The unit of specific activity (U/ml) is defined as µmol substrate consumed min⁻¹ ml⁻¹ of CbbT. The preparation of CbbT used in both experiments was performed using NaCl instead of ammonium sulfate for all but the PSHP column.
Figure 2.4
2.3.6 OXYGEN-MEDIATED INACTIVATION

It was reported that *E. coli* transketolase that had been covalently linked to a support matrix was susceptible to inactivation by air (47). As part of our goal to maximize transketolase stability, the activity of air sparged samples of purified *R. sphaeroides* CbbT in solution was compared with argon sparged samples after incubation at various times at 25 °C. It was demonstrated that CbbT undergoes an apparent oxygen-dependent inactivation (Figure 2.5 A). Inactivation was slowed by increasing the concentration of reductant in the buffer (either BME or DTT); however inactivation appeared to be irreversible. If near full inactivation was reached, CbbT began to precipitate out of solution. Although most of the oxygen-sensitivity studies focused on CbbT, Tkt was similarly inactivated (Figure 2.5 B). The apparent lag in inactivation observed for the air samples in Figure 2.5 occurs in the presence of reductant, with inactivation only beginning after consumption of the excess reductant by oxygen in the air sample. This lag is not observed in samples devoid of an exogenous source of reductant, examples of this can be seen in Figures 3.3 and 3.4.
**Figure 2.5.** Oxygen inactivation of transketolase from *R. sphaeroides*. (A) CbbT inactivation at 25 °C versus time. (B) Tkt inactivation at 25 °C versus time. Each enzyme was desalted into 20 mM HEPES-NaOH pH 7.8, 2 mM MgSO4, 0.1% BME, and 0.1 mM ThDP. Desalted samples (0.3 mg/ml) were sparged on ice with either N2 or air for 20 min and then placed in a 25 °C water bath at time point 0. Activity of both enzymes was assayed in 100 mM HEPES-NaOH pH 7.8, 2 mM MgSO4, 0.1% BME, and 0.1 mM ThDP at 25 °C. The specific activity of CbbT and Tkt was 11 and 33 µmol min\(^{-1}\) mg\(^{-1}\) respectively, as determined at pH 7.8.
Figure 2.5
2.3.7 THERMAL DEGRADATION OF CBBT AND TKT

Although our stability studies focused on enzymatic activity, we also explored the thermal degradation observed for CbbT after samples were heated in preparation for SDS-PAGE. This phenomenon has been reported to also occur with *E. coli* transketolase (71), however no further details were provided. To help rule out the presence of trace proteases, various protease inhibitors (bacterial protease inhibitor cocktail, Sigma P 8465) were added and shown to have no effect. It also appeared that the fragmentation was relatively insensitive to pH, occurring on a similar level within pH’s ranging from 4 to 8. The type of buffer used was also investigated but a similar level of fragmentation was observed whether MOPS, HEPES, KPO4, Tris-Cl, and Tris-SO4 buffers were used.

Interestingly, CbbT degradation was greatly reduced by first carboxamidomethylating the free thiols of CbbT with iodoacetamide, suggesting that a reactive cysteine was involved in this reaction. CbbT contains 6 Cys residues per subunit, and part of the work presented in Chapter 4 of this dissertation considers the role of Cys-160 and various site-directed mutants on catalysis and cofactor interaction. However, boiling of a C160A mutant enzyme prior to SDS/PAGE still exhibited the observed degradation with the wild-type protein, which may be prevented by alkylation with iodoacetamide (Figure 2.6). These results indicated that one or more of the remaining 5 Cys residues are involved in the CbbT degradation. The fragmentation was found to be dependent on temperature, occurring between 70 °C and 100 °C, following 20 min of incubation in the presence of SDS (Figure 2.7 A). Surprisingly, it was found that this process could also be prevented by the addition of a high concentration of acceptor substrate R5P (Figure 2.7 B) or the donor substrate F6P (data not shown). The fact that these substrates were
able to offer protection against cleavage even in the presence of SDS, suggests some involvement at or near the active site of the protein. Alternatively, these substrates may cause a conformational change that limits the accessibility of loci involved in degradation. A high concentration of EDTA (20 mM) also appeared to convey a slight protection from the degradation (Figure 2.7 B lane5), though lower concentrations as found in the protease inhibitor cocktail by Sigma (5 mM) had no affect. This could be a result of uneven gel staining (compare Figure 2.7 B lanes 4 and 6), or possibly the result of protection provided by apoenzyme formation that is enhanced by the presence of high concentrations of EDTA. Degradation of a prepared apoenzyme sample has not yet been investigated. It was also noted that the fragmentation would consistently produce two fragments in larger quantities, of approximate ~ 24 kDa and ~ 46 kDa sizes. These fragments were transferred to a nitrocellulose filter and submitted for N-terminal sequencing. The results show the 24 kDa fragment to be the N-terminal of the entire subunit, and the 46 kDa fragment appears to be the C-terminal end, making it quite possible that these two predominant fragments are generated by a single cleavage event. The sequence from the 24 kDa fragment shows the cleavage occurs on the carboxy side of Asp236 at an Asp-Pro peptide bond.
Figure 2.6. Protection of thermal cleavage of CbbT mutant C160A by alkylation with iodoacetamide (IAM). Samples (4 µg per lane) were prepared identically with or without IAM; samples were boiled for 20 min. Lanes: (1), no heat, - IAM; (2), heat, - IAM; (3), no heat, + IAM; (4), heat, + IAM. Protein bands were visualized by silver staining.
Figure 2.7. Thermal degradation of CbbT. (A) Temperature-dependence of CbbT autolysis. Samples of purified CbbT were heated at different temperatures for 20 min in SDS-PAGE loading buffer prior to electrophoresis; 10 µg of protein was loaded per lane and the gel was stained with Coomassie. Lanes (temperature): (1), 30 °C; (2), 40 °C; (3), 50 °C; (4), 60 °C; (5), 70 °C; (6), 80 °C; (7), 90 °C; (8), 100 °C. (B) Protection of degradation by substrates after boiling for 20 min prior to SDS-PAGE; 5 µg of protein was loaded per lane and the gel was silver stained. (Lane), additive; (1), no additives; (2), 55 mM R5P; (3), 1 mM Xu5P; (4), 55 mM R5P and 1 mM Xu5P; (5), 20 mM EDTA; (6), no additives; (7), no additives; (8), unheated control. In both gels, arrows indicate the two dominant cleavage products that were submitted for N-terminal analysis.
Figure 2.7
Similar denaturation-dependent intramolecular cleavage has been reported to occur for other proteins including egg-white lysozyme (72), porcine tubulin (73), \( \beta \)-galactosidase from *E. coli* (74), human \( \alpha_2 \)-macroglobulin (\( \alpha_2 \mathrm{M} \)) (75), and \( \alpha \)-amylase from *Aspergillus oryzae* (76). In \( \alpha_2 \mathrm{M} \), the fragmentation occurred specifically at one site in the subunit at temperatures above 37 \( ^\circ \mathrm{C} \) in the presence of SDS, or at 80 \( ^\circ \mathrm{C} \) in the absence of SDS (77). In addition this cleavage could be inhibited by the presence of methylamine, an inhibitor of \( \alpha_2 \mathrm{M} \). The site of cleavage was determined to be on the amino side of the methylamine-reactive glutamyl residue (77). Therefore this appears to be a mechanism unique to proteins that incorporate methylamine. A study on the cleavage of tubulin demonstrated that the cleavage under high temperature and denaturing conditions occurs primarily at Asp-Pro or Asp-Cys peptide bonds above pH 6, however not all fragments could be identified and cleavage at other Asp bonds was not ruled out (73). A detailed investigation in the fragmentation of \( \alpha \)-amylase confirmed a similar mechanism, in which specific cleavages at Asp-Pro and Asp-Cys bonds were displayed, with Asp-Cys cleavage requiring a free thiol as opposed to cystine (76). In all the literature reporting thermally-induced autolysis under denaturing conditions, there were no instances found where specific substrates inhibited degradation, as demonstrated here for CbbT and Tkt. Though the donor substrate Xu5P was not able to prevent degradation at a concentration of 1 mM, it is believed that higher concentrations of substrate would confer protection in a manner similar to F6P. This thermal degradation may not be physiologically relevant since it requires denaturant and relatively high temperatures to occur, however no long-term studies at physiological temperatures have been conducted as of yet.
2.3.8 TRANSKETOLASE OPTIMAL PH

Most transketolases exhibit a similar pH optimum in the range from 7.4 to 8.2 (78) with the exception of TkIA of *E. coli* which is optimally active from pH 8.0 to 8.5 (71). This was also true for CbbT and Tkt. Interestingly, when the buffer Tris was used to explore the pH profile of CbbT, activity showed little dependence on pH throughout the range explored (7.6 to 9.0). However due to the inhibitory effect of Tris on CbbT all specific activities were low. With the buffers HEPES-NaOH and Glycylglycine-HCl, CbbT exhibited a pH optimum around 7.8 (Figure 2.8 A). Tkt displayed more variation, depending on the buffer that was used, but no inhibition by Tris was observed. Upon comparing all buffers investigated it appears that the pH optimum for Tkt is more alkaline than for CbbT, falling in the range between pH 8.3 and 8.6 (Figure 2.8 B). Under their respective optimal activity conditions, both enzymes displayed similar specific activities; 36.9 U/mg for CbbT in HEPES-NaOH pH 7.8 and 34.3 U/mg for Tkt in Glycylglycine-HCl, pH 8.2.
Figure 2.8. Activity levels at different pH values for CbbT (A) and Tkt (B). All assays were performed in the presence of 200 mM buffer, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME at 25 °C. A unit of activity is defined as μmol substrate consumed min⁻¹. The pH was measured with a pH meter immediately after completion of the assay.
Figure 2.8
2.3.9 KINETICS

The standard physiological substrates used for kinetic analysis were R5P and Xu5P. For CbbT, data was collected at the optimal pH of 7.8 as well as at 6.7 and 8.5 for comparison. We found that with CbbT the $K_m$ values were significantly affected by pH, following a trend of increasing $K_m$ with increasing pH. A two- to three-fold increase in the Michaelis constants of both substrates was observed with each increasing pH value examined (6.7, 7.8, and 8.5). In spite of displaying the slowest turnover rate at pH 6.7, the accompanied increase in relative affinity for substrates resulted in similar catalytic efficiencies as at pH 7.8, whereas a decrease in $k_{cat}/K_m$ is observed at pH 8.5 for both substrates (Table 2.4). Tkt $K_m$ values for both substrates were determined at pH 7.8 and found to be within the range of values reported for other transketolases (Table 2.5). CbbT displayed the lowest reported $K_{Xu5P}$ that could be found in the literature, with a value of 11 µM at pH 7.8. R5P values are within the range of those reported for other organisms.
<table>
<thead>
<tr>
<th>pH</th>
<th>$k_{cat}$ (sec$^{-1}$)</th>
<th>$K_{R5P}$ (µM)</th>
<th>$k_{cat} / K_{R5P}$ (sec$^{-1}$ M$^{-1}$)</th>
<th>$K_{Xa5P}$ (µM)</th>
<th>$k_{cat} / K_{Xa5P}$ (sec$^{-1}$ M$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.7</td>
<td>14.2</td>
<td>356 ± 17</td>
<td>3.99 X 10^4</td>
<td>5.6 ± 1.2</td>
<td>2.54 X 10^6</td>
</tr>
<tr>
<td>7.8</td>
<td>42.7</td>
<td>949 ± 55</td>
<td>4.50 X 10^4</td>
<td>11 ± 2.0</td>
<td>3.89 X 10^6</td>
</tr>
<tr>
<td>8.5</td>
<td>29.0</td>
<td>2,320 ± 260</td>
<td>1.25 X 10^4</td>
<td>33.6 ± 5.8</td>
<td>8.63 X 10^5</td>
</tr>
</tbody>
</table>

**Table 2.4.** Kinetic constants calculated for CbbT at three separate pH values. The standard assay consisted of 100 mM buffer of either MOPS-NaOH (pH 6.7), HEPES-NaOH (pH 7.8), or glycylglycine-HCl (pH 8.5) with 2 mM MgSO$_4$ and 0.1% BME, and 0.1 mM ThDP. $K_m$ values were collected using variable amounts of one substrate while maintaining the other substrate constant at a saturating non-inhibitory concentration. A MW of 69.3 kDa was used for the $k_{cat}$ calculations.
<table>
<thead>
<tr>
<th>Source†, Taxonomy</th>
<th>$V_{\text{max}}$ (µmol min$^{-1}$ mg$^{-1}$)</th>
<th>$K_{R5P}$ (µM)</th>
<th>$K_{Xa5P}$ (µM)</th>
<th>pH</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human$^{a}$, <em>Homo sapiens</em></td>
<td>4.8</td>
<td>360</td>
<td>390</td>
<td>7.5</td>
<td>(79)</td>
</tr>
<tr>
<td>Maize$^{a}$, <em>Zea mays</em></td>
<td>25.3</td>
<td>581</td>
<td>403</td>
<td>7.5</td>
<td>(8)</td>
</tr>
<tr>
<td>Spinach chloroplasts, <em>Spinacia oleracea</em></td>
<td>14.0</td>
<td>330</td>
<td>67</td>
<td>7.5</td>
<td>(80)</td>
</tr>
<tr>
<td>Pepper$^{a}$, <em>Capsicum annuum L.</em></td>
<td>30.1</td>
<td>750</td>
<td>95</td>
<td>7.6</td>
<td>(81)</td>
</tr>
<tr>
<td>Tkl1$^{a}$, <em>S. cerevisiae</em></td>
<td>37.0</td>
<td>146</td>
<td>162</td>
<td>7.6</td>
<td>(82)</td>
</tr>
<tr>
<td>Torula yeast*, <em>Pichia jadinii</em></td>
<td>43</td>
<td>430</td>
<td>77</td>
<td>7.5</td>
<td>(83)</td>
</tr>
<tr>
<td>TktA*, <em>Escherichia coli</em></td>
<td>50.4</td>
<td>1,400</td>
<td>160</td>
<td>8.5</td>
<td>(71)</td>
</tr>
<tr>
<td>CbbT$, *R. sphaeroides$</td>
<td>37.0</td>
<td>949</td>
<td>11</td>
<td>7.8</td>
<td>This work</td>
</tr>
<tr>
<td>Tkt$, *R. sphaeroides$</td>
<td>27.5</td>
<td>420</td>
<td>87</td>
<td>7.8</td>
<td>This work</td>
</tr>
</tbody>
</table>

**Table 2.5.** Comparison of catalytic properties of CbbT and Tkt with transketolases from various sources. For CbbT and Tkt the standard assay consisted of 100 mM buffer of either HEPES-NaOH (CbbT) or Tris-HCl (Tkt) pH 7.8, 2 mM MgSO$_4$ and 0.1% BME, and 0.1 mM ThDP.

† $a$ – enzyme was from a recombinant expression system.

* – These enzymes were assayed at 30 °C. All others were reported at 25 °C.
2.3.10 APPLICATIONS OF TRANSKETOLASE FOR SUBSTRATE PREPARATIONS

During our research on transketolase, Sigma-Aldrich discontinued the availability of Xu5P. There were alternate vendors, however in all cases the purity of Xu5P was reported as 80% or less. This new decrease in availability also made the synthesis of Xu5P a cost effective alternative. As a demonstration of the synthetic potential of transketolase, we utilized CbbT to achieve the synthesis of Xu5P. The use of transketolase for synthesis of this compound was originally proposed by Racker in 1955 (84) but the reaction scheme was greatly improved recently by the availability of triose phosphate isomerase (TPI) (45, 84). The setup was further modified to improve the stability of CbbT throughout the reaction period by providing anaerobic conditions. The schematic in Figure 2.9 demonstrates the multi-enzymatic approach utilized in this one-pot synthesis.
Figure 2.9. Scheme showing the synthesis of Xu5P. F1,6bP, fructose 1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; G3P, glyceraldehyde 3-phosphate; β-Hp, β-hydroxypyruvic acid; FbpA, fructose bisphosphate aldolase; TPI, triosephosphate isomerase; TK, transketolase (CbbT).
The starting material, lithium β-hydroxypyruvate, was synthesized from bromopyruvate by alkalinization with LiOH (85) and then purified by first washing with activated charcoal, followed by precipitation. The final yield from this step was 2.83 gm of solid β-hydroxypyruvic acid lithium salt and corresponded to a 43% yield. This was similar to the published yield of 55%.

The enzymatic synthesis was carried out similarly to the published method with the exception that the reaction was constantly purged with a stream of N₂ gas. This anaerobic environment was employed to aid in the stability of CbbT throughout the reaction. Since a proton is consumed during each cycle of catalysis by transketolase, the reaction progress could be conveniently monitored by HCl consumption (86) (Figure 2.10). A pH controller was used throughout the 67 h reaction to maintain the pH at 7.3 (within 0.02 pH units).
The synthesis resulted in 330 mg of Xu5P (as determined enzymatically), which corresponds to a 40% yield based on the amount of F1,6bP used. No final pH adjustment was performed, and the product was stored at low pH in 2 ml aliquots at -80 °C. A sample was lyophilized for H1 NMR analysis. The NMR spectrum was not very clean, dominated by a large signal due to the high concentration of formate present. Degradation of Xu5P was believed to have occurred during lyophilization, based on the appearance of a dominant brown residue. At first it appeared that the Xu5P also behaved
differently enzymatically when compared to Xu5P obtained from Sigma. However it was later realized that the low pH of the synthesized Xu5P interfered with its assay. A 2 ml aliquot was carefully pH adjusted to 6.7 on ice and this sample compared quite well (i.e. exactly) to commercially obtained Xu5P. Though the purity could not clearly be validated by NMR, the newly synthesized Xu5P behaved identical to the commercially obtained substrate. Further, no inhibitory effects were observed in enzymatic assays that used the Xu5P preparation as an acceptor substrate in a transketolase reaction and identical $K_m$ and $V_{max}$ values were obtained using the newly synthesized Xu5P.

2.3.11 TRANSKETOLASE EXPRESSION IN R. SPHAEROIDES

The expression and accumulation of transketolase was also investigated in two different strains of *R. sphaeroides* grown under different conditions. *R. sphaeroides* is capable of aerobic chemoheterotrophic growth in the presence of an organic carbon source, photoautotrophic growth utilizing CO$_2$ as sole carbon source and light energy, and photoheterotrophic growth in the presence of an organic carbon source using light energy. Strain CAC is a spontaneous mutant capable of aerobic chemoautotrophic growth in the dark (52), utilizing CO$_2$ as a carbon source and H$_2$ as the energy source. Strain 1884 is an insertional mutant in the *cbbFI* gene of wild type strain HR, in which the *cbbI* operon was inactivated because of polar effects on downstream genes, causing a concomitant elevated expression of genes of the *cbbII* operon (53). The highest transketolase specific activities detected were for photoautotrophic cultures of both strains and the chemoautotrophic culture of CAC. Western analysis using rabbit antibodies generated against CbbT corroborated the enzyme activity results (Figure
2.11), suggesting an effect on enzyme synthesis. The antibodies have been shown to
cross react somewhat with purified Tkt, though to a weaker extent and require several µg
of protein to be noticeable. This reduced affinity for Tkt could explain why no
transketolase bands are detected under conditions of low CbbT expression. Also
surprising was the low transketolase activities measured for these same cultures, though
CbbT is reduced, Tkt would be expected to be expressed at some level. This could be an
effect of measuring the activity at pH 6.9 since we have shown Tkt has a higher optimal
pH range than CbbT. We have also been baffled by the occurrence of two bands
commonly observed by native PAGE analysis of a homogenous (by SDS-PAGE) CbbT
preparation. Though we have not identified the cause of the alternate bands, we have
shown that both bands display transketolase activity. As seen in Figure 2.11, two bands
were also observed in crude extracts from R. sphaeroides by Western analysis.
Figure 2.11. Preferential accumulation of CbbT under autotrophic conditions in *R. sphaeroides*. (A) Purified CbbT and *R. sphaeroides* extracts were run on a 3-step gradient native PAGE gel consisting of 8/10/12% polyacrylamide from top to bottom. (B) Western analysis performed using a duplicate gel. TK, purified CbbT; PA, photoautotrophic; PH, photoheterotrophic; CH, chemoheterotrophic; CA, chemoautotrophic. 1 µg of purified CbbT and 15 µg of extract were loaded per lane. Specific activities of the extracts are presented on the Western immunoblot (µmol min\(^{-1}\) mg\(^{-1}\)). Strains are indicated by the brackets (CAC and 1884). Purified CbbT would often display two bands after native PAGE analysis.
Figure 2.11
2.4.1 CONCLUSIONS

Most organisms studied to date display two to three transketolases encoded within their genome. Research on the differential expression of these genes is weak and biochemical data comparing different transketolases of the same organism is completely lacking. In fact, it appears that this is the first enzymatic study of two transketolases from the same organism. The cost effective method of protein over-expression in *E. coli* strain GJ1158 has proven to be a useful system for the production of large amounts of soluble CbbT and Tkt. An efficient purification scheme was designed for both CbbT and Tkt and both enzymes were purified to homogeneity. Despite the numerous techniques utilized to produce pure or partially purified preparations of transketolase from various sources, Tkt appears to be the first to be resolved using cation exchange chromatography.

Though similar in amino acid sequence, with an identity of 58% and a similarity of 69%, these proteins display quite different isoelectric properties as observed from the calculated pI values and by their different behavior during ion exchange chromatography. This difference can also be viewed by comparing the surface charge distribution of the molecular models generated for the two proteins (Figure 2.12). The electrostatic surface model generated for Tkt displayed large regions of positively charged residues, quite different from that generated for CbbT, which is dominated by negative charges. Of the three available transketolase structures, *E. coli* and *Zea mays* display a similar dominantly negative electrostatic surface, whereas *S. cerevisiae* shows similar patches enriched with positive charge. It is possible that these patches on the surface of Tkt allow for unique protein-protein or protein-substrate interactions. However, both the CbbT and Tkt models are theoretical, requiring that the actual structures first be solved to verify this
observation. In hopes of achieving this, a sample of purified CbbT was sent to a collaborating x-ray crystallography lab. Diffraction data has been collected for crystals diffracting to a resolution of 2.6 Å and is currently in the refinement phase of model building.
**Figure 2.12.** Molecular models of CbbT and Tkt displaying the electrostatic potential mapped to the surface. The computation was by the Coulomb method using Swiss-Pdb Viewer (87) v 3.7, which considers only the charged side chains (Arg, Lys, Glu, Asp) at pH 7.0. The kT/e electrostatic units used as coloring parameters were +1.5 kT/e for blue, 0 kT/e for white, and -1.5 kT/e for red.
Figure 2.12
We have found that CbbT and Tkt do indeed display different biochemical properties. Notably, they differ in pH profile and in substrate affinity, suggesting each enzyme may be tailored to perform in a preferred cellular environment. It was shown that the transketolase reaction operates near thermodynamic equivalence, therefore the direction the reaction proceeds would be determined by substrate concentration and enzyme affinity ($K_m$). Since CbbT functions in the Calvin-Benson-Bassham cycle, it would be expected to drive the reaction preferentially in one direction, which includes the production of R5P. This could explain the relatively high $K_{R5P}$ we observed for CbbT compared to Tkt which, participating in the pentose phosphate pathway, would be expected to more readily reversible. We have also observed that CbbT displays a $K_{Xu5P}$ value of 11 µM, the lowest value found reported in literature is for the rat liver transketolase with a value of 22 µM (88). However this rat enzyme was found to be unable to utilize β-hydroxypyruvate (89), which is a useful donor substrate for synthesis using transketolase. When using β-hydroxypyruvate, formation of the active glycolaldehyde results in the release of CO$_2$, rendering the reaction irreversible (90). We have shown that CbbT is able to effectively utilize β-hydroxypyruvate as a donor substrate in the synthesis of Xu5P. As a testament of transketolase’s attraction in enzymatic synthesis, 330 mg of Xu5P was produced in a simple one pot scheme that utilized two additional coupling enzymes to provide transketolase with the necessary substrates. The fact that CbbT displays the lowest reported $K_{Xu5P}$ may make this a useful enzyme in future substrate specificity studies, or even from an industrial standpoint.
Though it was once believed that proteins were stable for up to 30 min of boiling in preparation for SDS PAGE analysis (91), there have been reports of proteins that degrade under these boiling and denaturing conditions. We have shown that like *E. coli* transketolase TktA, CbbT and Tkt from *R. sphaeroides* also undergo a thermal degradation. Though it has only been briefly reported for the *E. coli* enzyme, due to the seemingly rare nature of this reaction it remains probable that this could be a trait common amongst all, or perhaps a sub-class of, transketolases.

Truly interesting is the observed oxygen inactivation of *E. coli* TklA bound on an insoluble support (47) and as we have shown, the soluble CbbT and Tkt from *R. sphaeroides* are also affected by oxygen. Though, somewhat similar in obscurity to the thermal-induced degradation, oxygen-mediated inactivation has not been reported for other well-studied transketolases. In general, oxygen-induced irreversible inactivation of enzymes is a poorly documented phenomenon. In hopes of understanding the mechanism involved in this type of enzyme inactivation, the major thrust of our research focused on characterizing the product of this reaction, namely the inactivated CbbT protein.
CHAPTER 3

OXYGEN INACTIVATION OF TRANSKETOLASE

3.1 INTRODUCTION

An *E. coli* mutant devoid of Fe- and Mn-containing superoxide dismutases (SODs), which caused an increase in the level of superoxide radicals, displayed an aromatic amino acid auxotrophy (92, 93). It was later found that this auxotrophy was due to the inactivation of transketolase since transketolase is necessary for the generation of E4P used in the first step of the aromatic amino acid pathway (94). Inactivation occurred only in the presence of donor substrate and an external source of superoxide and was believed to occur by oxidation of the 1,2-dihydroxyethyl thiamine diphosphate intermediate formed during transketolase catalysis. This paracatalytic oxidation of the intermediate has been known for some time and involves the transfer of two electrons to the intermediate, resulting in the production of glycolate. Initially discovered in fragmented spinach chloroplasts (95), it was later found that various sources of oxidants were capable of driving this reaction (96-98). In addition, it was noted that enzyme inactivation is irreversible and occurred concurrent with the oxidation reaction (97). The precise mechanism of inactivation is unknown, though the involvement of a transient reactive intermediate has been suggested (99).
Crude preparations of *E. coli* transketolase were shown to undergo inactivation in the presence of air, which is prevented by careful storage of the protein under anaerobic conditions (47). In this instance, the enzyme was immobilized to a support matrix and had not been exposed either to an exogenous source of reactive oxygen species (ROS) or to donor substrates during storage. This would suggest an oxidative inactivation mechanism distinct from the previously mentioned reaction with the intermediate; however this point was not further explored. We have shown that purified CbbT and Tkt both undergo similar substrate-independent oxygen inactivation processes. To better understand this type of inactivation we decided to biochemically characterize the inactive form of CbbT and compare it to the fully active form of the enzyme. To achieve this we investigated the effects of oxygen on enzyme activity during transketolase gene over-expression in *E. coli* strain GJ1158. We defined a consistent method to inactivate purified CbbT, and utilized several spectral techniques to compare the active and inactive forms of the enzyme. Our results indicated that the amount of oxygen present in cultures during recombinant protein accumulation is a key determinant of the nascent activity of the final CbbT product. The rate of inactivation of purified CbbT was dependent on the concentration of O₂ and the process was diminished by providing increasing concentrations of reductant, though inactivation was still irreversible. Further, it was found that the inactive form of CbbT exhibited a clear difference in fluorescence spectra, as well as a smaller difference by circular dichroism (CD), suggesting structural changes occurred during inactivation. In support of this, different free energies of unfolding (\(\Delta G(H_2O)\)) were determined for the active vs. the inactive form of CbbT.
3.2 MATERIALS AND METHODS

3.2.1 LOW OXYGEN FERMENTATIONS

The Ohio State University Fermentation Facility was used to conduct various small volume fermentations under controlled oxygen conditions. An Applicon 3 fermentor system was used in conjunction with the Applicon controller, which monitored various conditional variables. The rate of oxygen allowed to enter the fermentor was controlled by mixing air and argon gas as set by calibrated rotameters (Matheson Tri-Gas, Montgomeryville, PA). The rate of mixing was maintained by using a fixed impeller speed of 600 rpm. The 3 l sealed fermentor was jacketed, allowing for the use of a temperature controller. Variables monitored in all fermentations included pH, % soluble oxygen, and OD_{600}. In later fermentations, the redox potential of the growth medium was also monitored with a redox probe. The dO_{2} probe was internally calibrated prior to inoculation by sparging the media, first with 100% argon for 15 min where the probe was set to a value 0, and then by sparging with 100% air for 15 min, at which point the probe was set to a value of 100. OD_{600} was measured with the use of a peristaltic pump that constantly sampled culture from the fermentor; this was run through a flow cell cuvette placed in a spectrophotometer, with these bacterial cells then delivered back into the fermentor. A series of 2 l microaerobic fermentations were conducted over the period of a few months. All fermentations consisted of 2 l of media (LB minus salt), pH adjusted to 7.5 in the 3 l fermentor vessel. Inoculation was with 2 ml of an overnight culture and initial growth temperature was at 37 °C, using an impeller speed of 600 rpm. Immediately prior to induction the temperature was reduced to 25 °C. The culture was induced with 174 ml of 5 M NaCl, to yield a final concentration of 0.4 mM, when the
culture reached an OD$_{600}$ of 0.5 - 0.6. Growth was allowed to proceed for 16 h. The culture was then harvested by centrifugation and the pellet was washed once with buffer (20 mM HEPES-NaOH, pH 7.8, 0.1% BME), a spatula was used to remove a small amount of cell material for immediate analysis, with the remaining cell material stored frozen at -80 °C. All $cbbT$ over-expression experiments were performed with fermentations that employed the pET11CbbT plasmid transformed into the salt inducible $E. coli$ strain GJ1158 (51).

3.2.2 ANAEROBIC OVER-EXPRESSION IN ENHANCED MEDIA

Terrific broth (TB) media was prepared as outlined in Current Protocols In Molecular Biology (100). To produce one liter of media, 12 g tryptone and 24 g yeast extract were dissolved into 900 ml of H$_2$O and autoclaved. To this, 100 ml of a filter sterilized solution consisting of 0.17 M KH$_2$PO$_4$ and 0.72 M K$_2$HPO$_4$, was added. For the nitrate and glucose supplemented media, 3 g of KNO$_3$ and 9 g of glucose were added per liter of media. For the glycerol and fumarate enhanced media, 20 ml of glycerol and 4 g of fumarate were added per liter of media. The small scale cultures consisted of 110 ml of media in 125 ml stoppered Erlenmeyer flasks. They were inoculated with 110 µl of an overnight culture grown in TB containing 50 µg/ml of ampicillin. Initial growth was at 37 °C with an agitation speed of 170 rpm. Cultures were induced at 25 °C by adding 12.5 ml of a 4 M NaCl solution (autoclaved) at an OD$_{600}$ of 0.5. Following a 16 h induction, the cells were harvesting by centrifugation and washed once with buffer (20 mM HEPES-NaOH, pH 7.8, 0.1% BME); pellets were stored at -80 °C until further analysis.
3.2.3 TIME COURSE OF INACTIVATION

In order to standardize conditions, purified enzyme stored at -80 °C was thawed and then desalted into the buffer of choice using a 1.0 x 30 cm Superose-12 column (Amersham Biosciences, San Francisco, CA). Typically the buffer consisted of 20 mM HEPES-NaOH, pH 7.8, 2 mM MgSO₄, 0.1 mM ThDP, which was first filtered and degassed, and then BME was added to 0.1%. The desalted sample was then diluted to the desired concentration, typically a volume where approximately 10 - 20 µl could be used in an activity assay to yield a satisfactory rate. The sample was then placed in a serum vial at least 3 times the volume of the sample. This was placed on ice and sealed by crimping a metal seal around the rubber septa. A gentle stream of the desired gas was then allowed to sparge over the head space of the vial using high gauge needles to minimize damage to the seals. After 20 min sparging, the needles were removed and the vials were placed in a 25 °C water bath; this was designated time zero. At the desired time points thereafter aliquots were removed from the vials using a Hamilton gas tight syringe that had been sparged with the same gas in the headspace.

3.2.4 CIRCULAR DICHROISM

Near UV CD measurements were taken using an Aviv model 62A DS spectrometer. Spectra were obtained by scanning from 205 to 280 nm at 1 nm intervals with a bandwidth of 1 nm and a 2 sec signaling average; five scans were averaged for each sample. Data were collected using quartz cuvettes with a 1 mm pathlength and the temperature controller set to 25 °C. Protein concentration was 1 mg/ml in 20 mM HEPES-NaOH buffer pH 7.8, containing 2 mM MgSO₄, 0.1 mM ThDP, and 0.1% BME.
One sample was stored under N\textsubscript{2} gas and the other under an air headspace and allowed to inactivate over 24 h at 25 °C, at which point the sample exposed to air displayed 30% the activity of the argon stored sample.

3.2.5 FLUORESCENCE

A Jobin Yvon – Spex model Fluoromax-3 spectrofluorometer was used for fluorometric data collection. The automated sample changer and Neslab RTE-111 temperature controller accessories were used with Starna 10 mm quartz fluorometry cuvettes fitted with airtight septa screw caps. Standard conditions contained 2 mM MgSO\textsubscript{4}, 0.1% BME, 0.1 mM ThDP, 100 mM buffer, and 800 nM of enzyme (as calculated by monomeric molecular weight) at 25 °C. The sample was excited at 280 nm and fluorescence emission was scanned from 295 to 450 nm.

3.2.6 UREA UNFOLDING CURVES

Samples of CbbT in variable concentrations of urea were prepared using a 1 mg/ml stock solution of CbbT and a master mix of 10X strength that consisted of 500 mM HEPES-NaOH pH 8.0, 10 mM DTT, 20 mM MgSO\textsubscript{4}, and 10 mM ThDP. A urea stock of 8 M and water were used to bring the final concentration of urea to one of the following empirically determined values (M): 0, 0.15, 0.3, 0.45, 0.6, 0.75, 0.9, 1.2, 1.5, 1.65, 1.8, 1.95, 2.1, 2.4, 2.55, 2.8, 3, 3.2, 3.4, 3.8, 4.2, 4.6, 5, 5.4, 5.8, 6.2, 6.6, and 6.8. The samples were prepared in 2 ml serum vials that were immediately sealed and sparged with a gentle flow of nitrogen gas for 20 min. Following the gas sparge, the samples were placed in a 25 °C water bath and allowed to equilibrate for 4 h. Samples were then
transferred to quartz fluorometry cuvettes with a path length of 1 cm and data was collected according to the following parameters. Excitation at 280 nm through a 5 nm slit, with a 0.5 sec integration time, collecting emission data through a 5 nm slit in 0.5 nm increments from 295 to 450 nm. Peak emission wavelength was determined for each sample using Data Max software v2.20. Data analysis was conducted using Sigma Plot v8.0. Protein unfolding curves were fit to the following positive 6 parameter exponential equation:

\[ Y = \frac{1}{1+1/ \exp((m \cdot x + g) / \text{(RT)})) \cdot ((u_i + u_s \cdot x) - (l_i + l_s \cdot x)) + l_i + l_s \cdot x \]

where a value of 1.987 cal mol\(^{-1}\) K\(^{-1}\) was used for \(R\), the ideal gas constant and temperature, \(T\) was 298 K. The variables are defined as follows: \(Y\) is the fraction of protein folded; \(x\) represents the concentration of urea (mol L\(^{-1}\)); \(u_i\) is the upper intercept; \(l_i\) is the lower intercept; \(u_s\) is the upper slope (the post transition baseline); \(l_s\) is the lower slope (the pre transition baseline); \(g\) is the free energy of unfolding in H\(_2\)O (cal mol\(^{-1}\)); \(m\) is the dependence of \(g\) on the denaturant (kcal mol\(^{-1}\) M\(^{-1}\)).

### 3.3 RESULTS AND DISCUSSION

#### 3.3.1 LOW OXYGEN FERMENTATIONS

Previously we demonstrated that CbbT activity in crude extracts of cultures over-expressing the \(cbbT\) gene was strongly dependent on the amount of oxygen present during \(cbbT\) induction. Since initial attempts to scale up over-expression to a 10 l fermentor under standard aeration failed, \(cbbT\) was over-expressed using decreasing oxygen concentrations. Following an initial small scale experiment using Erlenmeyer flasks, several 2 l fermentations were conducted under controlled conditions. In addition,
the use of various rich media was explored under microaerobic conditions on a smaller scale.

*E. coli* typically utilizes O$_2$ as the terminal electron acceptor, but as a facultative anaerobe it is capable of utilizing alternative acceptors under oxygen-limiting conditions (101). For this reason three different kinds of rich growth media, two of which were enhanced with alternate electron acceptors, were also tested for their ability to increase the accumulation of active CbbT under semi-anaerobic conditions. These were Terrific Broth (TB), TB enhanced with fumarate and glycerol (TBfg), and TB enhanced with nitrate and glucose (TBng). The cultures were grown on a small scale, and induced overnight at room temperature. CbbT activities of crude extracts obtained from cells grown under all conditions was low, with values of 0.14 U/mg for TBng, 0.33 U/mg for TB, and 0.56 U/mg for TBfg. Growth in all cases was poor and resulted in low net cell mass yields. Rich media and anaerobic conditions were therefore no longer pursued for over-expression purposes.

Initial experiments tested the effect of aeration on CbbT accumulation. This was accomplished by using Erlenmeyer flasks of different volumes ranging from 125 ml to 1,000 ml while maintaining a constant volume of media (100 ml). The resulting head space to media ratio is directly related to the amount of air that is made available to the cells. All cultures were induced at an OD$_{600}$ of 0.5 and allowed to proceed at 25 °C for 16 h. Upon comparison of the specific activities of these cultures, it became apparent that the amount of oxygen introduced during *cbbT* gene over-expression drastically affected CbbT activity (*Figure 3.1*).
Figure 3.1. CbbT activity under variable aeration conditions. Transketolase specific activity of the crude extracts is plotted versus the Erlenmeyer flask size; 100 ml liquid culture volume was used in all cases.

These preliminary experiments, which demonstrated a clear effect of aeration on CbbT crude extract activity, was further explored using larger scale, highly controlled fermentations. To achieve variable oxygen concentrations, house air was mixed with argon at different proportions, controlled through the use of rotameters. Flow rates of the
gases were determined using calibration charts for the various rotameters. Seven fermentations were performed under similar conditions, differing only in the amount of air supplied during growth. A standard air flow rate for this size of fermentor is 1.5 l min\(^{-1}\). Since previous 10 l fermentor attempts failed under comparable air flows, we provided our first 2 l fermentation with equal flows of argon and air and were able to achieve a weak amount of activity. Lower concentrations of air were therefore chosen for all subsequent fermentations. The consequences of \textit{cbbT} over-expression under anaerobic conditions were also examined following a pre-induction growth in the presence of oxygen. The remaining cultures were over-expressed in the presence of air, with flow rates ranging from 0.01 to 1.5 l/min. Growth curves labeled by the net air flow are shown in \textbf{Figure 3.2}. 
Figure 3.2. Growth curves of *E. coli* 2 l fermentations cultures at variable O₂ levels.

The point of induction is denoted by the horizontal double arrow drawn around an OD₆₀₀ of 0.5. The insert depicts the rate of air bubbled through the fermentor cultures in l min⁻¹.

It is apparent that there is a general trend under conditions of higher air flow to result in a higher final OD or a denser cellular mass, showing that oxygen is limiting in all cases below the highest air flow rate. This was further supported by the lack of dissolved O₂ detected in any of the other cultures by the time post logarithmic growth had
been reached (data not shown). Upon analysis of transketolase activity of the crude extracts it was found that there is a trend of higher specific activities from cultures grown under lower air flows. This trend is disrupted by the fermentation over-expressed under anaerobic conditions as well as the fermentation over-expressed at the lowest air flow of 0.01 l min\(^{-1}\). These low activities are most likely the result of \textit{E. coli} being too stressed under anaerobic conditions to produce a significant amount of recombinant protein.

Redox probes are commonly used to monitor anaerobic conditions in culture media. They can be used to measure trace amounts of dissolved oxygen at levels too low for dO\(_2\) probes to detect. The redox probe therefore became a more useful sensor than the dO2 probe to report the relative levels of O\(_2\) in the media. This is reflected by comparing the final redox values obtained at the end of each culture, which ranged from +70 mV for the highest air flow to -70 mV for the lowest air flow. O\(_2\) is considered limiting in the range of -50 to -100 mV (and lower). Table 3.1 summarizes the data obtained from these fermentations.
<table>
<thead>
<tr>
<th>Air (l min$^{-1}$)</th>
<th>Ar (l min$^{-1}$)</th>
<th>Air/total</th>
<th>Air/Ar</th>
<th>Final OD_{600}</th>
<th>Final redox* (mV)</th>
<th>Activity (U mg$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.13</td>
<td>0.000</td>
<td>0.000</td>
<td>0.67</td>
<td>-70</td>
<td>2.67</td>
</tr>
<tr>
<td>0.01</td>
<td>0.70</td>
<td>0.016</td>
<td>0.016</td>
<td>0.62</td>
<td>-51</td>
<td>1.35</td>
</tr>
<tr>
<td>0.02</td>
<td>1.46</td>
<td>0.014</td>
<td>0.014</td>
<td>0.75</td>
<td>-38</td>
<td>3.69</td>
</tr>
<tr>
<td>0.09</td>
<td>1.46</td>
<td>0.059</td>
<td>0.063</td>
<td>1.18</td>
<td>nd</td>
<td>2.50</td>
</tr>
<tr>
<td>0.26</td>
<td>1.10</td>
<td>0.192</td>
<td>0.237</td>
<td>2.00</td>
<td>37</td>
<td>2.25</td>
</tr>
<tr>
<td>0.40</td>
<td>1.96</td>
<td>0.170</td>
<td>0.204</td>
<td>1.82</td>
<td>48</td>
<td>2.01</td>
</tr>
<tr>
<td>1.52</td>
<td>1.52</td>
<td>0.500</td>
<td>1.000</td>
<td>2.16</td>
<td>70</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 3.1. Gas flows and measured parameters at variable oxygen levels in 2 l fermentor vessels. Final OD$_{600}$ and final redox values were read prior to harvesting the cells.

Specific activity is listed as U mg$^{-1}$, where a U is defined as µmol min$^{-1}$.

*nd – not determined

In order to precisely control the ratio of gas supplied to future fermentations it is recommended that the rotameters are calibrated and a tank source of air used to better provide a consistent flow. In addition, only the ratio of gases should be changed while maintaining a consistent total gas flow for every culture. This possible source of error made it difficult when comparing some cultures that had similar air/argon ratios with different net gas flow as the total gas flow determines the amount of time air is introduced into the reactor and is permitted to dissolve into the media. It is hoped that these optimized conditions of the small-scale fermentation may be readily applied to a larger scale fermentation allowing for the accumulation of massive amounts of highly active protein for future applications.
3.3.2 OXYGEN INACTIVATION OF PURIFIED TRANSKETOLASE

Purified CbbT was used in order to quantitate inactivation in the presence of oxygen \textit{in vitro}. Either argon or N\textsubscript{2} gas were used as anaerobic controls or for mixing purposes to generate different concentrations of O\textsubscript{2}. By exposing a sample of CbbT to oxygen under different conditions and monitoring enzyme activity over time, rates of inactivation could be plotted and determined. Initially, samples were desalted into buffers devoid of added reductant, however under these modestly anaerobic conditions argon or N\textsubscript{2} sparged samples would eventually begin to lose activity, presumably from air that entered the vial during the time samples were taken for activity assays. This effect was minimized by the addition of reductant to both samples. The reductant also slowed inactivation of air-sparged samples; extended incubations in air at 25 °C would often result in CbbT precipitation. In order to increase the rate of inactivation and maintain a soluble inactive form of CbbT, O\textsubscript{2} gas was often used instead of air. A time course experiment comparing O\textsubscript{2} vs. argon sparged samples in the absence or presence of 1 mM DTT can be seen in \textbf{Figure 3.3}. 
Figure 3.3. Time course comparing inactivation of CbbT in the presence of argon (●) vs. O₂ (○) sparged samples. In addition, samples either contained no DTT (dashed lines) or 1 mM DTT (solid lines). The samples were desalted into 20 mM MOPS-KOH, pH 7.0, 2 mM MgCl₂, and 0.1 mM ThDP and were flushed with gas for 10 min on ice; incubation and activity assays were performed at 25 °C. Data is plotted as % Initial Activity versus time (A) and log (% Initial Activity) versus time (B).
Injecting CbbT into vials that had been previously sparged with different concentrations of O₂ resulted in different rates of inactivation, with rate of inactivation dependent on the concentration of O₂ used (Figure 3.4).

**Figure 3.4.** Time course comparing the % initial activity of CbbT in the presence of argon (●), a mix of 50% argon and 50% O₂ (●), or 100% O₂ (○) vs. time. The samples were desalted to remove excess reductant, into 20 mM MOPS-KOH, pH 7.0, 2 mM MgCl₂, and 0.1 mM ThDP and then injected into vials that were flushed with the corresponding gas mixture for 10 min; incubation and activity assays were performed at 25 °C. Data is plotted as % Initial Activity versus time (A) and log (% Initial Activity) versus time (B).
Recovery of O$_2$ inactivated CbbT was attempted by sparging the inactive enzyme with argon for 20 min, followed by incubation at 25 °C for up to 24 h. No increase in activity was observed under these conditions. The addition of the donor substrate R5P (up to 100 µM) or increased concentrations of either cofactors (Mg$^{++}$ to 10 mM and ThDP to 1 mM) had no detectable protective effects. Also, no visible degradation of the inactive samples was observed by SDS-PAGE (Figure 3.5).

Figure 3.5. SDS PAGE analysis of CbbT samples incubated under argon or O$_2$. Lane 1, molecular weight markers. Lane 2, CbbT incubated under an argon head space. Lane 3, CbbT incubated under an O$_2$ head space. Incubations were performed in 50 mM MOPS-NaOH, pH 7.0, 2 mM MgCl$_2$, 0.1 mM ThDP, and 1 mM DTT for 24 h at 25°C. The O$_2$-stored sample displayed 10% the activity of the argon-stored sample.
3.3.3 SPECTRAL PROPERTIES OF INACTIVE TRANSKETOLASE

Several spectrometric techniques were used to aid in determining differences between the active form of CbbT and its oxygen-inactivated form. CD in the near UV range is particularly sensitive to changes in protein secondary structure. Intrinsic fluorescence is sensitive to changes in the chemical environment of the chromophore and is therefore sensitive to changes in tertiary structure. This is true especially if changes occur near a chromophore. Typical chromophores in proteins are Trp, Phe, and Tyr; however the quantum yield of Trp is much greater than the other two and therefore dominates the spectra when present. CbbT contains eight Trp residues per subunit.

A comparison of emission scans collected for active and inactive forms of CbbT (Figure 3.6) clearly indicated a 30% reduction in peak fluorescence for the inactivated CbbT, accompanied by a red shift of the peak emission from 329 nm to 332 nm. This shift in fluorescence is suggestive of Trp being exposed to a more polar environment upon inactivation by O$_2$. Though this in itself is not capable of identifying the site of modification it is diagnostic of the oxygen-mediated inactivation of CbbT. In addition, the decrease observed suggests that ThDP is still retained by CbbT since the loss of ThDP is associated with an increase in fluorescence without an accompanied shift in emission wavelength (see section 4.3.4).
Figure 3.6. Fluorescence emission spectra of CbbT incubated under a head space of either N\textsubscript{2} (—) or O\textsubscript{2} (---) at 25 °C. The concentration of both samples was 0.6 mg/ml in 20 mM HEPES-NaOH, pH 7.8, 2 mM MgSO\textsubscript{4}, 0.1 mM ThDP, and 0.1% BME. Emission spectra were collected following a 40 h incubation period at which point the O\textsubscript{2}- incubated sample retained 14% the activity of the N\textsubscript{2}-incubated sample. The excitation wavelength used was 280 nm.
The CD spectrum of a CbbT sample which had been inactivated upon exposure to air and had retained 30% activity of the N₂-sparged sample, showed a shift to more negative values in the region from 215 to 230 nm in comparison to the fully active N₂ sample (Figure 3.7). It can be seen that the shift is most pronounced around 222 nm, a wavelength often chosen to monitor secondary structure in proteins.

**Figure 3.7.** CD spectra of CbbT incubated under a headspace of either N₂ (solid line) or air (dashed line). Sample concentrations were 1 mg/ml in 20 mM HEPES-NaOH, pH 7.8, 2 mM MgSO₄, 0.1 mM ThDP, and 0.1% BME. One sample was stored under N₂ and the other under an air headspace and allowed to inactivate over 24 h at 25 °C at which point the sample exposed to air displayed 30% the activity of the nitrogen-stored sample.
3.3.4 FREE ENERGY OF UNFOLDING

Free energy of unfolding has been used to determine the conformational stability of different forms of proteins. We investigated the use of typical denaturants such as guanidinium hydrochloride (GdnCl) and urea, as well as thermal denaturing. Interestingly, it was found that concentrations of GdnCl, in the range from 0.5 to 1.5 M, would induce CbbT to precipitate out of solution. Upon investigating thermal denaturation curves, it was found that samples exposed to temperatures required to reach a constant upper slope, ~ 60 °C, would not refold in a reversible manner, though no precipitate was observed by visible spectroscopy. For these reasons, urea was chosen as the denaturant of choice to compare an active sample of CbbT with a sample that had largely been inactivated through exposure to O₂.

Samples were prepared identically to the time course experiments (Section 3.2.3), and the O₂ sample was allowed to reach 20% the activity of the N₂ control. The N₂ sample did not lose any activity over the time course. Concentrations of urea that would provide useful data points as well as determinations of the length of time required to reach equilibrium were determined empirically. Since a 4 h incubation at 25 °C was used, it was decided to sparge all the samples with N₂ gas to prevent further inactivation of either sample. Though the intensity of fluorescence can also be used to quantitate the unfolding of a protein, we found that the shift in peak emission wavelength produced cleaner data. Plots of peak maximal emission wavelength versus concentration of urea yielded differences in unfolding between the active and inactive samples. From this data the fraction of sample folded was determined and an unfolding curve was generated to calculate ∆G (H₂O) values (Figure 3.8). Both samples showed near complete unfolding.
in 5.4 M urea. The oxygen-inactivated sample appeared to be in a structurally more stable state and required a higher concentration of urea to reach a similar level of unfolding as the active N₂-stored sample. Assuming a two state mechanism of unfolding, the calculated free energy of unfolding, $\Delta G(H_2O)$, was 1.01 kcal mol⁻¹ and 1.63 kcal mol⁻¹ for the active and inactive forms of CbbT respectively. The dependence of denaturant concentration on free energy is described by the magnitude of m (102). The m values obtained for the N₂ and O₂ samples were 0.99 and 0.84 respectively.
Figure 3.8. Urea denaturing curves for active and inactivated CbbT. Closed circles (●) represent O₂-incubated inactive CbbT and open circles (○) represent N₂-incubated active CbbT. Solid lines represent the non-linear fit results as described in section 3.2.6.

3.4 CONCLUSIONS

We have shown that CbbT is sensitive to oxygen, becoming inactivated over time of exposure. Though this process can be slowed by the presence of reductant, it is apparently irreversible. Furthermore, it appears that this process was not exclusive to purified enzyme but O₂-mediated affects influenced the activity of enzyme in *E. coli* cells.
over-producing the recombinant protein as well. A direct relationship of the dependence of activity on the amount of oxygen present during induction was clearly seen after comparing 2 l fermentor cultures. The specific activity in a crude extract increased considerably as the amount of air introduced during induction decreased. Compounding the data analysis are natural physiological changes that occur within the cells, which could affect protein accumulation as well as cause modification of the product (CbbT) directly. It is known that the redox state of the media is sensed by E. coli cells (103) and affects regulation of the synthesis of various enzymes (104, 105). This leads to the possibility that a build up of reducing equivalents triggers the expression of one or more proteins that are responsible for the increased stability of transketolase at these low O₂ conditions. It remains to be determined whether the level of transcript was altered by the change in O₂ tension.

Spectral differences have been observed between these two forms of CbbT. Care was taken to prevent any artifacts in these spectra from CbbT precipitation which has been observed in extended incubations in O₂ at 25 °C. This was achieved by reducing the incubation time as well as collecting visible spectra of samples throughout the time course. Under these conditions producing a soluble completely inactive sample is not feasible; however we have shown that a 70 to 80% reduction in activity was adequate to yield different spectral properties. Most obvious was the change observed in fluorescence. This change in fluorescence can be used to detect the oxygen-inactivated form of CbbT, however purified protein is required. The CD band in the far UV region around 222 nm also displayed a change upon inactivation, with a slight decrease in
ellipticity upon inactivation. CD is sensitive to secondary structure in this region and an increase in ellipticity is indicative of an increase in $\alpha$-helix and/or $\beta$-sheet content.

Also indicative of structural changes was the observed difference in $\Delta G(H_2O)$ between the inactive and active forms of CbbT. With this $\Delta \Delta G(H_2O)$ of 0.6 kcal mol$^{-1}$ the inactivated CbbT appears to have adopted a more stable structure, a feature that may explain the lack of reversibility of this mechanism. It has been shown that there is a strong correlation between m values and how much of the protein surface is accessible upon unfolding (106). This suggests that the inactive form of CbbT undergoes slightly less total unfolding with m values of 0.99 and 0.84 for the active and inactive forms respectively.

It remains unclear whether this is an actual physiological event occurring within *R. sphaeroides* or rather a process that is characteristic to all *in vivo* situations. Various mechanisms of protein oxidation have been identified in various systems, all of which require ROS (107). This makes the observed inactivation of purified CbbT, which occurs *in vitro* in the apparent absence of known sources of ROS, a perplexing problem. Though the Mg(II) cofactor is believed incapable of generating ROS, it is possible that trace amounts of metals capable of the Fenton reaction (i.e. Fe(II) or Cu(I)) might be present that could presumably generate ROS. This type of oxidation would be preferential to residues surrounding metal binding sites of the protein (108).
CHAPTER 4

MUTAGENESIS OF CBBT RESIDUE C160

4.1 INTRODUCTION

Previously a sequence analysis of 21 different transketolase sequences had shown residue C159 of the yeast enzyme to be conserved, but not invariant, across all branches of life; further it was suggested that this reactive side chain may be responsible for oxygen-mediated inactivation of the *E. coli* enzyme (6). Residue C160 of *R. sphaeroides* CbbT corresponds to residue C159 of the yeast enzyme, which is known to be within the active site as determined by x-ray crystallography (7). This residue is deep within the ThDP binding pocket and is situated near the pyrophosphate moiety of the bound cofactor. Previous studies identified several invariant amino acids responsible for ThDP binding and activation (58, 82, 109). To investigate the role that C160, a semi-conserved active site residue, plays in both catalysis and cofactor binding, as well as to investigate its possible function in oxygen-mediated inactivation, C160 of CbbT was targeted for mutagenesis. Introduced residues were selected based on their occurrence at this site in different sources of transketolase. Aspartate or glutamate residues are found in some transketolases at position C160; a conserved substitution from cysteine to serine and removal of the side chain by introducing an alanine residue was also accomplished by site-directed mutagenesis. The results of both catalytic and fluorescence-binding studies
with wild-type and mutant proteins presented in this chapter clearly establish that Cys-160 plays an important role in correctly positioning ThDP at the active site. However, no apparent role of this residue in solely influencing oxygen inactivation was noted.

4.2 MATERIALS AND METHODS

4.2.1 MOLECULAR MODELING

For visualization of the ThDP binding pocket, a theoretical model of CbbT based on the crystal structures of yeast transketolase (pdb IDs: 1AYO, 1GPU, 1ITZ, 1NGS, 1TKA, 1TKB, 1TKC, and 1TRK) (13, 17-19, 57-59) was constructed, using the program Swiss-Model. Alignments and mutations were analyzed using the programs SPDBV and O (61-63).

4.2.2 MUTAGENESIS

Site directed mutagenesis was achieved using the Quikchange kit from Stratagene (La Jolla, CA). For mutagenesis the cbbT gene was subcloned into a pUC19 vector, resulting in a plasmid of 4.9 kbps (pUC19CbbT). A series of complementary primers were designed to introduce the desired mutation (Table 4.1). Reactions contained 50 ng template, 2.5 units of Pfu polymerase, 1 X Pfu buffer, 5% DMSO, and 125 ng of each primer. The recommended thermocycling conditions described in the manual were used. An aliquot of the PCR reaction (20%) was then subjected to digestion by DpnI for 2 h at 37 °C to cleave methylated template DNA. This was then transformed into E. coli JM109 by electroporation. To reduce the chances of introducing random mutations by the polymerase, a BamHI – NcoI fragment of 840 bp containing the site targeted for mutation
was subcloned from pUC19CbbT into pET11CbbT upon completion of the mutagenesis protocol. Mutations (*) in the resulting pETCbbT* vector were then confirmed by sequencing of this entire 840 bp region in both directions utilizing the sequencing primers designed for \textit{cbbT} (section 2.2.2).

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Nucleotide sequence of mutagenic primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>C160</td>
<td>ATCGCGGACGGCTGCTTGATGGAGGGGCATC</td>
</tr>
<tr>
<td>C160A</td>
<td>CGCGGACGGCTTCCTTGATGGAGGGGC</td>
</tr>
<tr>
<td>C160S</td>
<td>CGCGGACGGCTCCCTTGATGGAGGGGC</td>
</tr>
<tr>
<td>C160D</td>
<td>CGCGGACGGCTGCCGCTTGATGGAGGGGC</td>
</tr>
<tr>
<td>C160E</td>
<td>CGCGGACGGCTGCTGCTTGATGGAGGGGC</td>
</tr>
</tbody>
</table>

\textbf{Table 4.1.} Primers used for mutagenesis at residue C160. First row consists of the native sequence with the codon corresponding to residue C160 in red. The introduced base pair changes are underlined; all primers were accompanied by a matching complimentary 28-mer for the actual mutagenesis.

\textbf{4.2.3 OVER-EXPRESSION OF THE TRANSKETOLASE GENE AND PURIFICATION OF RECOMBINANT PROTEIN}

Gene over-expression and recombinant protein purification procedure was identical to that described for purification of wild-type CbbT (section 2.2.5). In short, all mutated genes were over-expressed under micro aerobic conditions in \textit{E. coli} strain GJ1158 (51) by inducing with 0.4 M NaCl (89) at an \textit{OD}_{600} of 0.8 overnight at 25 °C. Harvested cells were washed two times with TEM (20 mM Tris-HCl, pH 7.0, 0.1% BME,
1 mM EDTA) buffer; the pellet was stored at -80 °C until further use. Cells thawed on ice were resuspended in 20 ml buffer A (20 mM MOPS-NaOH, pH 6.9, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME) and were lysed using a French press. This was followed by a low speed centrifugation (10,000g, 20 min), and then precipitation of nucleic acids with 0.1% streptomycin sulfate, incubated on ice for 20 min; after ultracentrifugation (100,000 g; 1 h), followed by filtration through a 0.45 micron filter, the extract was loaded onto a 2.6 x 14 cm DEAE fast-flow sepharose (DSFF) column and eluted using an increasing linear gradient of 2 – 15% buffer B (buffer A with 1 M ammonium sulfate). Active fractions, eluting between 85 and 130 mM ammonium sulfate, were pooled, and then treated with solid ammonium sulfate to a final concentration of 1 M. Following a 1 h equilibration on ice, the solution was filtered and then loaded onto a 1.6 x 90 cm phenyl Sepharose high performance (PSHP) column. Elution of most of the proteins was achieved using a decreasing linear gradient from 100% to 30% of buffer B. However the C160A, C160E, and C160 S mutant proteins were only partially eluted under these conditions. Complete elution of these three proteins required the use of a no-salt high pH buffer (20 mM glycylglycine, pH 9.0, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME). The PSHP fractions were loaded onto a 1.6 x 11 cm Q-sepharose HP (QSHP) column and active enzyme eluted using a linear gradient ranging from 0 – 20% buffer B. Active fractions were concentrated (Millipore spin concentrators, NMW 30 kDa) and loaded onto a 1.6 x 50 cm Superpose 200 gel filtration column and eluted with 5% buffer B. Glycerol was added to 20%, the aliquots were flash frozen in liquid N₂ and stored at -80 °C until needed; no noticeable loss in activity was observed after storage.
4.2.4 QUANTITATION OF THDP

For ThDP solutions, concentrations were determined using an $\varepsilon_{267}$ of 8520 M$^{-1}$ cm$^{-1}$ (110). For non-pure ThDP solutions, concentrations were quantitated by the fluorometric detection of ThDP oxidized to thiochrome in an alkaline medium with ferricyanide (111). In this method, the sample containing ThDP is first incubated in 50% EtOH for two min at room temperature. This is then oxidized by the addition of NaOH to 2% and ferricyanide to 0.02%; after two min of additional mixing, H$_2$O$_2$ to 0.1% is added. The fluorescence of the samples is then measured using an excitation wavelength of 365 nm and reading the emission at 450 nm. The concentration of ThDP was then calculated from a standard curve generated using a stock ThDP solution of known concentration.

4.2.5 FORMATION OF APOENZYME

Apoenzymes were prepared by two methods. The first method involved dialysis of the holoenzyme against 20 mM glycylglycine-HCl, pH 9.0, in the presence of 1 M ammonium sulfate and 10 mM BME at 4 °C. The first of four buffer exchanges also contained 1 mM EDTA. A second method was employed to save time; this consisted of apoenzyme formation while the protein was bound to a small 1 ml QSHP anion exchange column. The protein was bound to the column using pH 9.0 glycylglycine-HCl buffer containing 10 mM BME; this column was then washed with 30 ml of the same buffer containing 1 mM EDTA, followed by 30 ml of buffer without EDTA. The apoenzyme eluted after the application of a sharp gradient from 0 to 1 M ammonium sulfate. To remove the majority of the ammonium sulfate, the resulting concentrated apoenzyme was
then passed through a 5 ml Hi-trap G-25 desalting column (Amersham Biosciences, San Francisco, CA) into a buffer of 20 mM HEPES-NaOH, pH 7.8, containing 0.1% BME. If the sample was not used immediately, glycerol was added to 20% and the protein stored at -80 °C.

4.2.6 BIOCHEMICAL CHARACTERIZATION OF MUTANT PROTEINS

The enzymatic assay was based on the NAD\(^+\) dependent glyceraldehyde-3-phosphate dehydrogenase coupled system developed by Kochetov (56). Protein concentration was determined using the Bio-Rad protein microassay procedure, based on the method of Bradford (112). All kinetic data were collected at 25 °C using a Varian Cary 100 spectrophotometer with temperature controller. Dependence of pH on activity was measured using 3 different buffers at approximate 0.2 pH unit intervals; the buffers used were MOPS, HEPES, and glycylglycine. The exact pH of the reaction was measured immediately following the assay; all buffers were used at a final concentration of 200 mM. Substrate kinetics were determined in the presence of 0.1 mM ThDP and 5 mM MgSO\(_4\). Xylulose 5-P was used as the donor substrate and ribose 5-P as the acceptor. SigmaPlot v8.0 was used for data analysis.
4.2.7 FLUOROMETRIC THDP AFFINITY MEASUREMENTS

A Jobin Yvon – Spex model Fluoromax-3 spectrofluorometer was used for fluorometric data collection. The automated sample changer and Neslab RTE-111 temperature controller accessories were used with Starna 10 mm quartz fluorometry cuvettes with airtight septa screw caps. The inner filter effect was minimized by using a higher wavelength for excitation, where ThDP absorbs less (295 nm as opposed to 280 nm). The wavelength chosen for emission was 329 nm. In addition, the following correction factor was applied to the fluorescence values

\[ F_{c,i} = F_{o,i} \times 10^{-0.5(A_x + A_{m,i})} \]  \hspace{1cm} (113)

where \( F_{c,i} \) is the corrected fluorescence intensity at a particular wavelength \( i \), \( F_{o,i} \) is the observed fluorescence, \( A_x \) is the absorbance of ThDP at the excitation wavelength, and \( A_{m,i} \) is the absorbance of ThDP at the emission wavelength. This correction factor reportedly works well for small inner filter absorbances less than 2 OD (114). In addition to the inner filter effect, it was noted that data followed the quadratic binding equation much closer if the sample was allowed to equilibrate in the cuvette. For this reason sequential additions of ThDP were made to the cuvette, allowing for equilibration to occur after each addition. The amount of time required to reach equilibrium varied based on the amount of ThDP added, taking up to \( \sim 3 \) h for lower concentrations of ThDP and \( < 1 \) h for higher concentrations. The standard assay contained 2 mM MgSO\(_4\), 0.1% BME, 100 mM buffer, and 800 nM of enzyme (calculated by monomeric molecular weight). The equilibrium observed fluorometrically was confirmed enzymatically by removing small aliquots at each ThDP concentration for activity assays. For all binding studies
ThDP concentrations were determined using an $\varepsilon_{267}$ of 8520 M$^{-1}$ cm$^{-1}$ (110). SigmaPlot v8.0 was used for data analysis.

4.2.8 THDP AFFINITY DETERMINED ENZYMATICALLY

Plots of ThDP binding for CbbT and all C160 mutants were obtained in the following manner. Apoenzyme was prepared as discussed in section 4.2.5. Reconstitution was performed in 100 mM buffer (MOPS-NaOH pH 6.9, HEPES-NaOH pH 7.8, or glycylglycine-HCl pH 8.5), 5 mM MgSO$_4$, and 0.1% BME. Samples contained an empirically determined fixed amount of enzyme (CbbT – 0.04 mg/ml, C160D – 0.05 mg/ml, C160S – 0.2 mg/ml, C160E – 0.35 mg/ml, C160A – 0.5 mg/ml) and a concentration of ThDP varying up to a saturating amount (CbbT and C160D: 0 – 30 nM, C160S: 0 – 6 mM, C160E: 0 – 12 mM, C160A: 0 – 30 mM). These were incubated in a 25 °C water bath up to 4 h. Since the time of reconstitution was dependent on ThDP concentration and varied for each enzyme, the activity of each sample was monitored periodically over time. Assays continued until the activity of a sample reached a stable plateau, this activity was then used to represent the amount of holoenzyme formed. Data was plotted as specific activity (U/mg) versus ThDP concentration and analyzed by non-linear regression using Sigma Plot v8.0. The C160A mutant was additionally analyzed by constructing a Hill plot.

4.2.9 KINETIC BINDING OF THDP

Fluorescence detection of holoenzyme formation was utilized to obtain a first order rate constant of ThDP binding to apo C160D and apo CbbT enzymes.
Fluorescence instrument settings were identical to those for the equilibrium binding with the exception that excitation was at 280 nm; all measurements were obtained at 25 °C. Each sample contained 100 mM buffer, 0.1% BME, 2 mM MgSO4, and 800 nM of apoenzyme. Reconstitution was initiated by the addition of 5 µl of a 10 mM ThDP solution for a final concentration of 50 µM ThDP. The cuvette was quickly mixed by inversion three times and data collection begun immediately; the estimated dead time for this experiment was 8 sec. Data collection continued until a leveling off of fluorescence was achieved, ranging from 6 to 45 min depending on the pH. Data was exported into SigmaPlot v8.0 for data analysis. Fluorescence quenching data were fit to the single-exponential equation: \( Y = a_1 + a_2 \times \exp(a_3 \times X) + a_4 \times X \); with \( a_i \) representing fit parameters. For comparison, data was also fit to the double-exponential equation: \( Y = a_1 + a_2 \times \exp(a_3 \times X) + a_4 \times \exp(a_5 \times X) \).

4.2.10 CIRCULAR DICHROISM

Near UV CD measurements were taken using an Aviv model 62A DS spectrometer, scanning from 260 to 360 nm at 1 nm intervals with a bandwidth of 1 nm and a 2 sec signaling average; five scans were averaged for each sample. Data were collected using 1 cm quartz cuvettes with temperature controller set to 25 °C. All protein concentrations were in the range of 0.50 – 0.80 mg/ml in 20 mM HEPES-NaOH buffer, pH 7.8, containing 2 mM MgSO4, and 0.1% BME. ThDP was added to 25 µM final concentration and samples were allowed to equilibrate at 25 °C for 1 to 3 h. Data was smoothed by applying a negative exponential 10 degree polynomial regression with SigmaPlot v8.0.
4.3 RESULTS AND DISCUSSION

4.3.1 BIOINFORMATICS ANALYSIS

A previous detailed sequence analysis was performed on 21 transketolase sequences and was used to identify various conserved and semi-conserved residues (6). Based on these sequences, C159 (or C160 in CbbT) was labeled as a semi-conserved residue. To supplement these findings with newly available sequences, the alignment of 85 different complete transketolase and transketolase-related sequences homologous to CbbT was performed. Results indicated that 48 sequences (or 56%) had a preference for Cys at the position equivalent to residue 159 of the Saccharomyces enzyme; Asp was found in 22 different sequences (or 26%), Glu in 12 different sequences (or 14%), and finally Ala was shown to be present in only 3 of the 85 enzymes compared (or 4%). All three sequences containing Ala at this position were actually putative 1-deoxy-D-xylulose 5-phosphate synthases (DXPS), proteins that are not classically transketolase enzymes but catalyze a transketolase-type reaction. Not all DXPS sequences have Ala at this position, though it appears to be preferred for that enzyme. Mammalian transketolase sequences thus far all appear to contain Glu at this position, and this is also true for the two archaeal transketolases considered in this alignment. Thus far only the E. coli transketolase, and now R. sphaeroides CbbT, have been shown to be subject to oxygen-mediated inactivation; therefore there is no indication that enzymes from sources that naturally contain alternate residues at position 159 (or the equivalent residue) are less sensitive to oxygen.
4.3.2 MUTAGENESIS AND ENZYME PURIFICATION

Site directed mutagenesis was first attempted using the Chameleon kit (Stratagene, La Jolla, CA), which applies a modification of the unique-site elimination method. In this method two mutagenic primers are used, one to alter a restriction site within the antibiotic resistance gene of the plasmid, and the other to introduce the desired mutation. With both primers annealing to the same strand of the plasmid, T7 polymerase is then used to replicate the entire strand. This is then transformed into a repair-deficient host, the subsequent plasmid DNA can then be selectively digested to remove the non-mutated plasmid. This method was applied first to produce the mutant C160S, but resulted in an extremely low number of colonies and despite the selective mechanism, the mutation efficiency was low (10%, as based on 10 sequences). These results were attributed to the high GC content found in \textit{cbbT}, which may have contained regions of secondary structure at 37 °C, making it difficult for the polymerase to replicate. This temperature requirement was set by the presence T7 polymerase and also prevented repeated rounds of denaturation and annealing by thermocycling. A much higher rate of success was found using the Quikchange method of Stratagene, which was used to generate the remaining mutants (\textbf{Figure 4.1}).
Figure 4.1. Molecular models of the C160 mutants of *R. sphaeroides* CbbT. The region surrounding the residue of interest is displayed, showing the location in the active site near the cofactor ThDP. Typical pKa values for the corresponding amino acids are given. The model of CbbT was constructed as described in the text (section 4.2.1). The site-directed mutants were created using the SPDB viewer, the rotomer with the lowest energy was selected and no subsequent energy minimization calculations were performed. The ThDP was taken from the yeast transketolase (pdb 1GPU (18)) and positioned following an improved fit of the yeast enzyme with the theoretical CbbT model.
Figure 4.1

- Cys (wild type): pKₐ ~ 8.3
- Ser: pKₐ ~ 13
- Ala
- Asp: pKₐ ~ 3.9
- Glu: pKₐ ~ 4.3

111
In general, mutant CbbT proteins behaved similar to wild type CbbT throughout the purification with the exception of the PSHP column. It was found that the mutant proteins C160A, C160S, and C160E all displayed a much higher affinity for this hydrophobic matrix, with only a small fraction eluting under the no-salt wash. For these mutant proteins, it was necessary to elute the enzyme from the column with a pH 9.0 buffer solution. All mutant proteins were purified to homogeneity, as shown by SDS-PAGE, with the exception of C160E, which despite the various forms of chromatography, still contained several minor bands (Figure 4.2).
Figure 4.2. SDS/PAGE analysis of C160E through different stages of purification.

Lanes: MW, Purified mutant C160D was used as a molecular weight marker, all subsequent lanes refer to C160E in chronological order of purification; 1, crude extract; 2, DSFF anion exchange; 3, PSHP hydrophobic interaction; 4, QSHP anion exchange; 5, S200 gel filtration. Approximately 8 µg of protein were loaded per lane, the gel was Coomassie stained (R-250).

4.3.3 BIOCHEMICAL PROPERTIES OF THE MUTANT PROTEINS

With the exception of the C160A enzyme, which retained only 7 percent of the activity of the wild-type enzyme, all mutant CbbT proteins possessed reasonable levels of enzymatic activity (from 26 to 81 percent of the wild-type activity). All active mutant proteins maintained $K_m$ values for the substrates R5P and Xu5P similar to that seen for the wild-type enzyme, with the exception of C160E, which displayed $K_{Xu5P}$ values approximately 6-fold higher than wild-type CbbT (Table 4.2). With the surprising
exception of C160D, all active mutants displayed maximal enzymatic activity around pH 7.8; whereas C160D preferred a more basic pH value of 8.5 (Figure 4.3).

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$V_{\text{max}}$ (U/mg)</th>
<th>$K_{R5P}$ (µM)</th>
<th>$K_{Xa5P}$ (µM)</th>
<th>$% V_{\text{max}}$ of wT</th>
<th>Optimal pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbbT(WT)</td>
<td>37.0</td>
<td>949</td>
<td>11</td>
<td>100</td>
<td>7.8</td>
</tr>
<tr>
<td>C160D</td>
<td>30.1</td>
<td>837</td>
<td>20</td>
<td>81</td>
<td>8.5</td>
</tr>
<tr>
<td>C160S</td>
<td>21.8</td>
<td>436</td>
<td>33</td>
<td>59</td>
<td>7.9</td>
</tr>
<tr>
<td>C160E</td>
<td>9.5</td>
<td>739</td>
<td>65</td>
<td>26</td>
<td>7.6</td>
</tr>
<tr>
<td>C160A</td>
<td>2.7</td>
<td>882</td>
<td>15</td>
<td>7</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Table 4.2. Kinetic properties of CbbT and C160 mutant proteins. All values are based on an average of at least three independent experiments. The standard assay consisted of 100 mM buffer, 2 mM MgSO$_4$ and 0.1% BME, under optimal ThDP concentrations and pH for each enzyme. For determining the optimal pH, 200 mM HEPES-NaOH or MOPS-NaOH buffer was used.
Figure 4.3. Plot comparing activity versus pH for CbbT and three C160 mutant proteins. Activity is expressed as a percent maximum, the pH was determined immediately following the assay. The standard assay consisted of 200 mM HEPES-NaOH, MOPS-NaOH, or glycyglycine-HCl buffer with 2 mM MgSO₄, 0.1% BME, and optimal ThDP concentrations for each enzyme.
4.3.4 THDP AFFINITY OF MUTANT PROTEINS

While it was clear that the C160A enzyme had drastically reduced enzymatic activity, the potential importance of Cys160 in binding ThDP was assessed after determining the amount of ThDP that remained bound after passing the wild-type and mutant proteins through a G-25 desalting column. ThDP concentrations were quantitated by the fluorometric detection of ThDP oxidized to thiochrome in an alkaline medium with ferricyanide (111). Initial estimates were that the C160A and C160S enzymes retained 2% and 40%, respectively, of the amount of ThDP associated with the wild-type protein.

In order to accurately quantify differences in ThDP affinity ($K_D$ and $K_m$), both fluorescent and enzymatic methods were employed. The potential use of fluorometric methods to measure cofactor binding to transketolase was previously considered (115), however no other $K_D$ values were found in the literature. Fluorometry has previously been employed to measure ThDP binding in pyruvate decarboxylase (116) and the pyruvate dehydrogenase complex (117). Adapting these methods for CbbT was found to be convenient and data sets for determining $K_D$ values for CbbT and the C160D mutant at their respective optimal pH were obtained (Figure 4.4). The three other mutant proteins, C160S, C160A, and C160E, all required too much ThDP, and technical considerations precluded the ability to make an accurate inner filter correction (described in section 4.2.7). Thus, data was collected for CbbT and the C160D mutant enzymes only.

Initially, C160D protein data was collected at pH 7.8 (yielding a $K_D$ value of $2.50 \pm 0.42 \mu M$), however it was later realized that the optimal pH of this protein was quite different from the wild type enzyme (pH 8.5 and 7.8, respectively). For this reason, $K_D$ values for
CbbT and the C160D mutant enzymes were determined at three different pHs (Table 4.3), above and below the optimum of the wild-type protein. Clearly, both proteins exhibited similar $K_D$’s at their respective optimum pH ($K_D$ of 3.4 µM for CbbT and 3.6 µM for C160D).

**Figure 4.4.** Determination of the $K_D$ for ThDP. Fluorescence data used to calculate the $K_D$ of CbbT (●) and the C160D (○) enzymes at their respective optimal pH values. Data was expressed as the difference in fluorescence observed divided by the initial fluorescence ($\Delta F/F_i$) vs. the concentration of ThDP.
Table 4.3. CbbT and C160D fluorometrically determined ThDP dissociation constants. Numbers are expressed in µM ThDP. Fluorescence was excited at 295 nm while the disappearance of fluorescence emission was monitored at 329 nm under variable ThDP concentrations (0 – 16 mM at pH 6.7, 0 – 50 mM at pH 7.8, 0 – 117 mM at pH 8.5). The assays contained 2 mM MgSO₄, 0.1% BME, and 200 mM of either MOPS-NaOH (pH 6.7), HEPES-NaOH (pH 7.8), or glycylglycine (pH 8.5).

<table>
<thead>
<tr>
<th></th>
<th>pH 6.7</th>
<th>pH 7.8</th>
<th>pH 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbbT</td>
<td>2.28 ± .22</td>
<td>3.41 ± .21</td>
<td>37.8 ± .39</td>
</tr>
<tr>
<td>C160D</td>
<td>1.10 ± .10</td>
<td>2.50 ± .42</td>
<td>3.57 ± .47</td>
</tr>
</tbody>
</table>

In order to reinforce the K_D results, as well as to consider how the other C160 mutants interacted with cofactor, relative affinities for ThDP were also determined enzymatically. Substrate concentration vs. activity plots for CbbT and C160D were generated (Figure 4.5). Plots were also generated for the other mutant proteins and K_m values determined at pH 6.7, 7.8, and 8.5 (Table 4.4). For the C160E enzyme, the K_m could not be determined at pH 8.5 because this protein showed no activity at this pH. However, the C160S enzyme had an approximate 700-fold increase in K_m compared to both C160D and CbbT at pH 7.8. In agreement with the dissociation data, C160D displayed a lower apparent affinity for ThDP than CbbT at all pH values examined. The
C160E enzyme possessed an affinity for ThDP even weaker than the C160S mutant with a value approximately 1,400 to 1,700 times higher than CbbT and C160D, respectively. This $K_m$ value did decrease slightly when the pH was lowered from 7.8 to 6.7, as did the value for the wild-type and C160D enzymes. Using a large amount of C160A enzyme allowed $K_{\text{ThDP}}$ data collection. Interestingly, this mutant enzyme displayed positive cooperativity upon binding ThDP with a Hill constant of 1.9 in the range from 10% to 85% saturation (Figure 4.6). This was unlike all other mutant and wild-type proteins, which showed no cooperativity upon ThDP binding. The $K_{\text{ThDP}}$ value for C160A obtained from the Hill plot showed a tremendous decrease in affinity with a $K_m$ value nearly 8,000 fold higher than the wild-type enzyme at pH 7.8 (Table 4.4).
**Figure 4.5.** Determination of the $K_m$ for ThDP. Data used in the determination of $K_m$ ThDP for CbbT (●) and the C160D (○) enzymes at their respective optimal pH values. Data were expressed as U/mg vs. the concentration (µM) ThDP. C160D data continued to 1 mM ThDP and displayed a clear saturation, but values above 100 µM ThDP were omitted from this graph.
Table 4.4. Enzymatically determined affinity for ThDP of CbbT and C160 mutant proteins. Values are expressed in µM ThDP. The assays contained 2 mM MgSO$_4$, 0.1% BME, and 200 mM of either MOPS-NaOH (pH 6.7), HEPES-NaOH (pH 7.8), or glycylglycine-HCl (pH 8.5).

<table>
<thead>
<tr>
<th></th>
<th>pH 6.7</th>
<th>pH 7.8</th>
<th>pH 8.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CbbT</td>
<td>1.06 ± 0.05</td>
<td>1.77 ± 0.13</td>
<td>85.9 ± 10.2</td>
</tr>
<tr>
<td>C160D</td>
<td>0.50 ± 0.04</td>
<td>1.37 ± 0.07</td>
<td>27.8 ± 2.5</td>
</tr>
<tr>
<td>C160S</td>
<td>ND$^a$</td>
<td>1,130 ± 180</td>
<td>ND</td>
</tr>
<tr>
<td>C160E</td>
<td>1,280 ± 40</td>
<td>2,430 ± 80</td>
<td>NA$^b$</td>
</tr>
<tr>
<td>C160A</td>
<td>ND</td>
<td>13,500 ± 300</td>
<td>ND</td>
</tr>
</tbody>
</table>

$^a$ ND, not determined

$^b$ NA, no activity
**Figure 4.6.** Hill Plot of C160A reconstitution data within the range from 10 to 85% saturation using enzymatic activity data at pH 7.8. Linear regression of this data resulted in a line with slope 1.90.
To determine whether the observed differences of the C160D mutation were due to an alteration in binding mechanism, kinetic binding studies were performed. Since the reaction time of ThDP binding appeared to be on the order of min, an accurate kinetic rate of formation could be calculated by monitoring the quenching of intrinsic fluorescence upon addition of a saturating amount of ThDP to the apo enzyme. All data could be reasonably fit to a first order exponential equation (Figure 4.7 A, B). However, the data collected at pH 6.8 for both enzymes better fit a double exponential function (Figure 4.7 C, D), suggesting a pH-induced change in the binding mechanism. This change could be the result of either a newly introduced step in binding or a change in the reaction rate of a preexisting binding step. The latter would suggest that ThDP binding to the apo enzyme is a multi-step process and perhaps the result of one active site binding before the other to form the active dimer. A comparison of these first order rate constants showed that there was no significant difference between the C160D mutant and wild type (Table 4.5), suggesting that the mechanism of ThDP binding has not been altered in this mutant.
Figure 4.7. Upper plots showing typical fluorescence quenching during the reconstitution of CbbT (A) and C160D (B) at pH 7.8. Lower plots compare CbbT reconstitution under pH 7.8 (C) and pH 6.7 (D) conditions. For (C) and (D) the time scale has been plotted logarithmically to emphasize the early phase of the reaction. Dots (······) represent raw data; Thick dashed line (▬▬▬) represents the single exponential fit to the equation $Y = a_1 + a_2 \times \exp(a_3 \times X) + a_4 \times X$; Thick dotted line (····) represents the double exponential fit to the equation: $Y = a_1 + a_2 \times \exp(a_3 \times X) + a_4 \times \exp(a_5 \times X)$. Each sample contained 100 mM buffer, 0.1% BME, 2 mM MgSO$_4$, and 800 nM of apoenzyme. Reconstitution was initiated by the addition of 5 µl of a 10 mM ThDP solution to yield a final concentration of 50 µM ThDP. Fluorescence excitation was at 280 nm and emission was monitored at 329 nm; the estimated dead time for this experiment was 8 seconds.
Figure 4.7
<table>
<thead>
<tr>
<th>pH</th>
<th>CbbT $k_{obs}$</th>
<th>CbbT $t_{1/2}$</th>
<th>C160D $k_{obs}$</th>
<th>C160D $t_{1/2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.7</td>
<td>0.0029 ± 0.0008</td>
<td>464 ± 31</td>
<td>0.0040 ± 0.0017</td>
<td>359 ± 11</td>
</tr>
<tr>
<td>7.8</td>
<td>0.0160 ± 0.0025</td>
<td>44 ± 7</td>
<td>0.0198 ± 0.0030</td>
<td>36 ± 5</td>
</tr>
<tr>
<td>8.6</td>
<td>0.0015 ± 0.0001</td>
<td>250 ± 73</td>
<td>0.0019 ± 0.0001</td>
<td>193 ± 66</td>
</tr>
</tbody>
</table>

**Table 4.5.** Fluorometrically determined binding rates for ThDP of CbbT and C160D. Values for observed rate constants ($k_{obs}$) are expressed in sec$^{-1}$, values for half life ($t_{1/2}$) are expressed in seconds. Decay in fluorescence was measured by exciting at 280 nm and monitoring emission at 329 nm upon the addition of ThDP to 50 µM. The assays contained 2 mM MgSO$_4$, 0.1% BME, and 100 mM of either MOPS-NaOH (pH 6.7), HEPES-NaOH (pH 7.8), or glycylglycine-HCl (pH 8.5).

Finally, CD measurements, through the appearance of a new band (range 300 – 380 nm), may also be used to assess the ability of ThDP to bind to transketolase (118). Although the assignment of this band remains unclear, it was most recently proposed that the band results from the conversion of a 4'-amino to a 1',4' – imino tautomer of the ThDP pyrimidine (119). This would suggest that 1',4' – imino pyrimidine is the preferred tautomer in the hydrophobic pocket of the active site in transketolase. However, the E1 subunit of *E. coli* pyruvate dehydrogenase does not display a new CD band in this region upon ThDP addition (120). In this subunit the imino tautomer,
formed by the addition of the stable reaction intermediate analog phosphonolactylthiamin diphosphate, resulted in a new CD band centered at 305 nm as observed by difference spectra (121). In the CD spectra of CbbT, as with other transketolases, an inversion of ellipticity was observed around 330 nm upon ThDP binding. Qualitative examination of CbbT and the C160 mutant proteins indicated that inversion was observed only for CbbT and C160D in the presence of 25 µM ThDP at pH 7.8, in agreement with the above kinetic data (Figure 4.8).
**Figure 4.8.** Circular dichroism spectra of apo CbbT and the various apo C160 mutant proteins, comparing the extent of holoenzyme formation in the presence of 25 µM ThDP and 2 mM MgSO₄ at pH 7.8 for 1 – 3 h at 25 °C, as indicated by the inversion observed at 330 nm. Closed symbols represent the apoform and open symbols represent the addition of ThDP, C160E (○); C160A (▽); C160D (□); CbbT (◊); C160S (△).
4.3.5 OXYGEN SENSITIVITY OF MUTANT PROTEINS

The oxygen sensitivity of the C160D and C160S mutants was examined in comparison to wild-type enzyme. The C160D enzyme behaved similarly to wild-type CbbT in the presence of oxygen, displaying a similar rate of inhibition (Figure 4.9). One notable difference was the observed increase in activity of the C160D mutant in both the N₂ sparged sample as well as the O₂ sparged sample prior to inactivation. This would suggest that the purified mutant was not fully reconstituted upon purification. As can be seen, the C160D mutant under the air purge also displayed an increase in activity; this increase was short lived due to oxygen inactivation. The fact that this effect was not observed in the wild type sample suggests that following purification protocol (at 4 °C) the wild type enzyme appears to be closer to completely reconstituted whereas the C160D mutant is not. As all our calculations of ThDP affinity were performed at 25 °C we do not have a quantitative measurement of the effect of temperature on this value. In addition, such a slow rate of reconstitution observed by the C160D argon sample was not observed for the affinity measurements of this mutant (section 4.3.4). The affinity measurements were conducted in the presence of a low amount of ammonium sulfate, which could possibly increase the rate of reconstitution.

Similar inactivation studies with the C160S mutant enzyme were attempted; however this mutant protein proved to be too unstable for incubation at 25 °C under these conditions and both aerobic and anaerobic samples quickly lost activity. The oxygen sensitivity of the C160A and C160E mutants was observed during data collection for ThDP affinity as measured by enzyme activity. Apoenzyme incubated with variable concentrations of ThDP was assayed over time to determine the reaction velocity at
equilibrium for each ThDP concentration. Each of these two mutant enzymes displayed a short lived plateau of maximal activity (~1hr) after which activity began decreasing. This decrease was believed indicative of oxygen sensitivity and/or general instability of these two mutant enzymes.
Figure 4.9. Time-dependent inactivation of CbbT in an air (□) and argon (■) atmosphere compared with the C160D enzyme incubated in an air (▲) and argon (▲) atmosphere. Enzyme samples were desalted into 20 mM HEPES pH 7.8, 2 mM MgSO4, 0.1 mM ThDP, and 0.5 mM DTT, sparged with the gases on ice and then incubated at 25 °C. Though the same incubation buffer was used for both enzymes, each enzyme was assayed at its own optimal pH, 7.8 (HEPES-NaOH) for wildtype CbbT, and 8.5 (glycylglycine-HCl) for C160D. The activity is expressed in µmol min⁻¹ mg⁻¹ for both samples with CbbT plotted on the left-hand y-axis and C160D plotted on the right-hand y-axis.
4.4 SUMMARY

Transketolase catalyzes an important step in the CBB reductive pentose pathway, and functions with several other enzymes to facilitate the regeneration of the acceptor molecule for CO₂ fixation by this pathway. Despite its function in catalyzing important interconversions of sugar molecules in organisms that reductively assimilate CO₂, there have not been detailed structure-function studies of transketolase proteins that participate in the CBB pathway. The close homology of all transketolases allows the available yeast structural models (13, 17-19, 57-59) to be used to understand aspects of the structure and function of this enzyme from different sources. From these X-ray crystallography models, Cys160 presumably interacts with the first phosphate group of the ThDP cofactor via H-bonding through a water molecule. Indeed, this interaction can be visualized by looking at the electron density of the solved structure (Figure 4.10). The b-factor of the H₂O molecule (HOH 405 in pdb 1GPU, b = 18.39) allows for its confident placement. Many known interactions determined from the crystal structure have been explored and tested by other methods, such as site-directed mutagenesis. Due to its location in the active site as well as the potential of cysteine to act as a reactive side chain, we have prepared four mutant proteins with alterations at position C160 in CbbT of R. sphaeroides.
Figure 4.10. Region of interest showing stick structures of residue C160 and part of the ThDP cofactor overlaid with the electron density map. The electron density for pdb 1GPU (18), downloaded from the Uppsala Electron Density Server displayed the positioning of the water molecule responsible for bridging the side chain of residue C160 (equivalent to C159 of the *S. cerevisiae* enzyme) and the α-phosphate of ThDP.

Each mutant protein was produced in large quantities in *E. coli* strain GJ1158, which provided a convenient host to obtain significant amounts of purified recombinant
protein without having to resort to making tagged fusion proteins. The physicochemical properties of the mutant proteins clearly indicated that Cys160 plays a critical role in ThDP binding. The ability of serine to form a slightly stronger but shorter H-bond than cysteine typically makes this a conservative mutation. However the drastic effect this substitution had on the ability of the protein to bind ThDP suggested that there is not much flexibility in shortening the hydrogen-donating side chain. The C160A mutation further supported this, causing a severe decrease in ThDP affinity as well as an altered binding mechanism that displayed positive cooperativity. Aspartate proved to be a reasonably conservative substitution enzymatically, with the increased pH optimum of the C160D enzyme the most notable change. In addition to maintaining about 75% of the activity of wild type CbbT, the C160D protein displayed an equal if not better affinity for ThDP at all pH values examined. Measuring both substrate $K_m$ values of the C160D mutant at its preferred pH of 8.5 resulted in a catalytic efficiency similar to that of wild type. With its low pK_a, it may be assumed that aspartate acts as an H-bond acceptor rather than donor, still interacting through the water molecule to bind ThDP. Further, with kinetic binding rates similar to wild type, the changes observed in the aspartate mutant do not appear to be caused by an altered ThDP binding mechanism. Interestingly a deviation from first order binding was observed at reduced pH, possibly caused by a non-equivalence in cofactor binding of the two chemically equivalent active sites (122). Non equivalent active sites during catalysis are a feature believed to be common among all ThDP dependent enzymes (123). In contrast to aspartate, the glutamate-substituted enzyme displayed a weak affinity for ThDP, as well as a substantial increase in substrate $K_m$ values. Since glutamate naturally occurs in this position in mammalian and archaeal
transketolases, a different structure in this region of the ThDP binding pocket may be suggested for these two groups of transketolase enzymes. Compared to aspartate, lengthening of the side chain by one C-C bond appeared to place the side chain out of range for any beneficial interaction with the water molecule, perhaps even sterically interfering with ThDP binding (Figure 4.10). It is worth noting that the observed increase in affinity for the PSHP matrix during purification of the C160A, C160S, and C160E mutants could be the result of an exposed ThDP binding pocket. This pocket has been shown to be lined with hydrophobic residues which stabilize the pyrimidine of ThDP (16), making it possible that the three mutant proteins with hindered ThDP binding contained an extra exposed region of hydrophobicity, unlike CbbT and C160D.

Although Cys160 did not appear to be solely involved with the interesting and perplexing propensity of CbbT to be inactivated in the presence of oxygen, perhaps this residue, in combination with others, might be important for this purpose. Certainly, the importance of Cys160 in cofactor binding has been demonstrated by these studies. Though not directly involved in catalysis, this loosely conserved residue plays a major role in the binding and proper positioning of the cofactor ThDP. In order to maintain the constrained v-shaped structure required for activity, ThDP must be securely bound within its pocket through many interactions between the protein and the pyrophosphate and the protein and the pyrimidine, including the interaction of the $\alpha$-phosphate with C160 bridged through a solvent water molecule.
CHAPTER 5

ANALYSIS OF OXYGEN-MEDIATED CBBT INACTIVATION BY HYDROGEN/DEUTERIUM EXCHANGE

5.1 INTRODUCTION

Many hydrogen atoms bound to a protein are capable of readily exchanging with water hydrogen atoms in solution. Of the hydrogen atoms in a protein, S, O, and N bound atoms are the most labile at neutral pH and ambient temperatures and are more prone to such exchange. The exchange reaction is primarily catalyzed by OH\(^-\), H\(^+\), and H\(_2\)O (124) and the rate of the reaction is affected by the local chemical environment of the hydrogen, temperature, and pH. The mechanism used to describe the hydrogen exchange includes changes in protein conformation to allow for the exchange of hydrogens that are solvent inaccessible or hydrogens involved in hydrogen bonding interactions. This mechanism describes an H in a closed (protected) state that first changes to an open (solvent exposed) state before exchanging with an isotope, such as deuterium (D). The mechanism first proposed by Linderstrom-Lang (125) (recently reviewed in (126, 127)) and later refined by his students (128) is commonly represented as follows:
where \( C \) and \( O \) denote the closed and open states respectively, \( k_{op} \) is the rate constant for the opening reaction, \( k_{cl} \) is the rate constant for the closing reaction, and \( k_{int} \) is the rate constant for the intrinsic rate of exchange of \( H \) with \( D \) in an unprotected state. Proteins in which the native form is more stable \( (k_{cl} \gg k_{op}) \) and the conformational dynamics occur on a time scale faster than the intrinsic exchange \( (k_{cl} \gg k_{int}) \) follow the EX2 process of exchange. This is in contrast to the EX1 process of exchange, in which the intrinsic exchange occurs on a timescale faster than conformational change \( (k_{int} \gg k_{cl}) \). The EX1 process of exchange generally occurs under conditions favoring protein unfolding, such as high temperature, extreme pH, or in the presence of denaturants. In less extreme conditions, the exchange process is best described by the EX2 process, where the observed rate of exchange \( (k_{ob}) \) is given by,

\[
k_{ob} = \frac{k_{op}}{k_{cl} k_{int}}
\]

Amide (N-H) exchange in short peptides occurs very fast under physiological conditions. In proteins, however, amide hydrogens are the slowest exchangeable hydrogens, exchanging with half lives ranging from milli-seconds to years (127). Sensitive to local environment, the rates at which these amide hydrogens exchange can be used to infer their solvent accessibility. Comparing the exchange rates of the backbone amides for a protein that exists in two different structural states can therefore be used to identify amino acids of altered conformation. The amount of structural information
obtained is dependent on the amino acid resolution of the technique employed to measure hydrogen isotope incorporation into the protein.

Determining exchange rates of individual peptides within a protein using mass spectrometry (MS) is a rapidly evolving technique. It was first demonstrated in 1991, where the conformational changes of bovine ubiquitin induced by methanol were examined (129). These earlier experiments focused on whole proteins, limiting the analyses to samples of small molecular weight, and did not provide detailed structural information. The use of proteolytic fragmentation to increase the resolution of exchange calculations to specific peptides (130, 131) was first applied using high-performance liquid chromatography-fast atom bombardment MS (HPLC-FABMS). These experiments examined structural changes in horse heart cytochrome c induced by increasing temperature (132). Nine fragments generated by pepsin digestion of this 11.7 kDa protein were examined. This treatment afforded 58% coverage of the total 104 amino acid residues of the protein. Identification of the peptide sequences included analysis of the molecular weight and C-terminal sequencing with carboxypeptidases (133). The results agreed with previous heat denaturation studies of cytochrome c, showing a major conformational change that occurs at around 60 °C. The next advance in the use of MS in HD exchange rate calculations of peptides involved introduction of electrospray ionization (ESI). The increased sensitivity of ESI compared to FAB was credited for the 95% sequence coverage observed for horse skeletal myoglobin. This was applied in a comparison of HD exchange rates for the apo and holo forms of this 17 kDa enzyme (134). In this study, it was noted that the relative unpredictable nature of pepsin digestion compromised identification of peptic fragments by mass alone, thus MS/MS
analysis was necessary to identify the peptides. Improvements in protocols have reduced the undesirable back-exchange to 10 – 20% via optimization of sample handling (135), digestion (136), and chromatography (137). HD exchange by MS has been used to study protein folding, stability, dynamics, ligand binding and protein-protein interactions. Currently, the largest protein investigated by HD exchange is aldolase, which exists as a homo tetramer of 157 kDa (138).

To further probe the changes that occur upon oxygen inactivation of transketolase, a MS approach was employed. Through comparison of active and inactivated samples it was hoped that a difference between the two states could be identified. Two levels of analysis were engaged. The first involved direct comparison of the mass spectra generated for each sample. Covalent chemical changes to the inactive protein would be revealed by a difference in protein MW. Due to the large size of transketolase, proteolytic enzymes were used to generate more readily detectable peptides; these peptides were then compared in a qualitative manner. The second level of analysis employed calculations of hydrogen/deuterium (HD) exchange rates to examine structural differences in the active and inactive forms of transketolase. The work presented in this chapter focuses on establishing a method for HD experiments using the CbbT enzyme. To aid in these experiments, Fourier transform-ion cyclotron resonance mass spectroscopy (FTICR-MS) was utilized for data collection. FTICR-MS is currently the most sensitive MS available (139) and capable of measuring mass to < 1 ppm mass accuracy (140). Instrumentation and guidance for these experiments were kindly provided by Prof. Michael Freitas of the Department of Chemistry, The Ohio State University.
5.2 MATERIALS AND METHODS

5.2.1 PREPARATION OF INACTIVATED CBBT

A 1 ml frozen aliquot of purified CbbT was thawed on ice and desalted by gel filtration using a Superdex 200 (120 ml) column. The column was equilibrated and run with 20 mM HEPES-NaOH, pH 7.8, 0.1 mM ThDP, 2 mM MgSO₄, and 0.1% BME. The peak fractions were combined yielding 6 ml of CbbT with a concentration of 0.6 mg/ml. This desalted CbbT sample was divided into 3 equal fractions and placed in 10 ml serum vials. Two vials were sparged with N₂ and the 3rd with O₂ on ice for 20 min. One of the N₂ vials and the O₂ vial were placed in a 25 °C water bath and activities were measured at 3 separate time points spanning a 24 h period. The activity of the N₂ sample stored on ice was assayed at the end of the incubation, at 24 h. All three samples were concentrated using an ultrafree-4 centrifugal filter containing a Biomax-30 membrane (Millipore, Billerica, MA); this was achieved within 1.5 h at 4 °C spinning at a force of 1,000 g. The N₂ sample stored on ice was concentrated to 130 µl, to which 35 µl of glycerol was added, yielding a final concentration of ~18.5 mg/ml. The O₂ and N₂ samples incubated at 25 °C were concentrated to 120 µl, to which 30 µl of glycerol was added, yielding final concentrations of ~ 20 mg/ml. All concentrated samples were divided into 5 aliquots, flash frozen, and stored at -80 °C. Protein concentration determination was by the method of Bradford (BioRad, Hercules, CA) using bovine serum albumin as a standard.
5.2.2 PROTEOLYTIC DIGESTION OF CBBT

Lyophilized pepsin A from porcine stomach mucosa (Sigma, St. Louis, MO) was dissolved in H₂O to achieve a stock concentration of 100 µM. This was aliquoted and stored long-term (up to 3 months) at -80 °C and short term (less than 1 month) on ice. Pepsin digestions were typically performed on ice at a pH of approximately 2.6. The final pepsin to CbbT ratio used was 1:2. For the final HD exchange digest, 1 µl of freshly thawed pepsin stock was added to a 25 µl quenched solution with a pH of ~ 2.5. The 5X quench solution consisted of 10% acetic acid, 50% glycerol, and 10 mM EDTA. This was added to a sample containing 10 mM HEPES-NaOH pH 7.8, 1 mM DTT, 2mM MgSO₄, 0.1 mM ThDP, and 9 µM CbbT (monomer). Following the 2 min digestion, the peptides were then desalted by HPLC (described in section 5.2.4) or by using a desalting trap offline and then analyzed by mass spectrometry. A Peptide MicroTrap™ cartridge (p/n 004/25108/02) (Michrom BioResources Inc., Auburn, CA) was used for offline desalting. The digestion mix was loaded directly onto the proprietary reverse phase matrix of the cartridge by syringe. The bound peptides were then washed with 50 µl of 2% acetic or formic acid and then eluted with 50 µl of 60% acetonitrile or MeOH containing 2% acetic or formic acid.

Modified sequencing grade trypsin (V5111) was purchased from Promega (Madison, WI). Tryptic digestes were performed in 20 mM HEPES-NaOH buffer at pH 7.8 at 37 °C. No effort was made to remove the exogenous cofactors MgSO₄ and ThDP which were still present in concentrations of 1.6 and 0.08 mM respectively. The digest solution consisted of 0.18 µM trypsin and 225 µM CbbT. Following a 16 h digestion, the
solution was placed on ice. Aliquots were removed and prepared for desalting by adding formic acid to a final concentration of 2%.

5.2.3 HD EXCHANGE CONDITIONS

CbbT samples were diluted 1:20 in D$_2$O for various time periods and then quenched. 20 µl of the 20 mg/ml CbbT stock was diluted by adding 10 µl of an 80x incubation buffer (800 mM HEPES-NaOH, pH 7.8, 80 mM DTT, 160 mM MgSO$_4$, and 8 mM ThDP). HD exchange was initiated by adding 0.9 µl of this concentrated CbbT mix to 19.1 µl of D$_2$O or H$_2$O. Although these small volumes are difficult to pipette accurately, care was taken to perform the dilutions consistently. The D$_2$O was allowed to exchange for 8 different time points (0.5, 3, 7.5, 15, 30, 60, 120, 240 min) in addition to a non-exchanged sample that was prepared in water (time point 0). All samples were prepared in triplicate. At the end of the exchange time, the sample was placed on ice and 5 µl of a concentrated quench solution was immediately added (10% formic acid, 5 mM EDTA, and 50% glycerol). The pH was lowered to ~ 2.6, as determined by using a pH meter with larger volumes of incubation buffer and quench solution in the absence of protein at 25 °C. The quenched samples were then flash frozen in liquid N$_2$ and stored at -80 °C until analysis.

5.2.4 LC-MS

An Agilent HPLC system (HP1100 series) equipped with a 6-port manual inject valve was used for LC-MS. Use of a 1 X 50 mm C8 column in conjunction with a steep elution profile allowed for the rapid desalting and slight separation of the peptides.
Pepsin was added to samples freshly thawed on ice and then immediately injected into the loading loop where digestion was allowed to proceed to 2 min, followed by injection onto the column. The mobile phases used for analysis were an aqueous phase of 2% formic acid with 0.025% TFA, and an organic phase of 2% formic acid with 0.025% TFA in MeOH. The mobile phases were degassed by sonication under vacuum and then purged with a low flow of He. The gradient began at 40% organic and increased to 90% over approximately 200 µl in 4 minutes; this was followed by a 95% organic wash of 100 µl in 2 minutes to clean off the column prior to re-equilibrating under the 100% aqueous phase. The HPLC system contained a dead volume of 600 – 900 µl, depending on system pressure. To alleviate the gradient delay, a pre column flow splitter was added immediately after the pump, consisting of a 13 cm piece of I.D. 50 µ fused silica tubing. This produced a split ratio of 5:1 with a pump flow rate of 250 µl min⁻¹ of 50% B, so that approximately 200 µl min⁻¹ would pass through the added silica splitter into a waste container and 50 µl min⁻¹ would continue onto the column. A post column flow splitter was added prior to the ESI, consisting of a 9 cm piece of 100 µ I.D. fused silica tubing. This produced a 50:1 split ratio under the above conditions, resulting in an electrospray flow of approximately 1 µl min⁻¹. A diagram of this setup is shown in Figure 5.1.

Despite the use of He to degas, air bubbles would still gradually buildup in the pumps and eventually interfere with the electrospray. To reduce the number of air bubbles that entered the column, both pumps were purged with ~ 5 ml of mobile phase every 6 runs and then the column was cleaned by washing with 100% organic phase.
Figure 5.1. Scheme of LC-MS setup utilized for measurement of HD exchange. Flow rates during the elution of peptides are indicated along the path of the system. The Styrofoam box contained an ice/water bath to chill the injection valve, loading loop, pre-cooling loop (~ 50 µl), and the column. A series of ice packs were used to sandwich the tubing leading to the mass spectrometer (MS). The split ratios listed for the two three-way splitters were calculated by measuring the solvent flow of a 50:50 A:B isocratic mobile phase.
5.2.5 FTICR-MS DATA COLLECTION AND ANALYSIS

The data were collected with a commercially available 7 T electrospray ionization Fourier transform ion cyclotron resonance (ESI FTICR) instrument (Apex IIe, Bruker, Billerica, MA) configured for external ion accumulation. The ESI source was operated in positive ion mode. The capillary inlet was heated to 80 °C and the capillary voltage was varied between 2200 to 2600 V. The capillary exit voltage was varied from 75 to 125 V and the skimmer voltage was varied from 5 to 15 V. The ions were accumulated in an external hexapole for 0.6 s and then transferred to the ICR cell. Frequencies were scanned from 54550 to 311878 Hz (corresponding to an \( m/z \) range from 350 to 2000) with an excitation chirp rate of 150 Hz /µs. Under the above conditions, data collection progressed at the rate of approximately one scan every 2.3 seconds. The instrument events were controlled by a modified modular ICR data acquisition system (MIDAS) (141, 142). MIDAS Analysis software was also used for post-collection processing of the time domain data including data truncation, fast Fourier transform, Hanning apodization (smoothing function) and baseline correction. For offline data, samples were directly infused and the data averaged, typically consisting of 25 to 75 scans. The resulting file was then processed to generate a spectrum of signal intensity vs. \( m/z \). For LC-MS, data were collected in a time resolved mode where each scan was saved separately. Upon completing a run, the scans were processed to generate a chromatogram of total signal intensity versus data count for the run.
5.2.6 HD EXCHANGE ANALYSIS AND RATE CALCULATIONS

Each LC-MS run lasted a period of 8 min and the resulting chromatogram consisted of approximately 200 scans. To simplify data analysis and improve the signal to noise, every 10 consecutive scans of the time resolved data were averaged. The averaged scans were processed under the standard conditions and a list of monoisotopic peaks and corresponding charge was created for each well-resolved peptide.

Plots of average mass versus time were generated for each peptide in the rate analysis. The average mass was calculated by taking the weighted average abundance of each isotopic peak in the scan of maximum intensity. This process was initially performed in an automated fashion using MIDAS HD exchange batch processing routines. A variety of MIDAS processing settings were tried for comparison. The best results for many of the data were obtained by processing the peaks with absolute threshold and noise settings of 15 and 5, respectively. Most data sets showed improvement upon manual calculation of the average mass, which was performed for spectra displaying large standard deviations in HD exchange rates. To calculate the increase in mass, the average mass calculated for each HD exchange time point was subtracted from the average mass of the peptide in H₂O. These values were plotted versus time and fit to either a two or three order exponential (127), as the one shown below, using Sigma Plot v8.0:

\[ Y = N - Ae^{-k_1x} - Be^{-k_2x} - Ce^{-k_3x} \]

where \( Y \) represents the mass of the peptide at time \( x \), \( N \) is the mass of the fully exchanged peptide, \( A \), \( B \), and \( C \) represent the number of amides exchanging at the apparent rates of \( k_1 \), \( k_2 \), and \( k_3 \) respectively. In the case of the second order exponential,
the final term was omitted. The ability to fit either the three or two order exponential equations was evaluated for the data of each peptide by comparing the standard errors of the fit parameters and the t-statistic/P value for the coefficients.

5.3 RESULTS AND DISCUSSION

5.3.1 PREPARATION OF OXYGEN-INACTIVATED CBBT

It was realized that in the presence of even low concentrations of ammonium sulfate, CbbT would precipitate out upon addition of the acidic quench solution on ice. For the complete removal of salt, a large gel filtration column (1.6 x 50 cm Superdex 200) was used rather than a small G-25 desalting column. With the absence of the stabilizing ammonium sulfate, an additional control was used to evaluate any loss in activity of the N₂ sample incubated at 25 °C; this was stored under a N₂ headspace and stored on ice. To further prevent unnecessary inactivation of the N₂ sample the number of time points at which a sample was removed was kept to a minimum reducing any trace amount of air introduced by the syringe. After 24 h at 25 °C, the O₂-incubated sample had lost 78.8% of the activity of the N₂-incubated sample. The N₂ sample stored on ice during the time course displayed 80% the activity of the N₂ sample incubated at 25 °C for 24 h. This reduction in activity proved to be from incomplete reconstitution and the N₂ on ice sample displayed the same level of activity as the N₂ at 25 °C sample following a 15 min incubation at 25 °C. CbbT dissociation during gel filtration at 4 °C does not occur in the presence of ammonium sulfate. The final specific activities of the N₂ and O₂ samples were 26.2 and 5.5 U/mg respectively. A representative time-course of inactivation is illustrated (Figure 5. 2). Since HD exchange is achieved by diluting
sample into D$_2$O, a highly concentrated sample of CbbT was desired. This was achieved at high recovery using centrifugal spin concentrators. Protein determinations before and after concentration showed a sample recovery of 95%.
Figure 5.2. Preparation of CbbT samples for analysis by mass spectrometry. The plot displays specific activity (measured in $\mu$mol min$^{-1}$ mg$^{-1}$ versus time) after incubation in the presence (○) or absence (●) of oxygen. CbbT was desalted by gel filtration in 20 mM HEPES-NaOH, pH 7.8, 2 mM MgSO$_4$, 0.1 mM ThDP, and 0.1% BME. Samples were incubated in serum vials at 25 °C following a 20 min sparge time with either pure O$_2$ or N$_2$ gas on ice. CbbT concentration was 0.61 mg ml$^{-1}$. Assays were performed as described in Chapter 2, section 2.2.8.
5.3.2 PREPARATION OF PEPTIDASE-GENERATED PEPTIDES OF CBBT

Pepsin is a protease that exhibits a broad range of specificity. It cleaves preferentially on the carboxy-terminal side of large bulky hydrophobic residues. Though other residues may also act as substrates; it does not cleave next to Val, Ala, or Gly residues (143). Pepsin was used for peptide formation primarily because it has a pH optimum in the range from 1 – 4 and is still active at 0 °C (ideal conditions to quench HD exchange). The reduced efficiency of proteolysis encountered at these low temperatures is typically countered by increasing the amount of pepsin to a ratio of 1:1 enzyme:substrate. The broad specificity of this enzyme makes prediction of cleavage sites difficult. Digestion by pepsin was first optimized using myoglobin as a control. Since pH and temperature were selected based on quenching conditions, the primary factors changed included varying the ratio of pepsin to substrate and the digest reaction time. For these experiments, the resulting peptides were desalted using a small peptide trap cartridge in an offline manner using syringes. The homogenous mixture of peptides was then analyzed by FTICR-MS. The mass spectra were typically averaged over 50 scans, depending on signal strength, and masses then analyzed by hand to generate a list of peptides based on their monoisotopic masses. Several pepsin to CbbT ratios (1:1, 2:1, 1:2, 1:3, 1:4, and 1:6) were examined for their effect on the number of peptides produced. It was found that above a ratio of 1:2 there was little effect on the number of peptides produced after a two min digestion at 0 °C. The 1:3, 1:4, and 1:6 samples produced a similar number of peptides; though the relative abundance of many peptides were lower when higher pepsin ratios were used. To reduce any artifacts by pepsin, such as autolysis products, the lowest effective concentration (1:2) was used.
A standard digest performed in the absence of EDTA and analyzed offline would produce approximately 50 – 60 well resolved peptides in the range of 500 to 1,200 m/z. It has been demonstrated that holo-transketolase from yeast is more resistant than apo-transketolase to digestion by trypsin (144). It has also been noted that pepsin digestion of partially unfolded ERK2 resulted in a different cleavage pattern than native ERK2 (145). The possibility that higher order structure was preventing a more complete digest of CbbT by pepsin was thus examined. The effect of EDTA and urea additives on the number of peptides generated during pepsin digestion was examined. Each additive increased the number of peptides, now displaying 98 and 102 well resolved peptides for the EDTA and urea samples, respectively, compared to 61 observed in a control sample. The mass spectra of the peptide mixtures generated under these conditions are illustrated in Figure 5.3. The samples containing the additives retained most of the peaks present in the control. The signal intensity of the peaks in the EDTA sample decreased by roughly 50% in comparison to the control, whereas the urea sample decreased by 80%. For this reason EDTA was routinely added to subsequent pepsin digests. A considerable amount of spectral overlap also accompanied this increase in observed peptides.
Figure 5.3. Effect of additives on CbbT digestion by pepsin. Each digest was performed at pH 2.6 and contained (upon quenching) 2% acetic acid, 10% glycerol, 10 mM HEPES-NaOH 1 mM DTT, 0.1 mM ThDP, 2 mM MgSO$_4$, and 9 μM CbbT (monomeric); digestion was initiated by the addition of 4 μM pepsin. After 2 min, the peptides were desalted using a MicroTrap$^\text{TM}$ peptide cartridge as described in the text (5.2.2). The resulting peptide mixes were then analyzed by FTICR-MS; spectra plotted as absolute abundance (arbitrary units) versus m/z for each condition are shown. (A), control with no additives; (B), addition of 2 mM EDTA; (C), addition of 2 M urea.
Figure 5.3
5.3.3 Assignment of Petidase-Generated CbbT Fragments

The program MS-Digest, of the Protein Prospector suite of proteomics tools (146), was used to generate a list of possible peptides that could be produced upon pepsin digestion of CbbT. Various parameters may be considered by the software. These include an indication of the maximum number of missed cleavages, possible modified side chains (i.e. via oxidation of Met, conversion of N-terminal Gln to pyroGlu, or intentional chemical modifications of Cys), minimum and maximum fragment masses, minimum fragment length, and the presence of multiple charged peptides. Predicted peptides for a pepsin digestion of CbbT generated a list of approximately 25,000 fragments that included 15 potential missed cleavages and multiple charges on fragments from 500 to 1,200 m/z. Of the 50 – 60 well resolved peptides that were observed in a standard pepsin digest in the absence of EDTA, approximately 25 were matched with the predicted fragments within ± 10 ppm. Use of ubiquitin as an internal standard for calibration in offline data did not improve the number of matches. In the absence of an added internal standard, various peptides were selected as positive matches throughout the m/z range to be used for internal calibration. LC-MS data must be externally calibrated and the mass accuracy is dependent upon the abundance of the peptide.

The use of the more sequence-specific peptidase trypsin, which cleaves the carboxy-terminal side of Arg and Lys, was also examined for comparison. Under similar prediction conditions described for the above pepsin in silico digest, Protein Prospector generated a list of 29,000 possible peptides, however unlike the pepsin prediction ~ 90% of these peptides had a charge state well above any observed experimentally with CbbT (< +8). This digest was analyzed both online by LC-MS and offline using the
MicroTrap™ desalting cartridge. Each method yielded a similar number of well resolved ions, 53 and 47 for the offline and online methods, respectively. Comparison of the ions detected between these two methods showed that 20 were found in both runs.

Comparison with the theoretical peak list generated by Protein Prospector found only 20 and 22 of the total peaks as reasonable matches for the offline and online methods, respectively. Lists of assigned peaks for the trypsin offline and online analysis are presented in Table 5.1 and 5.2, respectively. For the offline peak list, 17 of the assigned peaks were the result of 0 missed cleavages, with the remaining 3 peaks containing only 1 missed cleavage. By contrast, only 11 of the online method peaks contained 0 missed cleavages and the remaining fragments contained up to 3 missed cleavages. It also appeared that the online method was better suited for the detection of larger peptides and was more conducive for peptides carrying more charge. This was also shown by the multiple occurrence of the same peptide fragment at different charge states, for example the peptide with an m/z value of 966.8 and a +3 charge was also observed at m/z 725.4 as a +4 and at m/z 580.5 as a +5. In addition to the multiple charge states, modified and unmodified forms of the same peptide were also detected; this was observed for both methods. The oxidized and unoxidized forms of Met 21 were observed by the offline method, whereas only the unoxidized form was observed online. Also the formation of pyroGlu was observed by both methods at position Glu 256; in both cases the unmodified form was also observed. Another instance of pyroGlu modification was observed for residue Glu 96 by the offline method; however this peptide was not detected online.

Originally it was believed that the peptide sequences could be deduced by m/z values alone. However, with less than a 50% match for a given set of trypsin- or pepsin-
generated peptides, this does not seem feasible. This lack of assignments was initially perplexing, though it has been reported that for confident assignments of pepsin-generated peptides, an alternate method of sequencing, such as tandem MS (MS/MS) is generally required (135, 147).
Table 5.1. Assignment of 20 of the 53 well-resolved tryptic CbbT peptides by the offline method. Highlighted sequences were detected in more than one form, containing a predicted modification. Peptides are listed by charge ($z$) and experimentally determined mass/charge ($m/z$) ratios. Closeness of match to the theoretical $m/z$ is given in parts per million (ppm).
<table>
<thead>
<tr>
<th>z</th>
<th>m/z exp. (mi)^a</th>
<th>m/z theo. (mi)^a</th>
<th>Mod.^b</th>
<th>Start^c AA#</th>
<th>End^c AA#</th>
<th>MC^d</th>
<th>Database Sequence</th>
<th>ppm^e</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>474.245</td>
<td>474.2475</td>
<td></td>
<td>18</td>
<td>26</td>
<td>0</td>
<td>ALAMDAVEK</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>480.7408</td>
<td>480.7414</td>
<td></td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>DIGAAQETR</td>
<td>-1.3*</td>
</tr>
<tr>
<td>2</td>
<td>482.2430</td>
<td>482.245</td>
<td>1Met-ox</td>
<td>18</td>
<td>26</td>
<td>0</td>
<td>ALAMDAVEK</td>
<td>-4.2</td>
</tr>
<tr>
<td>2</td>
<td>510.2786</td>
<td>510.2802</td>
<td></td>
<td>246</td>
<td>255</td>
<td>0</td>
<td>TVIGYGAPNK</td>
<td>-3.1</td>
</tr>
<tr>
<td>1</td>
<td>527.2948</td>
<td>527.2942</td>
<td>pyroGlu</td>
<td>98</td>
<td>102</td>
<td>0</td>
<td>QLGAR</td>
<td>1.1</td>
</tr>
<tr>
<td>2</td>
<td>537.8017</td>
<td>537.7957</td>
<td></td>
<td>49</td>
<td>58</td>
<td>0</td>
<td>FLTVDPAPK</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>627.6578</td>
<td>627.6508</td>
<td>pyroGlu</td>
<td>256</td>
<td>275</td>
<td>0</td>
<td>QGGHDVHGAPLGAIAEIAAAAR</td>
<td>11.2</td>
</tr>
<tr>
<td>3</td>
<td>633.3344</td>
<td>633.3263</td>
<td></td>
<td>256</td>
<td>275</td>
<td>0</td>
<td>QGGHDVHGAPLGAIAEIAAAAR</td>
<td>12.8*</td>
</tr>
<tr>
<td>1</td>
<td>675.3662</td>
<td>675.3612</td>
<td></td>
<td>12</td>
<td>17</td>
<td>0</td>
<td>MANAIR</td>
<td>7.4</td>
</tr>
<tr>
<td>3</td>
<td>685.0228</td>
<td>685.0162</td>
<td></td>
<td>388</td>
<td>407</td>
<td>0</td>
<td>GMVSVPAGAFAGSYIHGYR</td>
<td>9.6</td>
</tr>
<tr>
<td>3</td>
<td>696.6645</td>
<td>696.6662</td>
<td></td>
<td>29</td>
<td>48</td>
<td>0</td>
<td>SGHPGPMGMADVATVLFRN</td>
<td>-2.4</td>
</tr>
<tr>
<td>1</td>
<td>715.3991</td>
<td>715.399</td>
<td></td>
<td>344</td>
<td>350</td>
<td>0</td>
<td>LSJEAPK</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>725.3878</td>
<td>725.3802</td>
<td></td>
<td>246</td>
<td>275</td>
<td>1</td>
<td>TVIGYGAPNKQGGHDVHGAPLGAIAEIAAAR</td>
<td>10.5</td>
</tr>
<tr>
<td>3</td>
<td>756.7211</td>
<td>756.7252</td>
<td></td>
<td>386</td>
<td>407</td>
<td>1</td>
<td>SKGMVSVPAGAFAGSYIHGYR</td>
<td>-5.4</td>
</tr>
<tr>
<td>1</td>
<td>784.4629</td>
<td>784.4681</td>
<td></td>
<td>313</td>
<td>319</td>
<td>0</td>
<td>LQASPLR</td>
<td>-6.6*</td>
</tr>
<tr>
<td>1</td>
<td>818.4022</td>
<td>818.4008</td>
<td></td>
<td>529</td>
<td>535</td>
<td>0</td>
<td>DENLTAR</td>
<td>1.7</td>
</tr>
<tr>
<td>4</td>
<td>866.1865</td>
<td>866.184</td>
<td></td>
<td>103</td>
<td>137</td>
<td>0</td>
<td>TAGHPEYGHAEGIEVTTPGLQGIATAVGMALAER</td>
<td>2.9</td>
</tr>
<tr>
<td>3</td>
<td>983.185</td>
<td>983.1664</td>
<td></td>
<td>356</td>
<td>385</td>
<td>0</td>
<td>ASEMALGVVNEALPFAVGSGADLTGSNLTR</td>
<td>18.9</td>
</tr>
<tr>
<td>3</td>
<td>1025.8676</td>
<td>1025.865</td>
<td></td>
<td>355</td>
<td>385</td>
<td>1</td>
<td>KASEMALGVVNEALPFAVGSGADLTGSNLTR</td>
<td>2.8*</td>
</tr>
<tr>
<td>2</td>
<td>1090.0473</td>
<td>1090.05</td>
<td></td>
<td>320</td>
<td>341</td>
<td>0</td>
<td>AAFETAEAAADTSAHLPPAIAAYK</td>
<td>-2.6*</td>
</tr>
</tbody>
</table>

*a – m/z of peptide based on the monoisotopic mass (mi).
*b – Mod.: predicted modifications of the peptide.
*c – AA# : amino acid number based on CbbT sequence.
*d – MC : the number of missed cleavages to produce the peptide.
*e – * denotes this peptide was used for calibration.

Table 5.1
Table 5.2. Assignment of 22 of the 47 well-resolved tryptic CbbT peptides by LC-MS. Highlighted sequences were detected in more than one form, either with the predicted modification listed or in a different charge state. Peptides are listed by charge ($z$) and experimentally determined mass/charge ($m/z$). Closeness of match to the theoretical $m/z$ is given in parts per million (ppm).
<table>
<thead>
<tr>
<th>z</th>
<th>m/z exp.</th>
<th>m/z theo. Mod. (^a)</th>
<th>Start(^b)</th>
<th>End(^c)</th>
<th>MC(^d)</th>
<th>Database Sequence</th>
<th>ppm(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>474.2479</td>
<td>474.2475</td>
<td>18</td>
<td>26</td>
<td>0</td>
<td>ALAMDAVEK</td>
<td>0.8*</td>
</tr>
<tr>
<td>4</td>
<td>567.7969</td>
<td>567.7959</td>
<td>386</td>
<td>407</td>
<td>1</td>
<td>SKGMVSVPAGAFAGSYIHYGIR</td>
<td>1.8*</td>
</tr>
<tr>
<td>5</td>
<td>580.5034</td>
<td>580.5058</td>
<td>246</td>
<td>275</td>
<td>1</td>
<td>TVIGYGAPNKOQGHVDHVAPLGAEEAAAARR</td>
<td>-4.1</td>
</tr>
<tr>
<td>3</td>
<td>627.6411</td>
<td>627.6508 pyroGlu</td>
<td>256</td>
<td>275</td>
<td>0</td>
<td>QGGHDHVAGPLGAEEAAAARR</td>
<td>-15.5</td>
</tr>
<tr>
<td>3</td>
<td>633.3221</td>
<td>633.3263</td>
<td>256</td>
<td>275</td>
<td>0</td>
<td>QGGHDHVAPLGAEEAAAARR</td>
<td>-6.6</td>
</tr>
<tr>
<td>3</td>
<td>696.6627</td>
<td>696.6662</td>
<td>29</td>
<td>48</td>
<td>0</td>
<td>SGHPGMPMADVATVLNR</td>
<td>-5.0*</td>
</tr>
<tr>
<td>4</td>
<td>725.3741</td>
<td>725.3802</td>
<td>246</td>
<td>275</td>
<td>1</td>
<td>TVIGYGAPNKOQGHVDHVAPLGAEEAAAARR</td>
<td>-8.4</td>
</tr>
<tr>
<td>3</td>
<td>756.7198</td>
<td>756.7252</td>
<td>386</td>
<td>407</td>
<td>1</td>
<td>SKGMVSVPAGAFAGSYIHYGIR</td>
<td>-7.1*</td>
</tr>
<tr>
<td>4</td>
<td>844.4319</td>
<td>844.4181</td>
<td>65</td>
<td>94</td>
<td>0</td>
<td>FVLSAGHGSMLLYAIHILLGYADMDMQIR</td>
<td>16.3</td>
</tr>
<tr>
<td>5</td>
<td>858.8368</td>
<td>858.8368</td>
<td>65</td>
<td>102</td>
<td>2</td>
<td>FVLSAGHGSMLLYAIHILLGYADMDMQIR</td>
<td>0.0</td>
</tr>
<tr>
<td>7</td>
<td>861.3122</td>
<td>861.3178</td>
<td>481</td>
<td>535</td>
<td>3</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVLSQNLPTVRTEHRDLTAR</td>
<td>-6.5</td>
</tr>
<tr>
<td>4</td>
<td>866.1902</td>
<td>866.184</td>
<td>103</td>
<td>137</td>
<td>0</td>
<td>TAHGPEYGAEGIVTGGPQGIATAVGMAER</td>
<td>7.2*</td>
</tr>
<tr>
<td>3</td>
<td>940.9405</td>
<td>940.9152</td>
<td>481</td>
<td>524</td>
<td>1</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVLSQNLPTVR</td>
<td>26.9</td>
</tr>
<tr>
<td>3</td>
<td>966.8311</td>
<td>966.8377</td>
<td>246</td>
<td>275</td>
<td>1</td>
<td>TVIGYGAPNKOQGHVDHVAPLGAEEAAAARR</td>
<td>-6.8*</td>
</tr>
<tr>
<td>4</td>
<td>973.7937</td>
<td>973.7782</td>
<td>481</td>
<td>517</td>
<td>0</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVSLR</td>
<td>15.9</td>
</tr>
<tr>
<td>3</td>
<td>983.1959</td>
<td>983.1664</td>
<td>356</td>
<td>385</td>
<td>0</td>
<td>ASEMALGVVEALPFAVGGSADLTGSLTR</td>
<td>30.0</td>
</tr>
<tr>
<td>6</td>
<td>1004.7251</td>
<td>1004.703</td>
<td>481</td>
<td>535</td>
<td>3</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVLSQNLPTVRTEHRDLTAR</td>
<td>22.2</td>
</tr>
<tr>
<td>3</td>
<td>1025.8756</td>
<td>1025.865</td>
<td>355</td>
<td>385</td>
<td>1</td>
<td>KASEMALGVVEALPFAVGGSADLTGSLTR</td>
<td>10.6</td>
</tr>
<tr>
<td>9</td>
<td>1026.9973</td>
<td>1027.02</td>
<td>388</td>
<td>407</td>
<td>0</td>
<td>GMVSVPAGAFAGSYIHYGIR</td>
<td>-22.5</td>
</tr>
<tr>
<td>2</td>
<td>1090.0317</td>
<td>1090.05</td>
<td>320</td>
<td>341</td>
<td>0</td>
<td>AAFETAAADTSALPPAAAYK</td>
<td>-16.9</td>
</tr>
<tr>
<td>4</td>
<td>1175.8981</td>
<td>1175.892</td>
<td>481</td>
<td>524</td>
<td>1</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVLSQNLPTVR</td>
<td>5.1*</td>
</tr>
<tr>
<td>3</td>
<td>1298.0268</td>
<td>1298.035</td>
<td>481</td>
<td>517</td>
<td>0</td>
<td>AIPNLAVPADAVETAEAEFWEIAMATATSTPTLLVLSR</td>
<td>-6.3*</td>
</tr>
</tbody>
</table>

\(a\) – m/z of peptide based on the monoisotopic mass (mi).

\(b\) – Mod.: predicted modifications of the peptide.

\(c\) – AA#: amino acid number based on CbbT sequence.

\(d\) – MC: the number of missed cleavages to produce the peptide.

\(e\) – * denotes this peptide was used for calibration.

Table 5.2
5.3.4 ONLINE AND OFFLINE MS ANALYSIS OF HD EXCHANGE

A single time point comparison of the online versus offline technique was performed using CbbT-active and O₂-inactivated samples that were exchanged in D₂O for 2 h. These digests were performed in the absence of EDTA and all data was collected within 12 min of initiating the digest. Analysis of the offline data focused on 24 peptides that were observed consistently in the three replicates for each sample. However, in a manner similar to that observed for the tryptic digest, not all peaks observed offline could be found in the online data, and vice versa. Of the 24 peptides observed offline, 13 of these were found consistently in a single set of online data. Table 5.3 shows a comparison of the average masses by which these peptides increased following the 2 h incubation in D₂O.
<table>
<thead>
<tr>
<th>peptide charge</th>
<th>peptide m$_i$</th>
<th>N$_2$ Offline</th>
<th>N$_2$ Online</th>
<th>O$_2$ Offline</th>
<th>O$_2$ Online</th>
<th>Δ Mass</th>
<th>Δ Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>625.4311</td>
<td>2.7</td>
<td>9.8</td>
<td>3.7</td>
<td>12.1</td>
<td>1.0</td>
<td>2.3</td>
</tr>
<tr>
<td>2</td>
<td>642.4522</td>
<td>1.0</td>
<td>4.3</td>
<td>1.0</td>
<td>5.3</td>
<td>0.1</td>
<td>1.0</td>
</tr>
<tr>
<td>4</td>
<td>660.9267</td>
<td>1.0</td>
<td>5.4</td>
<td>1.6</td>
<td>7.0</td>
<td>0.6</td>
<td>1.5</td>
</tr>
<tr>
<td>1</td>
<td>683.5293</td>
<td>0.5</td>
<td>1.6</td>
<td>0.6</td>
<td>1.7</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>703.1286</td>
<td>0.6</td>
<td>2.6</td>
<td>1.1</td>
<td>4.6</td>
<td>0.5</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>760.2699</td>
<td>5.1</td>
<td>12.4</td>
<td>6.4</td>
<td>12.0</td>
<td>1.3</td>
<td>-0.4</td>
</tr>
<tr>
<td>4</td>
<td>766.7800</td>
<td>2.2</td>
<td>5.0</td>
<td>4.2</td>
<td>6.4</td>
<td>2.0</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>781.5759</td>
<td>3.0</td>
<td>9.7</td>
<td>3.4</td>
<td>9.7</td>
<td>0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>3</td>
<td>783.5395</td>
<td>4.9</td>
<td>10.4</td>
<td>4.8</td>
<td>9.9</td>
<td>-0.1</td>
<td>-0.5</td>
</tr>
<tr>
<td>3</td>
<td>797.8816</td>
<td>2.6</td>
<td>5.2</td>
<td>3.1</td>
<td>6.5</td>
<td>0.5</td>
<td>1.3</td>
</tr>
<tr>
<td>2</td>
<td>812.0456</td>
<td>0.6</td>
<td>1.5</td>
<td>0.7</td>
<td>1.6</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>819.5561</td>
<td>0.0</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1</td>
<td>1106.8160</td>
<td>1.9</td>
<td>3.3</td>
<td>1.9</td>
<td>3.4</td>
<td>0.0</td>
<td>0.1</td>
</tr>
</tbody>
</table>

**Table 5.3.** Comparison of HD exchange measured online versus offline. Peptides were generated with pepsin using either N$_2$-incubated active CbbT or O$_2$-incubated inactivated CbbT. The peptides are listed by the charge and monoisotopic mass (m$_i$) of the N$_2$ offline sample. The monoisotopic masses for the N$_2$ and O$_2$ samples in H$_2$O did not differ by more than 3 PPM. The increase in mass values were calculated from the average mass of the peptide following a 2 h incubation in D$_2$O after subtracting the average mass of the same peptide in H$_2$O. For comparison of deuterium (D) incorporated into the O$_2$ versus N$_2$ sample, the last two columns display the difference upon subtracting the N$_2$ mass increase from the O$_2$ mass increase.
Comparing the observed increase in mass calculated for each method indicated that the online method displayed 2 – 5 times more incorporation of D than the offline method. The decrease in incorporation observed for the offline method is most likely a result of back exchange during sample preparation and analysis. This was in contrast to a previous comparison of these two methods using FTICR-MS, which displayed similar incorporation of D into myoglobin peptic fragments by both methods (148). The LC-MS method used to compare the techniques in the myoglobin study utilized a larger column, which resulted in longer run times. The offline technique offers a mechanistically simpler alternative to LC-MS but requires constant sample manipulation prior to injection into the MS. The opportunity for back exchange is increased as well as the occurrence of artifactual forward exchange (127), making it difficult to ensure consistency between runs. It was noted that differences in D incorporation between the O$_2$ and N$_2$ samples followed a similar trend for both techniques, with the exception of one peptide ($^4$; mi = 760.2699). Though the lack of LC-MS replicates prevented a measure of standard error, it was still noted that for a few peptides, the O$_2$ sample had accumulated more D than the N$_2$ sample. The apparent reduction in back exchange during LC-MS made this the method of choice for kinetic analysis of HD exchange rates. In addition, with the enhanced number of peptides observed by the addition of EDTA to the digest, the ability to separate some of the overlapping peaks by LC-MS was deemed advantageous.
5.3.5 ANALYSIS OF CBBT HD EXCHANGE RATES BY LC-MS

Triplicate samples representing 9 time points of HD exchange were prepared for both the N₂-active and the O₂-inactivated CbbT. Preparation of fully exchanged control samples was attempted by incubating at a higher temperature for 4 h. Unfortunately, these samples precipitated out of solution. These samples would have been useful to estimate the extent of back exchange during sample processing, but were not necessary for comparing the O₂- and N₂-CbbT samples. Of the 54 samples, data were recorded for all but one (the 30 min time point for the N₂ sample), which was lost due to a program failure during the LC-MS run. To minimize back exchange, it was ensured that data for every run was collected within 12 min of initiating the digest. The entire data set was collected in two batches over a period of 2 days. In general, each digest eluted in a reproducible fashion and in some cases afforded the separation of peptides which would have overlapped in the mass spectrum. A chromatogram showing a representative example of an effective separation is shown in Figure 5.4.
Figure 5.4. Chromatographic separation of a CbbT pepsin digest. The chromatograph produced by one of the N2-CbbT samples in H2O is presented as an example. (A) The total ion chromatogram (TIC) is represented by a dotted line, plotted as relative intensity versus data count. The three smaller peaks labeled I, II, and III/IV are selected ion chromatograms (SICs) normalized to equal height. The $m/z$ ranges specified for the individual SICs are as described: I, 807.05 – 807.20; II, 806.4 – 806.5; III, IV – 804.6 – 805.3. Normalized mass spectra for each of these ranges are given in (B-D). (B) Peak I, spectrum 127. Consists of a peptide with $m/z$ of 807.1276 and a charge of +3. (C) Peak II, spectrum 149. Consists of a peptide with $m/z$ of 806.4482 and a charge of +2. (D) Peaks III/IV, spectrum 162. Consists of a peptide with $m/z$ of 804.7755 with a charge of +3 (III) and a peptide with $m/z$ of 807.2564 with a charge of +4 (IV).
Figure 5.4
Despite the apparent consistency in elution profile, there was variability in the number of peptides detected per run, which ranged from approximately 75 to 200 well-resolved peaks. Careful analysis of the data generated by the H$_2$O, O$_2$, and, N$_2$ samples was performed to select for the peptides reproducibly present. Lists of the peptides detected in each of the six runs were manually generated and compared. This list revealed 28 peptides that were present in all replicates for each sample and an additional 84 peptides that were present in at least 2 of each of the 3 replicates. These 112 peptides were the focus of further analysis.

Weighted averages of the isotopic distribution for each peptide at the various time points were calculated using the HD batch processing feature of MIDAS. The required parameters for each peptide consisted of a starting and ending $m/z$ range as well as the charge of the peptide. Starting ranges were selected to include the monoisotopic peak. These values as well as the charge states were already known from the peak list generated from the H$_2$O samples. The ending $m/z$ values were selected by examination of the last HD exchange time point, which displayed the largest increase in mass. An additional parameter allows for the selection of a discrete range of scans (i.e. chromatogram data count) in which to find the desired peptide. Though potentially useful, this feature was only applied in a few instances due to variability in the scan numbers from run to run. The HPLC injection, ESI voltage, and data collection were all initiated manually, which could be done reproducibly. Frequently however, after the HPLC run and ESI voltage were started, time was spent fine-tuning the electrospray prior to beginning the data collection. Care was taken to record the data count numbers at various times following
digestion, as the stop watch was always begun upon initiating the pepsin digest, it was believed that these recorded times could later be used to synchronize the data. However, it later became apparent that this was not feasible.

Of the list of 112 peptides, plots of increasing mass versus reaction time were generated using the MIDAS. The resulting plots were ascribed to one of three groups relative to their rates of HD exchange. In the first group, 45 peptides tended to exhibit first order exchange with an overall higher net exchange by the end point of 4 h. Upon closer inspection it appeared they were also relatively well separated by m/z values from peptides of similar charge. The second group of peptides displayed very little increase in mass with little deviation in error. This group most likely represents peptides that exhibit very little net exchange after 4 h. These peptides may posses a high degree of hydrogen bonding and/or are primarily located within the solvent inaccessible core of the protein (149). It is also possible that these peptides undergo rapid back-exchange during processing. For closer analysis of these peptides, sequence analysis is required and a back exchange control could also be helpful. The third group displayed seemingly random changes in mass with large deviations in error. It was concluded that the peptides in this group either overlapped with other peptides of similar charge, complicating an accurate quantitation by MIDAS, or displayed inconsistent intensities throughout the runs that were either completely or at least partially hidden from MIDAS. To reduce the requirement for manual recalculation of these data points, analysis focused on the first group of 45 peptides. An example displaying the increase in mass over time as observed by MS is shown in Figure 5.5 for a peptide that displayed different HD exchange rates.
Figure 5.5. Mass spectra demonstrating HD exchange data for peptide 703$^{+3}$ of CbbT at 25 °C. N$_2$-active CbbT is on the left hand side and O$_2$-inactivated CbbT is on the right hand side. The average mass calculated for each of these time points was used to calculate the rates of amide exchange for this peptide in these two samples. The time at which the exchange reaction was quenched is given to the right of the O$_2$ spectra in minutes. The scale on the x-axis is in units of m/z and spans from 702.5 to 707.0. Abundance, on the y-axis, has been normalized for each spectra.
Figure 5.5
Upon manual recalculation of data points that displayed large standard deviations from the first group, a final data set for 38 peptides was compiled. These peptides displayed deuterium incorporation from 1.5 to 20 mass units after the 4 h exchange period. The standard deviations in the calculation of peptide mass directly correlated with the quality of the data and ranged from 0.05 to 1 mass unit. Of the 38 plots, 14 displayed some degree of variation between the active N2- and O2-inactivated CbbT samples. Though no large differences in net exchange were observed, different rates of exchange were apparent. These plots were then fit to 2nd and 3rd order exponential functions for comparison as described in section 5.2.6. A sample set of data collected for five peptides that did not differ in exchange are presented in Figure 5.6.
Figure 5.6. Representative peptides from active N$_2$- and O$_2$-inactivated CbbT displaying similar deuterium exchange kinetics. In all plots, closed circles (●) represent N$_2$-incubated (active CbbT) and open circles (○) represent O$_2$-incubated (inactivated CbbT). The curve fit results are shown as a dashed line for the N$_2$ data and as a solid line for the O$_2$ data. The plots display the increase in mass observed for the peptide versus time on a linear scale (left hand side), as well as a logarithmic scale (right hand side). Plots were constructed as described in the text (5.2.6). Peptides were named based on monoisotopic mass and charge, values determined for the fit parameters can be found in Table 5.4. (A), 671$^{+2}$; (B), 798$^{+3}$; (C), 807$^{+4}$; (D), 864$^{+2}$; (E), 914$^{+4}$. 
Figure 5.6 (continued on the next page)
Figure 5.6
Graphical data for 11 of the 14 peptides that displayed differences between the N2 and O2 samples are presented in Figure 5.7. The remaining three peptides, 766+3, 725+4, and 1188+2 were omitted from this figure as these peptides displayed low abundance during various runs. This low abundance prevented the calculation of data at certain time points or standard deviations at most time points. For these reasons they were divided in a separate class of peptides that displayed potential differences (Figure 5.8). For most of the peptides that exhibited differences between N2 and O2 CbbT, the peptide from the O2 sample increased in mass faster than the N2 sample. This suggested the O2 sample contained amides which display an increased rate of deuterium exchange (Figure 5.7 A, D, E, G, H, J). One notable exception to this trend was observed for peptide 740+3 (Figure 5.7 F), which displayed increased exchange rates for the N2 sample. The non-linear fit parameters calculated for this peptide suggested this difference is the result of 2 amides in the O2 sample that experienced a decreased exchange rate ($k < 0.02 \text{ min}^{-1}$) compared to the N2 sample (Table 5.4). The cleanest graphical representation of a difference in exchange rates was observed for peptide 703+3 (Figure 5.7 E). Though this peptide only approached the complete exchange of 3 amide hydrogens, it was consistently detected in most of the runs. The higher abundance of this peptide allowed for the accurate calculation of exchange rates with small standard deviations at each time point.
Figure 5.7. Deuterium exchange of CbbT peptides displaying differences between active N$_2$- and O$_2$-inactivated CbbT. In all plots closed circles (●) represent N$_2$-incubated (active) CbbT and open circles (○) represent O$_2$-incubated (inactive) CbbT. The non-linear fit results for the N$_2$ data are shown as a dashed line and for the O$_2$ data as a solid line. Plots display the increase in mass observed for the peptide versus time on a linear scale (left hand side), as well as a logarithmic scale (right hand side). Plots were constructed as described in the text (5.2.6). Peptides were named based on monoisotopic mass and charge; values determined for the fit parameters can be found in Table 5.4. (A) 575$^{+4}$, (B) 644$^{+3}$, (C) 649$^{+3}$, (D) 660$^{+4}$, (E) 703$^{+3}$, (F) 740$^{+3}$, (G) 716$^{+4}$ (H) 729$^{+2}$, (I) 758$^{+2}$, (J) 792$^{+3}$, (K) 952$^{+2}$. 
Figure 5.7 (continued on the next page)
Figure 5.7 continued

D

Increase in mass

660^{+4}

Time (min)

E

Increase in mass

703^{+3}

Time (min)

F

Increase in mass

740^{+3}

Time (min)

Figure 5.7 (continued on the next page)
Figure 5.7 continued.

Figure 5.7
Figure 5.8. Deuterium exchange of CbbT peptides displaying potential differences between active N$_2$- and O$_2$-inactivated CbbT. In all plots, closed circles (●) represent N$_2$-incubated (active CbbT) and open circles (○) represent O$_2$-incubated (inactive) CbbT. The curve fit results for the N$_2$ data are shown as a dashed line and for the O$_2$ data as a solid line. The plots display the increase in mass observed for the peptide versus time on a linear scale (left hand side) as well as a logarithmic scale (right hand side). Plots were constructed as described in the text (5.2.6). Peptides were named based on monoisotopic mass and charge; values determined for the fit parameters can be found in Table 5.4. (A) 660$^{+4}$, (B) 725$^{+4}$, (C) 1188$^{+2}$. 
Figure 5.8
A table of HD exchange rate constants calculated for these 38 peptides is shown in Table 5.4. Most of the rate constants (72% of all amides measured) ranged from 0.01 to 5 min\(^{-1}\) similar to rates reported by others (150). There was also a large group (24%) that exchanged very quickly, with rates above 10 min\(^{-1}\) and represented hydrogens that are not protected or involved in hydrogen bonding. HD exchange rates in this fast range are similar to those calculated for unfolded peptides (151). Since these values could not be calculated accurately with the time points used for our data collection, amides exhibiting rates above 10 min\(^{-1}\) were grouped together and labeled as very fast. A smaller group (4%) displayed rates from 0.006 to 0.01 min\(^{-1}\) and represented more shielded amides; these were grouped together and labeled as very slow. Most peptides that displayed similar exchange rates graphically also displayed similar rate constants. There were a few exceptions that did display small differences in calculated constants. These peptides displayed large standard deviations and therefore were not classified as different based on the calculated HD exchange rate constants. An example of this was observed for peptide 664\(^{+4}\), in which approximately 10 hydrogens exchanged in both the O\(_2\) and N\(_2\) forms of CbbT. The fitted parameters suggest that upon oxygen inactivation, 4 amides experienced a decrease in exchange rate and 3 experienced an increase in exchange rate (Table 5.4). These different HD exchange rates can not be confidently designated as different since the data for the O\(_2\) and N\(_2\) CbbT peptides 664\(^{+4}\) exhibit overlapping standard deviations (Figure 5.9).
Table 5.4. Exchange rate constants for 38 peptides of active and inactive CbbT. The parameters were obtained by a non-linear fit to either a second or third exponential equation as described in the text. Peptides are listed by monoisotopic mass and charge of the non-deuterated sample. The labels N\textsubscript{2} and O\textsubscript{2} correspond to the active and inactivated forms of CbbT, respectively; those fit to a third order function have labels in \textbf{bold} type (i.e. ‘N\textsubscript{2}’). The peptides that displayed the greatest difference between N\textsubscript{2} and O\textsubscript{2} are listed first and appear in \textbf{bold} type (i.e. ‘706\textsuperscript{+3}’). For simplification, the parameters were grouped based on the exchange rates. Very slow (< 0.02 min\textsuperscript{-1}) and very fast (>10 min\textsuperscript{-1}) exchange rates, which could not be accurately calculated by the time points used, were grouped together quantitatively under the columns A and D respectively. Amides of intermediate (0.02 – 0.1 min\textsuperscript{-1}) and fast (0.1 – 10 min\textsuperscript{-1}) exchange were placed under the columns B and C respectively. The parameters N, A, B, C, and D are all expressed as number of amide hydrogens; standard error for N is given in parenthesis; Rates for exchange, \(k_{\text{intermediate}}\) and \(k_{\text{fast}}\) are expressed in the units of min\textsuperscript{-1} with standard error in parenthesis.
<table>
<thead>
<tr>
<th>Peptide / Sample</th>
<th>N&lt;sup&gt;a&lt;/sup&gt;</th>
<th>A&lt;sup&gt;b&lt;/sup&gt;</th>
<th>B&lt;sup&gt;c&lt;/sup&gt;</th>
<th>k&lt;sub&gt;intermediate&lt;/sub&gt;</th>
<th>C&lt;sup&gt;c&lt;/sup&gt;</th>
<th>k&lt;sub&gt;fast&lt;/sub&gt;</th>
<th>D&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>575&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>11.3 (0.7)</td>
<td>2.9</td>
<td>3.6</td>
<td>0.16 (0.05)</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>10.6 (0.3)</td>
<td>4.3</td>
<td>0.035 (0.010)</td>
<td>6.2</td>
<td>3.00 (0.73)</td>
<td></td>
</tr>
<tr>
<td>644&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>6.7 (0.7)</td>
<td>3.3</td>
<td>3.5</td>
<td>2.03 (0.70)</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>10.0* (4.9)</td>
<td>5.8</td>
<td>2.0</td>
<td>0.22 (0.15)</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>649&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>9.8 (0.2)</td>
<td>2.3</td>
<td>0.022 (0.007)</td>
<td>1.4</td>
<td>0.11 (0.03)</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>8.8 (0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>660&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>5.9 (1.1)</td>
<td>2.9</td>
<td>13.8</td>
<td>3.03 (0.72)</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>5.3 (0.4)</td>
<td>1.7</td>
<td>4.3</td>
<td>0.12 (0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>703&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>2.9 (0.1)</td>
<td>5.8</td>
<td>10.6</td>
<td>2.05 (0.61)</td>
<td>1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>3.0 (0.1)</td>
<td>1.3</td>
<td>7.0</td>
<td>2.62 (0.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>716&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>16.7 (1.1)</td>
<td>2.9</td>
<td>16.4</td>
<td>1.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>16.4 (0.3)</td>
<td></td>
<td></td>
<td>5.54 (1.82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>725&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>12.8 (1.1)</td>
<td>5.8</td>
<td>12.0</td>
<td>2.62 (0.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>12.0 (0.3)</td>
<td>5.6</td>
<td>7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>729&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>3.8 (0.3)</td>
<td>2.3</td>
<td>3.9</td>
<td>1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>3.9 (0.2)</td>
<td>2.2</td>
<td>1.7</td>
<td>3.91 (1.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>740&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>7.5 (0.2)</td>
<td>2.7</td>
<td>7.9</td>
<td>4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>7.9 (0.3)</td>
<td>2.0</td>
<td></td>
<td>5.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>758&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>2.0 (0.0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>2.2 (0.1)</td>
<td>0.4</td>
<td>0.035 (0.020)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>766&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>11.1 (0.2)</td>
<td>5.2</td>
<td>5.2</td>
<td>2.30 (0.28)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>10.7 (0.3)</td>
<td>3.0</td>
<td>0.037 (0.016)</td>
<td>5.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>792&lt;sup&gt;13&lt;/sup&gt;</td>
<td>N2</td>
<td>10.1 (0.3)</td>
<td>4.8</td>
<td>10.0</td>
<td>1.89 (0.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>10.0 (0.3)</td>
<td>3.3</td>
<td>7.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>952&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>3.3 (0.1)</td>
<td>1.9</td>
<td>4.2</td>
<td>5.09 (3.48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>4.2 (0.2)</td>
<td>2.7</td>
<td>1.6</td>
<td>3.41 (0.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1188&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>10.8 (0.3)</td>
<td>6.6</td>
<td>11.1</td>
<td>2.23 (0.74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>11.1 (0.5)</td>
<td>2.1</td>
<td>4.8</td>
<td>0.26 (0.07)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>642&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>4.2 (0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>4.1 (0.1)</td>
<td>1.3</td>
<td>0.035 (0.009)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>664&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>10.6 (0.3)</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>9.9 (0.3)</td>
<td>2.9</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>669&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>14.4 (0.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>14.6 (0.2)</td>
<td></td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>671&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>4.1 (0.1)</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>4.0 (0.1)</td>
<td>2.1</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>685&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>6.2 (0.4)</td>
<td>1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>6.2 (0.2)</td>
<td>1.4</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>699&lt;sup&gt;12&lt;/sup&gt;</td>
<td>N2</td>
<td>5.1 (0.1)</td>
<td>1.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>5.0 (0.2)</td>
<td>2.2</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>742&lt;sup&gt;14&lt;/sup&gt;</td>
<td>N2</td>
<td>6.7 (0.3)</td>
<td>3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>O2</td>
<td>6.8 (0.5)</td>
<td>3.0</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.4 (continued on following page)
<table>
<thead>
<tr>
<th>Peptide / Sample</th>
<th>N&lt;sup&gt;a&lt;/sup&gt;</th>
<th>A&lt;sup&gt;b&lt;/sup&gt;</th>
<th>B&lt;sup&gt;c&lt;/sup&gt;</th>
<th>k&lt;sub&gt;inter&lt;/sub&gt;</th>
<th>C&lt;sup&gt;c&lt;/sup&gt;</th>
<th>k&lt;sub&gt;fast&lt;/sub&gt;</th>
<th>D&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>N2 753+4</td>
<td>10.5 (1.4)</td>
<td>43</td>
<td>6.2</td>
<td>2.38 (0.44)</td>
<td>6.4</td>
<td>2.18 (0.86)</td>
<td></td>
</tr>
<tr>
<td>O2 753+3</td>
<td>9.4 (0.1)</td>
<td>1.7</td>
<td>0.018 (0.012)</td>
<td>7.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 754+3</td>
<td>6.4 (0.4)</td>
<td>2.3</td>
<td>4.1</td>
<td>3.44 (0.71)</td>
<td>7.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 754+3</td>
<td>6.9 (0.4)</td>
<td>2.8</td>
<td>4.1</td>
<td>3.99 (1.19)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 767+4</td>
<td>7.8 (0.8)</td>
<td>2.1</td>
<td>5.7</td>
<td>2.57 (0.61)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 767+4</td>
<td>7.7 (0.4)</td>
<td>2.5</td>
<td>5.2</td>
<td>3.17 (0.96)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 781+4</td>
<td>10.8 (0.1)</td>
<td>3.1</td>
<td>2.2</td>
<td>0.47 (0.04)</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 781+3</td>
<td>10.6 (0.7)</td>
<td>2.8</td>
<td>2.6</td>
<td>0.46 (0.40)</td>
<td>5.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 783+3</td>
<td>11.7 (0.8)</td>
<td>4.2</td>
<td>7.4</td>
<td>1.81 (0.42)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 783+3</td>
<td>11.4 (0.4)</td>
<td>5.0</td>
<td>6.3</td>
<td>2.64 (0.67)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 786+4</td>
<td>8.8 (0.8)</td>
<td>2.6</td>
<td>6.2</td>
<td>2.64 (0.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 786+4</td>
<td>8.5 (0.2)</td>
<td>3.1</td>
<td>5.4</td>
<td>3.00 (0.72)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 797+3</td>
<td>8.1 (0.3)</td>
<td>5.1</td>
<td>3.0</td>
<td>2.27 (0.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 797+3</td>
<td>8.1 (0.3)</td>
<td>5.5</td>
<td>2.6</td>
<td>2.58 (0.86)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 806+2</td>
<td>9.7 (0.2)</td>
<td>1.7</td>
<td>1.30 (0.076)</td>
<td>8.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 806+2</td>
<td>9.5 (0.1)</td>
<td>3.4</td>
<td>0.320 (0.055)</td>
<td>6.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 807+3</td>
<td>11.3 (0.3)</td>
<td>3.5</td>
<td>7.8</td>
<td>2.72 (0.33)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 807+3</td>
<td>11.4 (0.4)</td>
<td>3.8</td>
<td>7.6</td>
<td>4.12 (1.08)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 807+4</td>
<td>20.3 (0.4)</td>
<td>7.2</td>
<td>7.4</td>
<td>0.37 (0.11)</td>
<td>5.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 807+4</td>
<td>20.5 (0.3)</td>
<td>6.8</td>
<td>8.3</td>
<td>0.33 (0.07)</td>
<td>5.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 825+2</td>
<td>3.7 (0.1)</td>
<td>1.3</td>
<td>0.038 (0.016)</td>
<td>2.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 825+2</td>
<td>3.9 (0.1)</td>
<td>1.2</td>
<td>0.038 (0.020)</td>
<td>2.7</td>
<td>3.87 (1.41)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 834+4</td>
<td>14.2 (1.2)</td>
<td>6.2</td>
<td>8.0</td>
<td>2.39 (0.56)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 834+4</td>
<td>14.5 (1.4)</td>
<td>6.4</td>
<td>8.1</td>
<td>1.99 (0.74)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 864+2</td>
<td>1.5 (0.0)</td>
<td>0.8</td>
<td>0.086 (0.038)</td>
<td>0.7</td>
<td>3.22 (2.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 864+2</td>
<td>1.5 (0.1)</td>
<td>1.0</td>
<td>0.113 (0.035)</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 914+4</td>
<td>22.4 (1.2)</td>
<td>7.5</td>
<td>8.3</td>
<td>0.24 (0.16)</td>
<td>6.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 914+4</td>
<td>21.9 (0.8)</td>
<td>6.4</td>
<td>9.4</td>
<td>0.26 (0.12)</td>
<td>6.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 952+3</td>
<td>6.5 (0.2)</td>
<td>3.5</td>
<td>2.9</td>
<td>3.42 (1.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 952+3</td>
<td>6.6 (0.3)</td>
<td>3.7</td>
<td>2.9</td>
<td>2.21 (0.62)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N2 954+3</td>
<td>16.1 (0.6)</td>
<td>3.2</td>
<td>12.9</td>
<td>2.94 (0.57)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>O2 954+3</td>
<td>15.9 (0.4)</td>
<td>3.7</td>
<td>0.055 (0.025)</td>
<td>12.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> - This parameter (N) was forced to 10 for this peptide to achieve reasonable values for the other parameters. N represents the total number of amide hydrogens exchanged.

<sup>b</sup> – The parameters A and D represent the number of amide hydrogens exchanging at very slow or very fast rates respectively.

<sup>c</sup> – The parameters B and C represent the number of amide hydrogens exchanging at the rates k<sub>intermediate</sub> and k<sub>fast</sub>, respectively.
Figure 5.9. Data displaying similar exchange for peptide 664+4 due to large standard deviations. Closed circles (●) represent N$_2$-incubated (active CbbT) and open circles (○) represent O$_2$-incubated (inactive CbbT). The non-linear fit results are shown as a dashed line for the N$_2$ data and as a solid line for the O$_2$ data. Plots display the increase in mass observed for the peptide versus time on a linear scale (left hand side), as well as a logarithmic scale (right hand side). Plots were constructed as described in the text (5.2.6).

Though the data demonstrate differences between the two forms of CbbT, further interpretation of the results requires sequence determination for the peptides. Mapping the results of HD exchange to a static structure would provide insight into which regions of the protein have been modified upon the oxygen inactivation of CbbT. We have mentioned that the assignment of CbbT peptides to specific amino acid sequences based on theoretically derived pepsin cleavage products resulted in a low number of matches
Assignment of the above mentioned 38 peptides produced possible matches for 28, within a broad error range from -22 to 115 ppm. Of these 28 matches, 23 contained multiple putative sequences within a similar range of error and could not be distinguished. The remaining 5 peptides matched a single sequence within an error range from -22 to 51 ppm and two of these displayed differences in HD exchange rates upon O₂ inactivation of CbbT (766<sup>+</sup> and 952<sup>+</sup>). Four of the peptides were mapped to the Pyr domain of CbbT and the fifth to the C-terminal domain, all in relative proximity to one another as observed in the modeled structure of CbbT (Figure 5.10). The locations of these five peptides are discussed following figure 5.10.
**Figure 5.10.** Mapping peptides to the CbbT structure. (A) Coil model displaying the locations of the 5 peptides with high confidence assignments. The model displays a single subunit of CbbT with two molecules of ThDP in the incomplete active sites. The assigned peptides are labeled with lines indicating both ends of each peptide. (B) Sequence alignment of CbbT with the *E. coli* transketolase TkI1. Secondary structural elements obtained from the TkI1 structure (pdb 1QGD) are indicated above the TkI1 sequence. Matching residues are boxed with a shaded background, and assigned peptides are indicated below the CbbT sequence by shaded bars. Only a relevant stretch of sequence has been shown for convenience. Peptides predicted to change upon oxygen inactivation are 766+3 and 725+4. Peptides 671+2, 807+4, and 864+2 displayed no difference. All peptides were mapped to the Pyr domain with the exception of 766+3, which was found in the C-terminal domain.
Figure 5.10
Peptide 864+2 would consist of a 16 amino acid fragment in the Pyr domain, including the invariant residue Glu 408. This residue was found deeply buried within the molecule and interacts with ThDP through hydrogen bonding (7). In support of this assignment is the low amount of net exchange detected for this peptide. With only 1.5 amides exchanged, this peptide demonstrated the lowest net exchange of all the peptides examined. Additionally, the rate constants calculated for this peptide indicated that 1 amide from each sample (N$_2$- and O$_2$-CbbT) exchanged at the slower rate of approximately 0.1 min$^{-1}$, with only fractional amounts that exchanged at faster rates (Table 5.4).

The mass assignment of 807+$^+$ locates this 31 amino acid peptide on the surface of the Pyr domain, comprising half of an $\alpha$-helix with 7 residues, followed by a 6 residue loop and then a 14 residue $\alpha$-helix as part of a helix-turn-helix. HD exchange in 807+$^+$ was measured for 20 of the 29 exchangeable amides (minus 2 Pro) in both CbbT samples. The rate constants for this peptide demonstrated that in both O$_2$- and N$_2$-CbbT, 7 hydrogens exchanged at an intermediate rate (0.021 min$^{-1}$), 7 – 8 exchanged at a fast rate (0.37 min$^{-1}$), and the remaining 5 – 6 exchanged at a very fast rate.

The 13 amino acid peptide 671+$^+$ was predicted to overlap within the 27 amino acid peptide 725+$^+$, in the Pyr domain. Since 725+$^+$ exhibited different (though listed as possible) exchange whereas 671+$^+$ remained constant between the two forms of CbbT, these two peptides exemplify the advantage of sequence overlap. The overlap between these two fragments reduces the number of amino acids implicated to change in 725+$^+$ from 27 amino acids to either the 7 N-terminal amino acids and/or the 7 C-terminal
amino acids of this peptide. The occurrence of such peptide overlap would be important
to improve the resolution of mapping HD exchange rates that demonstrate change upon
oxygen inactivation of CbbT.

Peptide 766+3 was predicted to span two small helices in the C-terminal domain.
The C-terminal domain reportedly does not directly play a role in the subunit interface or
active site and its function remains unresolved (7). Exchange for 11 of the 18
exchangeable amides was detected in this 21 amino acid fragment. In the N2 sample, 5
amides exchanged at 0.023 min\(^{-1}\) and 6 at 2.3 min\(^{-1}\), whereas in the O2 sample 3 amides
exchanged at 0.037 min\(^{-1}\) and 8 at 1.9 min\(^{-1}\), representing nearly a 100-fold increase in
exchange rates for 2 of the amides in the O2 sample.

5.4 CONCLUSIONS

HD exchange by MS was investigated as a technique to probe structural
differences between active- and oxygen-inactivated forms of CbbT. Methods of sample
preparation and data collection have been defined. The importance of LC separation to
resolve peptides of overlapping \(m/z\) was observed. The gradient used in the collection of
this data was optimized for speed, though many groups have reported adequate exchange
data collected within 20 – 30 min of initiating the digest (135, 152, 153). It is therefore
recommended that extra time be added to decrease the slope of the elution gradient and
improve separation, finding a median between optimal speed and separation.
Reproducibility in the chromatography phase of data collection would also allow one to
fully capitalize on the Midas analysis software, greatly reducing the data reduction phase.
As many as 215 clearly resolved peptides were observed for a single CbbT peptic digest by LC-MS, though a reduced consistency of these ions limited the data analysis to 118 peptides. Much of this inconsistency is believed to be a result of the HPLC system used with the pre-column flow splitter, in effect converting it into a low flow system. On the positive side, this effectively reduced the dead volume of the HPLC and allowed for the rapid elution of peptides. Negatively, this led to variation in system back pressure, which resulted in a variable flow rate that particularly affected the stability of the spray. Such variations could be eliminated through the use of a nano-flow chromatography system. Recently, a set-up utilizing a low flow system demonstrated an increased sensitivity in HD exchange experiments by MS (137). Consistency between samples could also be improved by performing the exchange reaction in one sample from which aliquots are removed and quenched at separate time points.

Since much of the HD exchange occurred rapidly in the peptides examined, collecting more data points within the first 3 minutes of exchange would allow for more accurate rate calculations of the fast exchanging amides. The current method of sample preparation limits the shortest accurate time point to approximately 30 seconds. Calculations for faster exchange rates have been achieved collecting data at shorter time points through rapid stopped-flow quenching (154). Other possibilities include slowing down the rate of exchange to a readily measurable time scale by a slight reduction in pH or temperature (151). These would represent valid ways to slow down HD exchange as long as major structural changes are not induced by the decrease in pH or temperature. To aid in the calculation of the very slow exchange rates, the observation window should
be widened by the inclusion of exchange time points up to 8 hours. In order to keep the data set both manageable and complete in such a large window, single data points at various times could be collected rather than fewer time points in triplicate.

Identification of the analyzed peptides is a requirement for the correlation of the exchange results to the tertiary structure of the enzyme. Though confident identification of the peptides was not possible by molecular weight alone, we were still able to evaluate changes between the two forms of CbbT. Rates determined for the 5 peptides that could be matched with higher confidence corresponded with their location in the modeled structure. Though only a subset of the peptides analyzed, the locations of these 5 included amino acids were found to be deeply buried near the active site, in subunit-subunit interfaces, as well as surface accessible. This is supportive that the designed method will allow coverage of CbbT throughout the enzyme and is not limited by tertiary structure, though only a sequence analysis of the peptides can confirm this. The changes observed in both the mapped peptides (725<sup>14</sup>, 766<sup>3</sup>) were derived from data sets lacking defined standard deviations, but more confident differences were observed for other peptides. For the peptides that displayed more confident differences in HD exchange rates, similar and lesser differences have been reported as significant in identifying locations of change and flexibility in the extracellular regulated protein kinase-2 (ERK2) upon phosphorylation, as detected by HD exchange using MS (155). It is therefore hopeful that with sequence information these peptides will reveal regions of CbbT sensitive to oxygen inactivation. Though a more complete data set would be desired, the current kinetic data gathered for the 38 peptides can be used to build a working
hypothesis.

In addition to information regarding differences between the N₂ and O₂ forms of CbbT, the exchange rates will also provide information concerning the flexibility of transketolase. The degree of flexibility in transketolase is currently an unresolved issue, with evidence in support of both rigid and dynamically adaptive structures in solution. X-ray structure analysis of yeast transketolase crystals containing a reaction intermediate in the active site suggested that no structural changes occur during formation of the intermediate. It was also noted that the intermediate bound to both active sites equally, suggesting dynamically equivalent active sites (18). In contrast, kinetic (122) and dynamic modeling (156) studies suggest non-equivalent dynamics during substrate binding at the two chemically equivalent active sites. Crystal structure analysis of the homologous ThDP-dependent enzyme pyruvate dehydrogenase was in support of the kinetic and molecular modeling analysis. Based on this structure, the authors proposed a flip-flop mechanism in which the two active sites alternately accept and catalyze substrate (123). Whether a similar mechanism applies to transketolase remains to be seen. Calculations of HD exchange dynamics by MS would allow for the direct measurement of transketolase flexibility in solution.
CHAPTER 6

SUMMARY AND FUTURE DIRECTIONS

Both of the transketolase genes from *Rhodobacter sphaeroides*, *cbbT* and *tkt*, were cloned into separate pET-11a expression vectors. The separate induction of these genes, in the salt-inducible *E. coli* strain GJ1158 under microaerobic conditions, allowed for the accumulation of large amounts of soluble, active protein. Subsequent purification and biochemical analyses allowed, for the first time, comparative studies to be undertaken of transketolase isozymes from the same organism. Despite the high degree of sequence homology, these two proteins displayed different biochemical properties. Differences were noted in affinities for the common substrates Xu5P and R5P, along with distinct pH optima, and surface charge distribution. It appears that these proteins evolved to operate at high efficiency under different environmental conditions. Therefore it was not surprising that activity assays for *in vitro* transketolase in *R. sphaeroides* clarified extracts, under pH and substrate conditions optimal for CbbT, displayed little apparent contribution by Tkt.

We also identified several factors which led to the stabilization of both CbbT and Tkt, most notably the presence of sulfate and the absence of oxygen. Inactivation by oxygen had previously been reported for the *E. coli* transketolase, TktA, which had been covalently linked to a support matrix (47). Support matrices are often utilized to improve
enzyme stability by preventing denaturation and allow for a convenient method of separating synthesized substrates from the utilized enzymes (157). Despite the use of a support matrix, TktA was still inactivated in the presence of air; in this study it was noted that activity could be stabilized by storing the column under gaseous N₂ or with a high concentration of reductant (0.5 % BME). We demonstrated that CbbT displayed a similar susceptibility to oxygen inactivation, both after its synthesis in *E. coli* and then after storage of the purified recombinant protein. To further our understanding of this inactivation, we characterized various properties of the oxygen-inactivated form of CbbT. Differences in the intrinsic fluorescence spectra, between the O₂-inactive and N₂-active forms, were suggestive of the intrinsic chromophore shifting towards a more polar environment (Sec 3.3.3). Fluorescence was also used to generate urea unfolding curves for both forms of CbbT; these results suggested that the O₂-inactivated CbbT had adopted a slightly more stable configuration (Sec. 3.3.4). Differences in far UV CD spectra, around 222 nm, were suggestive of an increase in secondary structure in the inactivated enzyme (Sec. 3.3.3). Since the inactivated sample used for CD still retained 30 % the activity of the active sample, collection of CD data using a sample which has reached a lower level of activity would ensure that no other differences are observed in this region. Collecting CD data in the near-UV range could also be performed to reinforce the fluorometry data, in which Trp is the primary intrinsic fluorophore, since CD is particularly sensitive for this residue around 280-300 nm. Though our application of these spectral techniques has been limited to endpoint samples, either fully active or mostly inactivated, they could be applied to monitor the inactivation directly over time.
This would provide some advantage over assaying enzyme activity versus time, in which the repeated removal of sample by syringe seemingly introduced air into the incubation vial. For vials that had many aliquots sampled during the time course, the resulting plots of activity versus time displayed an unexpected inactivation of the N$_2$-active sample. With direct spectrometric measurements, data could be collected without the risk of introducing air, utilizing gas-tight sealed cuvettes. The resulting plots of change in spectrometric property (CD or fluorescence) versus time could then be used to investigate rates of oxygen inactivation under different conditions. Also, the effect of oxygen on the apoenzyme should be investigated to note any differences in the rates of inactivation compared to the holoenzyme. Should inactivation of the apoenzyme be similar to holoenzyme, ThDP involvement in the mechanism of inactivation might be ruled out.

To explore the possibility that Cys160 of CbbT, the only cysteine in the active site, was responsible for oxygen inactivation, this residue was targeted for site-directed mutagenesis. Altering this residue to Ala, Ser, Asp, or Glu did not affect oxygen-mediated inactivation of CbbT. However, the contribution of this residue in ThDP binding was clearly established. ThDP affinity measurements conducted at 25 °C demonstrated the similarity of the C160D mutant with wild-type enzyme at 3 separate pH values; however the other mutant proteins displayed decreased affinities for ThDP. Interestingly, the time course of ThDP reconstitution for the C160D enzyme suggested that this protein was not fully reconstituted at the end of the purification (performed at 0-4 °C), existing as a mixture of apo- and holoenzyme. Therefore, after incubating this enzyme at 25 °C in the presence of ThDP and Mg$^{++}$ cofactors, activity was increased as
more holoenzyme was formed (Figure 4.9). The increase in activity was very slow for the completely desalted sample, with the N₂-stored sample continuing to increase in activity even after 90 h at 25 °C. In the reconstitution assays used for ThDP affinity measurements, all mutant and wild type enzymes were incubated in the presence of a residual amount of ammonium sulfate and did not exhibit such a slow increase in activity at any of the pH values investigated. It is therefore possible that the observed stability afforded by sulfate is related to ThDP binding; in particular, sulfate appeared to affect the rate of binding. It is interesting to note that 2 sulfate ions are present in the solved structure of the *E. coli* transketolase (TktA) pdb (1QGD). No explanation regarding the presence of these sulfate ions was presented by the authors of the pdb since a supporting publication for this structure has yet to be published (9). The location of these ions in the pdb reveals that one ion binds in each of the substrate channels, apparently interacting with the side chains responsible for binding to the phosphate of the substrate (Figure 6.1) (59). Though too far away to interact directly with ThDP, it is possible that by binding to the substrate channel, the sulfate ion provides stability to the surrounding region and/or blocks ThDP from diffusing out of the active site. This could also explain the inhibitory nature of sulfate as well as phosphate, reported for the yeast transketolase (70). We have also observed inhibition by phosphate for CbbT, but not sulfate; in CbbT sulfate may be easily displaced by substrate, or perhaps it binds in a different location. To investigate the effect of sulfate on ThDP binding, it would be desirable to establish *Kₘ* or *Kₛ* values at different concentrations of sulfate, using wild-type CbbT. If any effect on ThDP binding by sulfate is observed, this could also be investigated in Tkt from *R. sphaeroides*. 
**Figure 6.1.** Sulfate ions in the *E. coli* TktA structure solved by crystallography.  

(A) Molecular surface model emphasizing the accessibility of ThDP at the bottom of the substrate channel and displaying the location of the sulfate ion.  

(B) Stick representation displaying residues located within 6 Å of the sulfate ion in addition to the ThPD cofactor; the four amino acids involved in substrate binding are labeled based on the *E. coli* numbering. The figures were created using the pdb coordinate 1QGD.
A method to measure the rates of hydrogen/deuterium (HD) exchange for the backbone amides of CbbT by mass spectrometry was designed. Evaluation of the data showed promise for detecting conformational differences between N₂-active and O₂-inactivated forms of CbbT. Essential to the data analysis is the assignment of the pepsin generated peptides to their respective amino acid sequences. Identification of the peptides by mass alone has proven to be insufficient, producing a low number of high confidence matches. It is possible that this low number of matches has been complicated by post-purification modifications of CbbT. We have shown that denatured purified CbbT is susceptible to degradation by heating upon preparation for SDS-PAGE, much in the manner of acid-induced Asp-Pro bond cleavage. It is therefore possible that similar cleavage events occur upon formation of the peptic peptides, which is performed under acidic conditions. Alternatively, cleavage might occur during desalting of the peptides, which are eluted by MeOH, exposing the peptides to both low pH and a denaturing solvent. Such cleavage events are not predicted by the software used to generate the theoretical list of peptides. In order to absolutely identify the peptides, a MS/MS approach should be employed in future studies. Standard instrumentation to achieve such sequencing typically involves a capillary HPLC and a quadrupole-time-of-flight mass spectrometer (Q-TOF MS) capable of tandem MS/MS experiments. In tandem MS/MS, the first MS event isolates the peptide of unknown sequence, this ion is then fragmented with the use of a collision gas (argon), and the resulting fragments are analyzed by the second MS event. The fragmentation pattern can then be analyzed by software (i.e. Sequest, Thermo Electron Corp.) to elucidate the amino acid sequence of that peptide.
Further enhancement of the mass spectral data will be provided by the forthcoming crystal structure of CbbT. A 50 mg sample of purified CbbT was sent to the laboratory of Professor David Eisenberg (UCLA-DOE Institute for Genomics and Proteomics, California). The crystallization and data collection are already completed and the structure is currently in the refinement phase of model building. With the advent of a CbbT structure, we would no longer be forced to rely on theoretical homology models to map the HD exchange data for analysis. In addition to identifying regions that are affected upon oxygen inactivation, the HD exchange experiments have the potential to shed light on the currently debated topic of flexibility in transketolase.
REFERENCES


43. **Crestia, D., C. Demuynck, and J. Bolte.** 2004. Transketolase and fructose-1,6-bis-phosphate aldolase, complementary tools for access to new ulosonic acid analogues. Tetrahedron **60**:2417-2425.


